



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspHI (560)

**AgeI (552)**

501 TCTGTTCTGCGCAGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCATCATGATGATGATCTCAGAAATACCCAGCCAATCTCTGGA  
1 M M D L R N T P A K S L D  
601 CAAGTTCATTGAAGACTATCTTTGCCAGACAGTGTTCGCCATGCAAAATCAACCATGCCATTGACATCATCTGTGGTTCTGAAGGAAAGGTGCTTC  
13 K F I E D Y L L P D T C F R M Q I N H A I D I I C G F L K E R C F  
701 CGAGGTAGCTCCTACCCTGTGTGTGTGCAAGGTGGTAAAGGGTGGCTCCTCAGGCAAGGGCACCACCTCAGAGGCCGATCTGACGCTGACCTGGTTG  
47 R G S S Y P V C V S K V V K G G S S G K G T T L R G R S D A D L V  
801 TCTTCTCAGTCTCTACCACCTTTTCAGGATCAGTAAATCGCCGGGAGAGTTCATCCAGGAAATAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAG  
80 V F L S P L T T F Q D Q L N R R G E F I Q E I R R Q L E A C Q R E R  
901 AGCATTTCCTGGAAGTTGAGGTCCAGGCTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCTGACTGAGTTCGCTCCAGCTCGGGGAGGGGTGGAG  
113 A F S V K F E V Q A P R W G N P R A L S F V L S S L Q L G E G V E  
1001 TTCGATGTGCTGCCTGCCTTTGATGCCCTGGTTCAGTTGACTGGCGGTATAAACCTAACCCCAAATCTATGTCAAGCTCATCGAGGAGTGCACCGACC  
147 F D V L P A F D A L G Q L T G G Y K P N P Q I Y V K L I E E C T D  
1101 TGCAGAAAGAGGGCGAGTTCTCCACCTGCTTACAGAACTACAGAGAGACTTCTGAAGCAGCGCCCCAACAGCTCAAGAGCCTCATCCGCCTAGTCAA  
180 L Q K E G E F S T C F T E L Q R D F L K Q R P T K L K S L I R L V K  
1201 GCCTGGTACCAAAATGTAAGAAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAAA  
213 H W Y Q N C K K K L G K L P P Q Y A L E L L T V Y A W E R G S M K  
1301 ACACATTTCAACACAGCCAGGGATTTCCGACGGTCTTGAATTAGTCATAAACTACCAGCAACTCTGCATCTACTGGACAAAGTATTATGACTTTAAAA  
247 T H F N T A Q G F R T V L E L V I N Y Q Q L C I Y W T K Y Y D F K  
1401 ACCCATTATTGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCTGTGATCCTGGACCGGGGACCTACAGGAACTTGGTGGTGGAGACCC  
280 N P I I E K Y L R R Q L T K P R P V I L D P A D P T G N L G G G D P  
1501 AAAGGGTGGAGGCAGCTGGCACAAGAGGCTGAGGCCTGGCTGAATTACCCATGCTTTAAGAATTGGGATGGTCCCCAGTGAAGTCTGGATTCTGCTG  
313 K G W R Q L A Q E A E A W L N Y P C F K N W D G S P V S S W I L L  
1601 GCTGAAAGCAACAGTGCAGACGATGAGACCCGACGATCCAGGAGGTATCAGAAATATGTTACATTGGAACACATGAGTACCCTATTCTCTCATAGAC  
347 A E S N S A D D E T D D P R R Y Q K Y G Y I G T H E Y P H F S H R

**NheI (1778)**

1701 CCAGCACACTCCAGGCAGCATCCACCCACAGGCAGAAGAGGACTGGACCTGCACCATCCTCTGAATGCCAGTGCATCGTAGCTGGCCAGACATGATAA  
380 P S T L Q A A S T P Q A E E D W T C T I L •  
1801 GATACATTGATGAGTTTGACAAACCACAACACTAGAATGCAGTGAAGGAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTAT  
1901 AAGCTGCAATAACAAGTTAACAACAACAAATGCATTCAATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACTCTAC  
2001 AAATGTGGTATGGAATTTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGC  
2101 ATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGAGGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTGCTT  
2201 CATTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAAATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTT  
2301 TTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGAACCTTTAATAGAAATTTGACAG  
2401 CAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTATCTCAATGAGCACA  
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  
2501 AAGCAGTCAGGAGCATAGTCAGAGATGAGTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGA  
114 F C 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  
2601 CAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAAT  
81 A V 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  
2701 GTGACAGCAGAGATGATCTCCCGTCTGGTCTGATGGCCGCCGACATGGTGGTGTGTCCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACC  
48 H V A S I I 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  
BspHI (2840)  
2801 TCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGTCTTTCATGATGGCCCTCTATAGTGAAGTCTATTATACTATGCCGATATACTATGCCGATGAT  
14 E V L E L D Q Q S I N F T K M  
2901 TAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCCTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCCA  
3001 TTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATGACGTCATGGGGTGGAGACTTGG  
3101 AAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCATGTAATAGCGATGACTAATACGTAGATGACTGCCAA

3201 GTAGGAAAGTCCCATAAGGTCATGTA  
3301 GATGTA  
3401 CGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAAITAAAGAACATGTGAGCAAAAGGCC  
3501 AGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG  
3601 GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCCGTTACCGGATACCTG  
3701 TCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGC  
3801 ACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC  
3901 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATC  
4001 TCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACACCGCTGGTAGCGGTGTTTTTTGTTTGAAGC  
4101 AGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGAACGAAAACACAGTTAAGGGATTTT  
4201 GGTGCTGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACTAA  
4301 CATACGCTCTCCATCAAACAAACGAAACAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA