



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAACCTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

NcoI (566)

**BspEI (558)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGCTCCGGACACCATGGAGAACAACAAAACCTCAGTGGATTCAA  
111 M E N N K T S V D S K  
601 ATCCATTAATAATTTGAAGTAAAGACCATACATGGGAGCAAGTCACTGGACTCTGGGATCTATCTGGACAGTAGTTACAAAATGGATTATCCTGAAATG  
111 S I N N F E V K T I H G S K S V D S G I Y L D S S Y K M D Y P E M  
701 GGCATATGCATAATAATTAATAAAGAAGTCCATAAGAGCACTGGAATGTCATCTCGCTCTGGTACGGATGTGGACGCAGCCAACCTCAGAGAGACAT  
141 G I C I I I N N K N F H K S T G M S S R S G T D V D A A N L R E T  
801 TCATGGGCTGAAATACCAAGTCAGGAATAAAAATGATCTTACTCGTGAAGACATTTTGAATTAATGGATAGTGTTCCTAAGGAAGATCATAGCAAAAG  
171 F M G L K Y Q V R N K N D L T R E D I L E L M D S V S K E D H S K R

NcoI (927)

901 GAGCAGCTTTGTGTGTGATTCTAAGCCATGGTGTGATGAGGGGTCATTTATGGGACAAATGGGCTGTTGAACTGAAAAAGTTGACTAGCTTCTTCAGA  
111 S S F V C V I L S H G D E G V I Y G T N G P V E L K K L T S F F R  
1001 GCGGACTACTGCCGAGTCTGACTGAAAGCCGAAACTCTTCATCATTAGGCCTGCCGGGTACGGAGCTGGACTGTGGCATTGAGACAGACAGTGGGA  
141 G D Y C R S L T G K P K L F I I Q A C R G T E L D C G I E T D S G  
1101 CTGATGAGGAGATGGCTTGCCAGAAGATACCGGTGGAGGCTGACTTCTGTATGCTTACTCTACAGCACCTGGTACTATTCTGGAGAAATCAAAGGA  
171 T D E E M A C Q K I P V E A D F L Y A Y S T A P G Y Y S W R N S K D  
1201 CGGGTCGTGGTTCATCCAGTCCCTTTCAGCATGCTGAAGCTGTACGCGCACAAAGCTAGAATTTATGCACATTCTCACTCGCGTTAACAGGAAGTGGCA  
211 G S W F I Q S L C S M L K L Y A H K L E F M H I L T R V N R K V A  
1301 ACGGAATTCGAGTCTTCTCCCTGGACTCCACTTCCACGCAAGAAACAGATCCCGTGTATTGTGTCATGCTCACGAAAGAACTGACTTTTATCACT  
241 T E F E S F S L D S T F H A K K Q I P C I V S M L T K E L Y F Y H

**NheI (1420)**

1401 AGAGGAATGATTGGGGGTGGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTAT  
271 •  
1501 TTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACAACAACAATTGCATTTCATTTTATGTTTCAGGTTTCAG

1601 GGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGC

1701 CTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAG

1801 TTTAAGATATAGTGATTTTTCCAAGGTTTGAAGTCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATGAAAAATATTC

1901 AGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGT

2001 TGGACTTAGGGAACAAAGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCA  
141 • N R T Y K L P I L E E

2101 ATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCA  
121 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 V V

2201 CCCTGATGATCTGTCCACCTCATCAGATAGGGGTCCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAAT  
91 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 A I

2301 GGCTTACAGCAGACAGTACCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGTCTTGGTCTGATGGCCGCCCGACATGGTGC  
61 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 H H

2401 TTGTTGCTCTCATAGAGCATGGTGTCTCTCAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTAT  
28 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

2501 AGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACCTAAACGAG

2601 CTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGGAGTGTGTACGACATTTTGGAAAGTCCGTTGATTTACTAG

2701 TCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATC

2801 ATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAAGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGT

2901 CATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAA

3001 AGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTAT

3101 GTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCGTAAAAGGCCGCGTGTGCGGCTTTTTCCATAGGCTCCG

3201 CCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTC  
3301 GTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATC  
3401 TCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC  
3501 CAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG  
3601 TGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCA  
3701 AACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG  
3801 GTCTGACGCTCAGTGGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTT  
3901 CATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCC  
4001 CAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA