



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTGCTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCAGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCATCATGATGATGATCTCAGAAATACCCAGCCAATCTCTGGA  
1 M M D L R N T P A K S L D  
601 CAAGTTCATTGAAGACTATCTTTGCCAGACAGTGTTCGCCATGCAAAATCAACCATGCCATTGACATCATCTGTGGGTTCTGAAGGAAAGGTGCTTC  
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 CGAGGTAGCTCCTACCCTGTGTGTGTGCAAGGTGGTAAAGGTGGCTCCTCAGGCAAGGGCACCACCTCAGAGGCCGATCTGACGCTGACCTGGTTG  
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 TCTTCCTCAGTCTCTCACCACCTTTTCCAGTACAGTAAATCGCCGGGAGAGTTCATCCAGGAAATAGGAGACAGCTGGAAGCTGTCAAAGAGAGAG  
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 AGCATTTCCTGGAAGTTGAGGTCCAGGCTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCTGACTGAGTTCCGCTCCAGCTCGGGGAGGGGTGGAG  
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 TGCAGAAAGAGGGCGAGTTTCCACCTGCTTACAGAACTACAGAGAGACTTCTGAAGCAGCGCCCCAACAGCTCAAGAGCCTCATCCGCCTAGTCAA  
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 ACACATTTCAACACAGCCAGGGATTTCGGACGGTCTTGAATTAGTCATAAACTACCAGCAACTCTGCATCTACTGGACAAAGTATTATGACTTTAAAA  
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 ACCCATTATTGAAAAGTACCTGAGAAGCAGCTCACGAAACCCAGGCTGTGATCCTGGACCGGGGACCTACAGGAACTTGGTGGTGGAGACCC  
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 AAAGGTTGGAGGCAGCTGGCACAAGAGGCTGAGGCCTGGCTGAATTACCCATGCTTTAAGAATTGGGATGGTCCCCAGTGAAGCTCTGGATTCTGCTG  
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

NheI (1673)

1601 GTGAGACCTCCTGCTTCCCTGCCATTCCCTGCCCTCTCCATGAAGCTTGGAGCATATAGCTGGAGAGCTAGCTGGCCAGACATGATAAGATAC  
347 V R P P A S S L P F I P A P L H E A •  
1701 ATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAGAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCT  
1801 GCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAATG  
1901 TGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAG  
2001 GGGCTGTTGCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGAGTAAAGATATAGTGTATTTTCCCAAGTTTGAAGTACTGCTTCTCATT  
2101 CTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTAT  
2201 TAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTAAATAGAAATTGGACAGCAAGA  
2301 AAGCGAGCTTCTAGCTTTAGTCTCGGTGACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAAGCA  
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  
2401 GTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCC  
113 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 A  
2501 ACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCAATGTAGGCCTCAATGTGGA  
79 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 H V  
2601 CAGCAGAGATGATCTCCCGAGTCTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGCTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCAC  
46 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 E V  
BspHI (2735)  
2701 CAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCTGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAAT  
13 L E L D Q Q S I N F T K M  
2801 GTCAAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTGTC  
2901 GTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATC  
3001 CCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGG  
3101 AAAGTCCCATAAGGTCATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACTTGTATG

3201 ACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCA  
3301 ATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAA  
3401 AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC  
3501 GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGC  
3601 CTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAA  
3701 CCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA  
3801 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGC  
3901 TCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAG  
4001 ATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACTCACGTTAAGGGATTTTGGTCA  
4101 TGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATAC  
4201 GCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA