



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGGCGCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCAGGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTATGGAAGGAGTGTCCAGAAAAGAGTAGCAGCAG  
1 M E K K E C P E K S S S S S  
601 TGAGGAAGAGCTGCCGAGACGGGACAGTGGATCCAGTAGGAACATAGATGCATCCAACTCATTAGACTACAAGGATCACGGAAACTGTTGGTTGACAAT  
13 E E E L P R R D S G S S R N I D A S K L I R L Q G S R K L L V D N  
701 AGCATACGGAACTGCAATACACAAAAGTGAATTTTTTTCAGGCTGAGGCTGTGTGACAAATGATACGGTATACAGGGAGTACCTGTGTTTCTG  
47 S I R E L Q Y T K T G I F F Q A E A C V T N D T V Y R E L P C V S  
801 AGACCCTTTGTGACATCTCACATTTTTTCCAAGAAGATGATGAGACAGAGGACAGCCATTATTGTTCCGTGCTGTTCTGAGTGTCAACTATCTGGGG  
80 E T L C D I S H F F Q E D D E T E A E P L L F R A V P E C Q L S G G  
901 GGACATCCAGTGTATCAGAAGAGCAGGAATCTCAGAGGGACAAGATTGAGGACATTTGCTCAGAAGAGATCAAATAGTTTCTTATGCTTCT  
113 D I P S V S E E Q E S S E G Q D S G D I C S E E N Q I V S S Y A S  
1001 AAAGTCTGTTTTGAGATCGAAGAAGATTATAAAAATCGTCAGTTTCTGGGGCTGAAGGAAATGTGGATGTTGAGTTGATTGATAAGAGCACAAACAGAT  
147 K V C F E I E E D Y K N R Q F L G P E G N V D V E L I D K S T N R  
1101 ACAGCGTTTGGTCCCCACTGCTGGCTGGTATCTGTGGTCAGGCACAGGCCTCGCTTCTGGTAAGGGATGAGGTACAGTACGATTGCGTTTGGTTC  
180 Y S V W F P T A G W Y L W S A T G L G F L V R D E V T V T I A F G S  
1201 CTGGAGTCAGCACCTGGCCCTGGACCTGCAGCACCATGAACAGTGGCTGGTGGGGCGCCCTGTTTGTGATGCTACTGCAGAGCCAGAGGGGCTGCGCC  
213 W S Q H L A L D L Q H H E Q W L V G G P L F D V T A E P E E A V A  
1301 GAAATCCACTCCCACTTCTCCTCCCAAGCAGGTGAGGTGGACGTCTCTGTTTCTGTTGCCATTTTAAAGATGAAGGGATGGTCTGGAGC  
247 E I H L P H F I S L Q A G E V D V S W F L V A H F K N E G M V L E  
1401 ATCCAGCCGGGTGGAGCCTTTCTATGCTGTCTGAAAGCCCGACTTCTCTGATGGGCATCTGCTGCGGATCGCCAGTGGGACTCGCCTCTCCAT  
280 H P A R V E P F Y A V L E S P S F S L M G I L L R I A S G T R L S I  
1501 CCCCATCACTTCAACACATTGATCTATTATCACCCCCACCCGAAGATTAAGTCCACTTGTACCTTGTCCCCAGCGACGCTTGTCTAACAAAGGCG  
313 P I T S N T L I Y Y H P H P E D I K F H L Y L V P S D A L L T K A  
1601 ATAGATGATGAGGAAGATCGCTTCCATGGTGTGCGCCTGCAGACTTCGCCCAATGGAACCCCTGAACTTTGGTCCAGTTATATTGTGTCTAATTCTG  
347 I D D E E D R F H G V R L Q T S P P M E P L N F G S S Y I V S N S  
1701 CTAACCTGAAAGTAATGCCCAAGGAGTTGAAATTTGCTCAGAGGCCTGGAGAAATTCAGCACTTCTCAAAATTTCTATGCTGGGAGATGAAGGAAAC  
380 A N L K V M P K E L L K L S Y R S P G E I Q H F S K F Y A G Q M K E P  
1801 CATTCACTTGTAGATTACTGAAAAAGACATGGGACTTTGGTGTGGGATACTGAGGTGAAGCCAGTGGATCTCCAGCTTGTAGCTGCATCAGCCCTCT  
413 I Q L E I T E K R H G T L V W D T E V K P V D L Q L V A A S A P P  
1901 CCTTCTCAGGTGCAGCCTTTGTGAAGGAGAACCACCGGCACTCCAAGCCAGGATGGGGACCTGAAAGGGGTGCTCGATGATCTCCAGGACAATGAGG  
447 P F S G A A F V K E N H R Q L Q A R M G D L K G V L D D L Q D N E  
2001 TTCTTACTGAGAATGAGAAGGAGCTGGTGGAGCAGGAAAAGACACGGCAGAGCAAGAATGAGGCCCTTGTGAGCATGGTGGAGAAGAAAGGGGACCTGGC  
480 V L T E N E K E L V E Q E K T R Q S K N E A L L S M V E K G D L A

**NheI (2185)**

2101 CCTGGACGTGCTCTCAGAAGCATTAGTGAAGGGACCTTACCTCGTGCCTATCTTAGACAGCAGAATTTGTAATGAGTCAGTAGCTGGCCAGAC  
513 L D V L F R S I S E R D P Y L V S Y L R Q Q N L  
2201 ATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAA  
2301 CCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTCATTCTTTATGTTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAA  
2401 CCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGG  
2501 CATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCGAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAC  
2601 AGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAA  
2701 AATGTTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTAAATAGAAAT  
2801 TGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAAT  
141 N R T Y K L P I L E E I T T K V L K G N M E I  
2901 GAGCACAAGCAGTCCAGGAGCATAGTCAAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGG  
117 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 Y P  
3001 TGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAAATGGCAATGGCTTCCAGCAGACAGTACCTGCAATGTAGG  
83 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 Y A  
3101 CCTCAATGTGGACAGCAGAGATGATCTCCCACTTGGTCTGATGGCCGCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGT  
50 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 E T

3201 GCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGC  
17 A V E V L E L D Q Q S I N F T K M  
3301 CGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTA  
3401 CCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGA  
3501 GACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGA  
3601 CTGCCAAGTAGAAAGTCCATAAAGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGA  
3701 TACACTTGATGTACTGCCAAGTGGGAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCA  
3801 TTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAGAACATGTGAGCA  
3901 AAAGGCCAGAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAA  
4001 GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGG  
4101 ATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGC  
4201 TGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAG  
4301 CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATT  
4401 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGCAAACAACCCACCGCTGGTAGCGGTGGTTTTTTTGT  
4501 TGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAG  
4601 GGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGCCGAATAAAATATCTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTAATCG  
4701 TAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATC  
4801 GAA