



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

**BspHI (560)**

**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCATCATGAAGTGGCACATGATCATCTCGGGGCTTATCGTAGT  
1 M N W H M I I S G L I V V  
601 AGTGATCAAAGTTGTTGGAATGACCTTTTTCTGCTGATTTCCACAGGTTTTGGCAAAGTAATGATGGCTTCGTCGCCACGAGAGCTACGGAACC  
13 V I K V V G M T F F L L Y F P Q V F G K S N D G F V P T E S Y G T  
701 ACTAGTGTGAGAATGTCTCACAGATCTTTGGGAGAAATGACGAAAGTACCATGCCTACAAGGAGCTATGGAACAGTCTGTCCAGAAACTGGGATTTTC  
47 T S V Q N V S Q I F G R N D E S T M P T R S Y G T V C P R N W D F  
801 ACCAAGGAAAATGCTTTTTCTCTCTCTCCGAATCACCTTGAAAGACAGCATGGATTATTGTGCAACACAAGGATCCACACTGGCAATTGTCAACAC  
80 H Q G K C F F F S F S E S P W K D S M D Y C A T Q G S T L A I V N T  
901 TCCAGAGAAACTGAAGTATCTTCAGGACATAGCTGGTATTGAGAATTACTTTATTGGTTGGTACGTCAGCCTGGAGAGAAAAAGTGGCGCTGGATCAAC  
113 P E K L K Y L Q D I A G I E N Y F I G L V R Q P G E K K W R W I N  
1001 AACTCTGTGTTCAATGGCAATGTACCAATCAGGACCAAGAACTTCGACTGTGTCACTATAGGTCGACGAAGACATATGATGTCATCATGTGAAGTCA  
147 N S V F N G N V T N Q D Q N F D C V T I G L T K T Y D A A S C E V

**NheI (1135)**

1101 GCTATCGTGGATCTGCGAAATGAATGCCAAATGAGCTAGTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAAATGCAGTG  
180 S Y R W I C E M N A K •  
1201 AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCATTTT  
1301 ATGTTTCAGGTTCAAGGGGAGGTGTGGGAGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAATACAGCATAGCAAACTTTA  
1401 ACCTCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCA  
1501 CCTTCTTCATGGAGTTAAGATATAGTGTATTTCCCAAGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCAGTACCTCCCACATTCCCTT  
1601 TTTAGTAAATATTCAGAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATC  
1701 CCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGG  
141 • N R T Y K L  
1801 GGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGC  
133 P I 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  
1901 CACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGC  
100 C 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  
2001 AGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCC  
67 S G 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  
2101 GCCCGACATGGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCA  
33 A G 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 TGGTGGCCCTCTATAGTGAAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGAC  
0  
2301 GGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTC  
2401 CCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCTGAGTCAAACCGCTATCCACGCCATTGATGTAC  
2501 TGCCAAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGG  
2601 CGGGCCATTTACCGTCATTGACGTCAATAGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCCA  
2701 TTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCA  
2801 TTTACCGTAAAGTTATGTAACGCTGCAGGTTAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT  
2901 TTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCGACAGGACTATAAAGATACAGGCGTTTCCC  
3001 CCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGCT  
3101 CACGCTGAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAA

3201 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA  
3301 GAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA  
3401 GCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTT  
3501 GATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGGGCCGCAATA  
3601 AAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGC  
3701 AAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA