



1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCCAGCGCCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT  
**AgeI (552)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGTACTCATTGTTTCGGAAGAGACAGGTACGTTAAT  
601 AGTTAATAGCGTACTTCTTTTTCTGCTTTCGTGATTCTTCTAGTTACACTAGCCATCCTTACTGCGCTTCGATTGTGTGCGTACTGCTCAATATT  
13▶ V N S V L L F L A F V V F L L V T L A I L T A L R L C A Y C C N I  
**NheI (791)**  
701 GTTAACGTGAGTCTTGAAAACCTTCTTTTTACGTTTACTCTCGTGTAAAAATCTGAATCTTCTAGAGTTCCTGATCTTCTGGTCTAAAGCTAGCTGG  
47▶ V N V S L V K P S F Y V Y S R V K N L N S S R V P D L L V •  
801 CCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTAT  
901 TTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAA  
1001 GTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCTCTACTTGAATCCTTTTCTGAGGGATGA  
1101 ATAAGGCATAGGCATCAGGGGCTGTTGCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGTT  
1201 TGAAGTACTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATG  
1301 AAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAAT  
1401 AGAAATTTGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCAT  
141◀ • N R T Y K L P I L E E I T T K V L K G N M  
1501 CTCAATGAGCACAAGCAGTCCAGGAGCATAGTCCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCTGATGGATCTGCCACCTCATCAGAG  
119◀ E I 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  
1601 TAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCAGACAGTGCACCTGCCAA  
85◀ Y P 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  
1701 TGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTGTTGCTCCTATAGACATGGTATCTT  
52◀ Y A E I H V A S I I E G T K T R I A A G V H H K N D E Y L M T I K  
1801 CTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCTGTTGGCCCTCTATAGTGAGTTCGATTATACTATGCCGATATA  
19◀ E T A V E V L E L D Q Q S I N F T K M  
1901 CTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACA  
2001 CGCCTACCGCCCATTTGCGTCAATGGGCGGAGTGTACGACATTTGGAAAGTCCCCTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGG  
2101 GGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTA  
2201 GATGTAAGTCCAGTAAAGTCCCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCGTACTGGC  
2301 ATATGATACACTTGATGTAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACAT  
2401 ACGTCAATATTGACGTCAATGGGCGGGGTCGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCTGCAGGTTAATTAAGAACATG  
2501 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCGCGTGTGCTGGCGTTTTCCATAGGCTCCGCCCTGACGAGCATCACAATAATCGAC  
2601 GCTCAAGTCCAGAGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGTTTTCCCTGGAAGTCCCTCGTGCCTCTCCTGTTCCGACCTGCCGCT  
2701 TACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAG  
2801 CTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGTGGCGTTATCCGGTAACTATCGTCTTGAAGTCCAAACCGGTAAGACACGACTTATGCCAC  
2901 TGGCAGCAGCCTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAAC  
3001 AGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACACCGCTGGTAGCGGTGTTTT  
3101 TTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACTCAC

3201 GTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGT

3301 GAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTTC

3401 TCTATCGAA