



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTTCGGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTAGGAGGGCACCATGGCTGAAGGGGAGATCACACCTTCGCAG  
601 CCCTGACCAGAGGTTAACCTGCCTCTAGGAACTACAAAAAGCCAACTGCTCTACTGCAGCAACGGGGCCACTTCTTGAGGATCCTTCTGATGG  
11▶ A L T E R F N L P L G N Y K K P K L L Y C S N G G H F L R I L P D G  
701 CACCGTGGATGGGACAAGGGACAGGAGCGACCAGCACATTCAGCTGCAGCTCAGTGGGAAAGTGGGGCGAAGTGTATATAAAGGGTACGGAGACCGGC  
44▶ T V D G T R D R S D Q H I Q L Q L S A E S A G E V Y I K G T E T G  
801 CAGTACTTGCCATGGACACCGAAGGGCTTTTATACGGCTCGCAGACACCAAATGAGGAATGTCTGTTCTGGAAAGGCTGGAAGAAAACATTATAACA  
78▶ Q Y L A M D T E 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  
901 CTTACACCTCCAAGAAGCATGCGGAGAAGAAGTGGTTTGTGGCCTCAAGAAGAACGGGAGCTGTAAGCGCGTCTCGGACTCACTATGGCCAGAAAAGC  
111▶ T Y T 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

**NheI (1060)**

1001 CATCTTGTCTGCCCTCCCGGTGCTTCTGACTAGAGGAGTCTGTTCTGAGTGTCTCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTG  
144▶ I L F L P L P V S S D •  
1101 GACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGTCAATAAACAAGT  
1201 TAACAACAACAATTGCATTCAATTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTC  
1301 TAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAA  
1401 TGTGCATTAGCTGTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAGTACTGCTTTCATTCTTTATGTTTTAA  
1501 ATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCA  
1601 GATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTA  
1701 GCTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAG  
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  
1801 TCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAA  
108▶ 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 H R V A V I T D F  
1901 AGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGAT  
75▶ D 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  
2001 CTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTTCCTCATAGAGCATGGTGTATCTCTCAGTGGCGACCTCCACCAGTCCAGATCC  
42▶ 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 A V E V L E L D  
2101 TGCTGAGAGATGTTGAAGGCTTCATGATGGCCCTCTATAGTGAAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGT  
8▶ Q Q S I N F T K M  
2201 GGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTCGCTCAATGGGGCGG  
2301 AGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCTGAGTCAA  
2401 CCGTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATGTTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGAAAGTCCATAAG  
2501 GTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGG  
2601 CAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGTCAATGGCGGGGGTCT  
2701 GTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTATGTAACGCTGACGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACC  
2801 GTAAAAAGCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGG  
2901 ACTATAAAGATACCAGCGTTTTCCCGTGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCG  
3001 GGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTCCGCTCAAAGTGGGCTGTGTGCACGAACCCCGTTACG  
3101 CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA

3201 AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA  
3301 GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGCAAGCAGCAGATTACGCGCAGAA  
3401 AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATT  
3501 AACATTTAAATCAGCGCGCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAA  
3601 ACAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCGAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA