



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCCATGGCTGAAGGGAAATCACACCTTACAGCCCTGAC  
1▶ M A E G E I T T F T A L T  
601 CGAGAAGTTAATCTGCCTCCAGGAATTACAAGAAGCCAAACTCCTCTACTGTAGCAACGGGGCCACTTCTGAGGATCCTTCCGGATGGCACAGTG  
13▶ E K F N L P P G N Y K K P K L L Y C S N G G H F L R I L P D G T V  
701 GATGGGACAAGGGACAGGAGCGACCAGCACATTAGCTGCAGCTCAGTGGGAAAGCGTGGGGAGGTGTATATAAGAGTACCGAGACTGGCCAGTACT  
47▶ D G T R D R S D Q H I Q L Q L S A E S V G E V Y I K S T E T G Q Y  
801 TGGCCATGGACACCGACGGGCTTTTATACGGCTCACAGACACCAAATGAGGAATGTTTGTCTGAAAGGCTGGAGGAGAACCATTACAACCTATAT  
80▶ L A M D T D G L L Y G S Q T P N E E C L F L E R L E E N H Y N T Y I  
901 ATCCAAGAAGCATGCAGAGAAGAATTGGTTTGTGGCCTCAAGAAGAATGGGAGCTGCAAACGCGTCTCGGACTCACTATGGCCAGAAAGCAATCTTG  
113▶ S K K H A E K N W F V G L K K N G S C K R G P R T H Y G Q K A I L

**NheI (1057)**

1001 TTTCTCCCGTCCAGTCTCTTCTGATTAAGAGATCTGTTCTGGGTGTGACCACTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGAC  
147▶ F L P L P V S S D •  
1101 AAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTATTTGTAACCATTATAAGTGAATAACAAGTTAA  
1201 CAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAATGTGGTATGGAATTCTAA  
1301 AATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGT  
1401 GCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTGCTTCTTATTTTATGTTTTAAATG  
1501 CACTGACCTCCCACATCCCTTTTATAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAAATCCAGAT  
1601 GCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCT  
1701 TTAGTTCCTGGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGTTCGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCA  
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 D P A Y D  
1801 GAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGT  
107▶ 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 V I T D F D  
1901 CCTTCTGCCCCTGCTCAGCAGACACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTC  
74▶ 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 A S I I E  
2001 CCCAGTCTGGTCTGATGGCCGCCGACATGGTGTCTGTTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCCTGC  
41▶ 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 L E L D Q  
2101 TGAGAGATGTTGAAGGCTTTCATGATGGCCCTCTATAGTGGTCTGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGA  
7▶ Q S I N F T K M  
2201 TGGCGTCTCCAGCTTATCTGACGTTCACTAAACGAGCTCTGTTATATAGACCTCCACCGTACACGCTACCGCCATTGCGTCAATGGGGCGGAGT  
2301 TGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGGTGAGTCAAACCG  
2401 CTATCCACGCCATTGATGACTGCCAAAACCGCATCATGTTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCCATAAAGTGC  
2501 ATGTAAGTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGGCTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCAG  
2601 TTTACCGTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTCGTT  
2701 GGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACCGCTGAGGTTAATTAAGAACATGTGAGCAAAAAGCCAGCAAAAAGCCAGGAACCGTA  
2801 AAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT  
2901 ATAAAGATACCAGGCGTTTTCCCGTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGA  
3001 AGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCG  
3101 ACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGC

3201 GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT  
3301 ACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAA  
3401 AAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTGGTCATGGCTAGTTAATTAAC  
3501 ATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACA  
3601 AAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA