



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCCGCTTTCGCCCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCGGCGCTCCCTTGAGAGCTACCTAGACTCAGCCGGCTCTCCAGGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

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

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCTACCTGAGATCACCGGTCCGGGAGTGTGGCCGAGTGCCTGCAGCAGGA  
601 GACCACCTGCCCGTGTGCCTGCAGTACTTCGAGAGCCCATGATGCTCGACTGCGGCCATAACATCTGTTGCGCGTGCCTCGCCCGTGTGGGGCAGC  
13▶ T T C P V C L Q Y F A E P M M L D C G H N I C C A C L A R C W G T  
701 GCAGAGACTAACGTGTCGTGCCCGCAGTCCCGGAGACCTCCCGCAGAGGCACATGCGGCCAACCGGCACCTGGCCAACGTGACCCAACCTGGTAAAGC  
47▶ A E T N V S C P Q C R E T F P Q R H M R P N R H L A N V T Q L V K  
801 AGTGCACCCAGGAGCGCGCTCGGGGCCGGCGGAGATGGGCGTGTGCGAGAAGCACCGGAGCCCTGAAGTGTACTGCGAGGAGGACAGATGCC  
80▶ Q L R T E R P S G P G G E M G V C E K H R E P L K L Y C E E D Q M P  
901 CATCTGCGTGGTGTGCGACCGCTCCCGGAGCAGCCGCGCCACAGCGTGTCCGCTCGAGGAGCGGTGGAGGGCTTCAAGGAGCAATCCAGAACCAG  
113▶ I C V V C D R S R E H R G H S V L P L E E A V E G F K E Q I Q N Q  
1001 CTCGACCATTTAAAAAGAGTGAAGATTTAAAGAAGAGACGTGGGCCAGGGGAACAGGCACGAGCTGAAGTCTTGAAGCTAACCCAGATGGAGAGGG  
147▶ L D H L K R V K D L K K R R R A Q G E Q A R A E L L S L T Q M E R  
1101 AGAAGATTGTTGGGAGTTTGAAGAGTGTATCACTCTTAAAGGAGCATGAGTATCGCTCTGGCCCGCTTGAAGAGCTAGACTTGGCCATCTACAA  
180▶ E K I V W E F E Q L Y H S L K E H E Y R L L A R L E E L D L A I Y N  
1201 TAGCATCAATGGTGCATCACCCAGTCTCTTGAACATCTCCACCTCAGCAGCCTGATCGCTCAGCTAGAAGAGAAGCAGCAGCAGCCACCCAGGGAG  
213▶ S I N G A I T Q F S C N I S H L S S L I A Q L E E K Q Q P T R E  
1301 CTCCTGCAGGACATTGGGAGACATTGAGCAGGCTGAAAGAAATCAGGATTCCTGAACCTTGGATCACACCTCAGATTTGCAAGAGAAAAATCCACATTT  
247▶ L L Q D I G D T L S R A E R I R I P E P W I T P P D L Q E K I H I  
1401 TTGCCAAAAATGTCTATTCTTACGGAGAGTCTAAAGCAGTTCACAGAAAAATGCAGTCAGATATGGAGAAAAATCAAGAATTAAGAGAGGCTCAGTT  
280▶ F A Q K C L F L T E S L K Q F T E K M Q S D M E K I Q E L R E A Q L  
1501 ATACTCAGTGGACGTGACTCTGGACCCAGACACGGCCTACCCAGCCTGATCCTCTGATAATCTCGGCAAGTGGGTACAGTTACCTCCAACAGGAC  
313▶ Y S V D V T L D P D T A Y P S L I L S D N L R Q V R Y S Y L Q Q D  
1601 CTGCTGACAACCCGAGAGGTTCAATCTGTTCCCTGTGCTTGGGCTCTCCATGCTTATCGCCGGGAGACATTATTGGGAGGTAGAGGTGGGAGATA  
347▶ L P D N P E R F L A G F C V L G S P C F I A G R H Y W E V E V G D  
1701 AAGCCAAGTGGACCATAGGTGTCTGTGAAGACTCAGTGTGCAGAAAAGGTGGAGTAACCTCAGCCCCAGAATGGATTCTGGGAGTGTCTTTGTGGTA  
380▶ K A K W T I G V C E D S V C R K G G V T S A P Q N G F W A V S L W Y  
1801 TGGAAAGAATATTGGGCTCTTACCTCCCAATGACTGCCCTACCCCTGCGGACCCGCTCCAGCGGGTGGGATTTTCTGGACTATGATGCTGGTGTAG  
413▶ G K E Y W A L T S P M T A L P L R T P L Q R V G I F L D Y D A G E  
1901 GTCTCTTCTACAACGTGACAGAGAGGTGTACACCTTCACTTTCTCTCATGTACTCTTTGTGGCCTGTCCGGCCCTACTTCAGTCTGAGTACTCGG  
447▶ V S F Y N V T E R C H T F T F S H A T F C G P V R P Y F S L S Y S  
2001 GAGGAAAAGTGCAGCTCTCTGATCATCTGCCCATGAGTGGGATAGATGGGTTTTCTGGCCATGTTGGGAATCATGGTCATTCCATGGAGACCTCCCC  
480▶ G G K S A A P L I I C P M S G I D G F S G H V G N H G H S M E T S P

**NheI (2125)**

2101 TTGAGGAGGTGAATTCAGGCCAAAAGCTAGTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGC  
513▶ .  
2201 TTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTCATTCTTTATGTTTCAGG  
2301 TTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAATGTGGTATGGAATTTAAAATACAGCATAGCAAAACCTTAACTCCAAT  
2401 CAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCA  
2501 TGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTCTTCACTTTCTTTATGTTTTAAATGCAGTACCTCCACATTCCCTTTTTAGTAAAA  
2601 TATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTA  
2701 GTAGTTGGACTTAGGGAACAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTT  
2801 CCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGCT  
130▶ 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 C P S  
2901 GACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATG  
97▶ 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 S G I  
3001 GCAATGGCTTACGACAGACAGTACCCTGCCAATGTAGGCCCAATGTGGACAGCAGAGATGATCTCCCACTTGGTCTGATGGCCGCCCGACAT  
63▶ 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 A G V H  
3101 GGTGCTTGTCTCATAGAGCATGGTGTCTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCAATGGTGGCCCT  
30▶ 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

3201 CCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAA  
3301 ACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTT  
3401 ACTAGTCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACC  
3501 GCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAAGGCATGTACTGGGCATAATGCCAGGCGGGCCATTT  
3601 ACCGTATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAA  
3701 TGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAA  
3801 GTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGG  
3901 CTCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCT  
4001 CCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAG  
4101 GTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTT  
4201 GAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGA  
4301 AGTGGTGGCCTAACTACGGTACACTAGAAGAAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC  
4401 CGGCAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCT  
4501 ACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTT  
4601 ATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCT  
4701 GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA