



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

**BspHI (560)**

**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCATCATGATGATGATGGCGCTGAGCAAGACCTTCGGGCAGAA  
1 M M M M A L S K T F G Q K  
601 GCCCGTGAAGTTCAGCTGGAGGACGACGGCGAGTTCTACATGATCGGCTCCGAGGTGGGAACTACCTCCGTATGTTCCGAGTTCTCTGTACAAGAGA  
13 P V K F Q L E D D G E F Y M I G S E V G N Y L R M F R G S L Y K R  
701 TACCCCTACTCTGGAGGCGACTAGCCACTGTGGAAGAGAGGAAAGAAATAGTTGCATCGTCACATGATCACGGATACAGACTCTAGCCACCAGTGTGA  
47 Y P S L W R R L A T V E E R K K I V A S S H D H G Y T T L A T S V  
801 CCCTGTTAAAAGCCTCGGAAGTGAAGAGATTCTGGATGGCAACGATGAGAAGTACAAGGCTGTGCCATCAGCACAGAGCCCCACCTACCTCAGGGA  
80 T L L K A S E V E E I L D G N D E K Y K A V S I S T E P P T Y L R E  
901 ACAGAAGGCCAAGAGGAACAGCCAGTGGGTACCCACCTGCCAACAGCTCCACCACCTTAGATGCCGTGCCATGCTCCACAACCATCAACAGGAACCGC  
113 Q K A K R N S Q W V P T L P N S S H H L D A V P C S T T I N R N R  
1001 ATGGGCCGAGACAAGAAGAGAACCTTCCCCCTTGTCTTGTGACCATGACCCAGCTGTGATCCATGAGAACGCATCTCAGCCGAGGTGCTGGTCCCCA  
147 M G R D K K R T F P L C F D D H D P A V I H E N A S Q P E V L V P  
1101 TCCGGCTGGACATGGAGATCGATGGGCGAGAAGTGGGAGAGCCTTCCCTGGAACATGAATGAGAAGTTGATGACGCTGAGATGTTTTAGAAATCCT  
180 I R L D M E I D G Q K L R D A F T W N M N E K L M T P E M F S E I L  
1201 CTGTGACGATCTGGATTTGAACCGCTGACGTTTGTGCCAGCCATCGCTCTGCCATCAGACAGCAGATCGAGTCTACCCACCGACAGCATCTGGAG  
213 C D D L D L N P L T F V P A I A S A I R Q Q I E S Y P T D S I L E  
1301 GACCACTCAGACAGCGCTCATCATCAAGCTGAACATCCATGTGGGAAACATTTCCCTGGTGGACCAAGTTGAGTGGGACATGTGAGAGAAGGAGAACT  
247 D Q S D Q R V I I K L N I H V G N I S L V D Q F E W D M S E K E N  
1401 CACCAGAGAAGTTTGCCTGAAGCTGTGCTCGGAGCTGGGTTGGGCGGGGAGTTTGTACCACCATCGCATAACAGCATCCGGGGACAGCTGAGCTGGCA  
280 S P E K F A L K L C S E L G L G G E F V T T I A Y S I R G Q L S W H  
1501 TCAGAAGACTACGCCTTCCAGCGAGAACCCTCTGCCACAGTGGAGATTGCCATCCGGAACACGGGCGATGCGGACCAAGTGGTGCCACTGCTGGAGACT  
313 Q K T Y A F S E N P L P T V E I A I R N T G D A D Q W C P L L E T  
1601 CTGACAGACGCTGAGATGGAGAAGAAGTCCGCGACCAAGGACAGGAACACGAGGCGGATGAGGCGTCTTGCCAACACGGCCCCGGCCTGGTAACCGCCC  
347 L T D A E M E K K I R D Q D R N T R R M R R L A N T A P A W •

**NheI (1710)**

1701 ATCAGCACACGCTAGCTGGCCGACAGATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATT  
1801 TGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGTGT  
1901 GGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAAAATTAACCTCAAATCAAGCCTCACTTGA  
2001 ATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTCTTTTATGGAGTTAAGATAT  
2101 AGTGTATTTTCCAAAGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATGCAGTACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATT  
2201 TAAATACATCATTGCAATGAAAAATAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGG  
2301 GAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTT  
141 • N R T Y K L P I L E E I T T K  
2401 TGACCAGCTTGCCATTCTCAATGAGCACAAGCAGTCAAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCTGATGGA  
125 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 R I S  
2501 TCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAAGCA  
92 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 A E A  
2601 CAGACAGTACCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCACAGTCTGGTCTGATGGCCGCCGACATGGTCTTGTGTCTT  
58 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 K N D E  
2701 CATAGAGCATGGTGTCTTCTCAGTGGCAGCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGTGAGTCTG  
25 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  
2801 ATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTAT  
2901 ATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAAA  
3001 CTCCCATTGACGTCATGGGGTGGAGACTTGGAAATCCCCTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAA  
3101 TAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTC

3201 AATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATT  
3301 GGC GTTACTATGGGAACATACGTCATTATTGACGTCAATGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTG  
3401 CAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGAC  
3501 GAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCGCTCTC  
3601 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT  
3701 GTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTA  
3801 AGACACGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT  
3901 ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC  
4001 CGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCT  
4101 CAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCT  
4201 GTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGT  
4301 GCAGGTGCCAGAACATTTCTCTATCGAA