



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

**BspHI (560)**

**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCATCAGTGGGGACTACAGATTCTCCCTGCCTGAGCCT  
1 M S W G L Q I L P C L S L  
601 AATCCTTCTTCTTTGGAACCAAGTGCCAGGGCTTGAGGGTCAAGAGTTCGATTTGGGTCTTGCCAAGTGACAGGGGTGGTTCTCCAGAAGTGTGGGAG  
13 I L L L W N Q V P G L E G Q E F R F G S C Q V T G V V L P E L W E  
701 GCCTTCTGGACTGTGAAGAACACTGTGCAAACCTCAGGATGACATCACAAAGCATCCGGCTGTTGAAGCCGACAGTTCTGCGGAATGTCTCGGGTCTGAGA  
47 A F W T V K N T V Q T Q D D I T S I R L L K P Q V L R N V S G A E  
801 GCTGTTACCTGCCCACAGCCTGCTGAAGTCTACTTGAACACTGTTTCAAGAACTACCACAGCAAATAGCAAATTAAGGTCTTGAGGTCATTCTC  
80 S C Y L A H S L L K F Y L N T V F K N Y H S K I A K F K V L R S F S  
901 CACTCTGGCCAACTTATAGTCATCATGTCAACTACAGCCAGTAAGGACAATCCATGCTTCCATTAGTGAGAGTGACACCAGCGGTTTTTG  
113 T L A N N F I V I M S Q L Q P S K D N S M L P I S E S A H Q R F L  
1001 CTGTTCCGAGAGCATTCAAACAGTTGGATACAGAAGTCGCTTGGTAAAGCCTTTGGGAAAGTGGACATTCCTGACCTGGATGCAGAAATCTACC  
147 L F R R A F K Q L D T E V A L V K A F G E V D I L L T W M Q K F Y

**NheI (1130)**

1101 ATCTCTGACTGCTGATTGGATAACTTCTCCTGCTAGTGGCCAGACATGATAAGATACATTGATGAGTTGGACAAACCACAACCTAGAATGCAGTGAAAAA  
180 H L •  
1201 AATGCTTATTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAAACAACAATTGCATTCATTTTATGTT  
1301 TCAGTTTCAGGGGAGGTGTGGGAGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAATACAGCATAGCAAACTTTAACCTC  
1401 CAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTC  
1501 TTTTCATGGAGTTTAAAGATATAGTGATTTTCCCAAGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATGCAGTACCTCCCACATTCCCTTTTAG  
1601 TAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCCA  
1701 GTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGAT  
141 • N R T Y K L P I  
1801 GAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAG  
132 L 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  
1901 GGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCAAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACC  
98 P S 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  
2001 CAATGGCAATGGCTTACAGCAGACAGTACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCCC  
65 I 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  
2101 GACATGGTCTTGTGCTCATAGAGCATGGTGTCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTG  
32 V 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 GCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTC  
2301 ACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTT  
2401 GATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCA  
2501 AAACCGCATCATCATGGAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGCATAATGCCAGGCGGGC  
2601 CATTTACCGTCATTGACGTCAATAGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGAC  
2701 GTCATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTGTTGGGCGGTGAGCCAGGCGGGCCATTAC  
2801 CGTAAGTTATGTAACGCTGACGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCC  
2901 ATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGGCGTTTTCCCCTGG  
3001 AAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGC  
3101 TGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGCTTATCCGGTAACTATC

3201 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT  
3301 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCT  
3401 TGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCT  
3501 TTTCTACGGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGGGCCGCAATAAAATA  
3601 TCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAT  
3701 AGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA