



1 GGATCTGCGATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

**BstEII (555)**  
**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGAGTACGACGCTTACAACGACTCCGGCATCTATGA  
601 TGATGAGTACTCTGATGGCTTTGGTACTTTGTGGACTTTGGAGGAGGCGAGTCCGTGGGAGGCCAAGGTGGCCCCGGTCTTCTGGTGGTATCTACAGC  
13▶ D E Y S D G F G Y F V D L E E A S P W E A K V A P V F L V V I Y S  
701 TTGGTGTGCTTCTCGGTCTCCTAGGCAACGGCCTGGTATTGTCATCGCCACCTTCAAGATGAAGAAGACCGTGAACACTGTGTGGTTTGTCAACCTGG  
47▶ L V C F L G L L G N G L V I V I A T F K M K K T V N T V W F V N L  
801 CTGTGGCCGACTTCTGTTCAACATCTTTTGGCGATGCACATCACCTACGCGCCATGGACTACCACTGGGTGTTCCGGAAGGCCATGTGCAAGATCAG  
80▶ A V A D F L F N I F L P M H I T Y A A M D Y H W V F G K A M C K I S  
901 CAACTTCTTGCTCAGCCACAACATGTACACCAGCGTCTTCTGCTGACTGTATCAGCTTTGACCGCTGCATCTCCGTGCTGCCCGTCTGGTCCAG  
113▶ N F L L S H N M Y T S V F L L T V I S F D R C I S V L L P V W S Q  
1001 AACCAACGAGCATCCGCTGGCTACATGACTGCTCGCCGTCTGGTCTGGCTTTCTTCTGAGTCCCGTCCCTTGTCTCCGGGACACCGCA  
147▶ N H R S I R L A Y M T C S A V W V L A F F L S S P S L V F R D T A  
1101 ACATTCATGGGAAGATAACCTGTTCAACAACCTCAGCTTGGCCGCGCTGAGTCTCCCCACATCCCGCCACTCGCAAGTAGTTTCCACAGGGTACAG  
180▶ N I H G K I T C F N N F S L A A P E S S P H P A H S Q V V S T G Y S  
1201 CAGACAGTGGCGTCACTGTACCCGCTTCTTTGGCGTCTCTGATCCCGTCTTATCATCACGGCCTGCTACCTTACCATCGTCTTCAAGCTGCAG  
213▶ R H V A V T V T R F L C G F L I P V F I I T A C Y L T I V F K L Q  
1301 CGCAACCGCTGGCCAAGAACAAGAAGCCCTTCAAGATCATTATCACCATCATCATCACCTTCTTCTGCTGGTGGCCCTACCACACCCTTACCTGC  
247▶ R N R L A K N K K P F K I I I T I I I T F F L C W C P Y H T L Y L  
1401 TGGAGCTCCACCACAGCTGTGCCAAGCTCTGTCTTACGCTGGGGTACCCTGGCCACGGCGTCCCATCGCCAACAGCTGCATGAACCCATTCT  
280▶ L E L H H T A V P S S V F S L G L P L A T A V A I A N S C M N P I L  
1501 GTACGCTTTCATGGGCCAGACTTCAGAAAATTCAAGTGGCCCTTCTTCCCGCTGGCCAACGCCCTGAGTGAGGACACAGGCCCTCTCTACCCC  
313▶ Y V F M G H D F R K F K V A L F S R L A N A L S E D T G P S S Y P

**NheI (1686)**

1601 AGTCACAGGAGCTTACCAAGATGTCGTCTTTGAATGAGAAGGCTTCGGTGAATGAGAAGGAGACCAGTACCCTCTGAACCTCACCGCTAGCTGGCCAGA  
347▶ S H R S F T K M S S L N E K A S V N E K E T S T L •  
1701 CATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTGTGA  
1801 ACCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTGCATTATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAA  
1901 ACCTTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAG  
2001 GCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCAAGGTTTGAAC  
2101 TAGCTCTTATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATGATAAATATTAGAAAATAATTTAAATACATCATTGCAATGAAAAT  
2201 AAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAA  
2301 TTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAA  
141▶ • N R T Y K L P I L E E I T T K V L K G N M E I  
2401 TGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGCCACCTCATCAGAGTAGGG  
117▶ 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 R D V E D S Y P  
2501 GTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAAGCAGACAGTACCCTGCCAATGTAG  
84▶ 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 C V T V R G I Y  
2601 GCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCCGACATGGTGCTTGTGTCTCATAGAGCATGGTGATCTTCTCAG  
50▶ A E I H V A S 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  
2701 TGGCACTCCACAGCTCAGATCTGCTGAGAGATGTTGAAGTCTTCAAGTGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATG  
17▶ A V E V L E L D Q Q S I N F T K M  
2801 CCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCT  
2901 ACCGCCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCATTGACGTCAATGGGGTGG  
3001 AGACTTGGAAATCCCGTGGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGT  
3101 ACTGCCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATTGACGTCAATAGGGGGGCTACTTGGCATATG

3201 ATACACTTGATGTACTGCCAAGTGGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTC  
3301 ATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTACGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGC  
3401 AAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCA  
3501 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGTCCCTCGTGCGCTCTCTGTTCCGACCCTGCCGTTACCG  
3601 GATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGTGGG  
3701 CTGTGTGCACGAACCCCCGTTAGCCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA  
3801 GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTAT  
3901 TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTGT  
4001 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACACTCACGTTAA  
4101 GGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATC  
4201 GTAACATAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTAT  
4301 CGAA