



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCCGCTTGTCCGCCCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCTTTGTCCGGCGCTCCCTTGAGAGCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

**NcoI (560)**  
**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCAccggtCACCATGGCGAGGGGCTTGATCTAGCACCGCTGCTACTGCT  
1▶ M A R G L D L A P L L L L L  
601 ACTGCTGGCGATGGCGACCCGCTTTTGACGGCTCAGAGCAACTGTACATGCCCCACCAACAAGATGACGGTCTGCGACACAAATGGCCAGCGGGGTG  
13▶ L L A M A T R F C T A Q S N C T C P T N K M T V C D T N G P G G V  
701 TGCCAATGTCGGGCAATGGGCTCACAGGTATTGGTCGACTGTCCACGCTAATTCCAAGTGCCTGCTGCTCAAGGCGCGCATGAGCGCCCGAAGAGCG  
47▶ C Q C R A M G S Q V L V D C S T L T S K C L L L K A R M S A R K S  
801 GCCGACGCTGTTGATGCCGAGCGAGCACGCGATACTGGACAACGATGGCCTCTACGACCCGGAGTGTGACGACAAGGGCCGCTTCAAGGCGCGCAGTG  
80▶ G R S L V M P S E H A I L D N D G L Y D P E C D D K G R F K A R Q C  
901 CAACCAGACCTCGGTGTGCTGGTGCCTAACTCGTGGGCGTGCGCCGACGACAAGGGAGACCAAGCCTGCGCTGCGACGAAGTGGTCCGAACCCAC  
113▶ N Q T S V C W C V N S V G V R R T D K G D Q S L R C D E V V R T H  
1001 CACATCTCATTGAGTTGCGCCACCGCCGACCGAGCCTTCAACCACTTGACCTAGACTCCGAGCTGCGGGGCTCTTCCAAGAACGTACAAGC  
147▶ H I L I E L R H R P T D R A F N H S D L D S E L R R L F Q E R Y K  
1101 TGCACCCAGCTTCTATCCGCGGTACTATGAGGAGCCACCACTCAGATAGAGCTTCGGCAGAACGCGTGCAGAAGGGCTTGAAGAGCTGGACAT  
180▶ L H P S F L S A V H Y E E P T I Q I E L R Q N A S Q K G L R D V D I  
1201 CGTGTATGCCGCTACTACTTCGAAAGGACATTAAGGGCAGTCACTGTTTATGGCCGCGCGGCTGGACGTGACGGTGCCTGGGGAACCCCTGCAT  
213▶ A D A A Y Y F E R D I K G E S L F M G R R G L D V Q V R G E P L H  
1301 GTGGACGGGACGCTCATCTACTGAGCGAGAGAGCCCGGAGTTCATGAAGCGCCTACCCGCGGCTCATTGCCGTATCGCTGTGCTTCGG  
247▶ V E R T L I Y Y L D E K P P Q F S M K R L T A G V I A V I A V V S  
1401 TAGCGGTAGTGGTGGTGGTGGTCTTGGTGGTCCACAAACGGAGGAAGTCGGGCAAATACAAAAGGTGGAGCTTAAAGGAGCTGGGGGAGATGAGAAG  
280▶ V A V V A G V V V L V V T K R R K S G K Y K K V E L K E L G E M R S

**NheI (1537)**

1501 CGAACCTAGCTTGTAGTTTCTGTAGGATGCCCGAgCTAGTCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCAACTAGAATGCAG  
313▶ E P S L •  
1601 TGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATT  
1701 TTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTT  
1801 TAACCTCAAATCAAGCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCT  
1901 CACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTACTCTTCAATTTCTTATGTTTTAAATGCACTGACCTCCCACATCCC  
2001 TTTTATGATAAATATTAGAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTCATAATA  
2101 TCCCCAGTTTGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTCTCCTGGTACTTGA  
141◀ • N R T Y K L  
2201 GGGGATGAGTTCTCAATGGTGGTTTTGACAGCTTGCATTTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACAT  
134◀ P I L E 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  
2301 GCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACA  
101◀ G C P S 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  
2401 GCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGATGGCCTCAATGTGGACAGAGATGATCTCCCAGTCTTGGTCTGATGG  
67◀ A S G I 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  
2501 CCGCCCCGACATGGTCTTGTGCTCATAGAGCAGTGGTATCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGCTCT  
34◀ A G V H 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  
2601 CATGGTGGCCCTCTATAGTGTGATTATACTATGCGCATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTG  
1◀ M  
2701 ACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAG  
2801 TCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGT  
2901 ACTGCCAAAACCGCATCATGGAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGCATAATGCCA  
3001 GCGGGCCATTTACCGTCAATGAGGCGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCAGTTTACCGTAAATACTCCACC  
3101 CATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTTGACGTCAATGGGGGGGTCGTTGGGCGGTACGCCAGGCGGGC

3201 CATTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCG  
3301 TTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC  
3401 CCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAG  
3501 CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCCTGCGCCTTATCCGGT  
3601 AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA  
3701 CAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGG  
3801 TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCT  
3901 TTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAA  
4001 TAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTA  
4101 GCAAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA