



100

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGGCGCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCCGGTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

**NcoI (560)**

**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCTACCTGAGATCACCGGTCCACATGGCGCCCGCCCTCTGGGTGCGCTGGTCTTCGA  
1 M A P A A L W V A L V F E  
601 ACTGCAGCTGTGGCCACCGGGCACACAGTGCCCGCCAGGTTGTCTTGACACCCTACAAACCGAACCTGGGTACGAGTGCCAGATCTCACAGGAATAC  
13 L Q L W A T G H T V P A Q V V L T P Y K P E P G Y E C Q I S Q E Y  
701 TATGACAGGAAGGCTCAGATGTGCTGTGCTAAGTGTCTCCTGGCCAAATATGTAAACATTTCTGCAACAAGACCTCGGACACCGTGTGTGCGGACTGTG  
47 Y D R K A Q M C C A K C P P G Q Y V K H F C N K T S D T V C A D C  
801 AGGCAAGCATGTATACCAGGTCTGAAACAGTTCGTACATGTTGAGTGCAGTTCCTCTGTACCACTGACCAGGTGGAGATCCGCGCCTGCACTAA  
80 E A S M Y T Q V W N Q F R T C L S C S S S C T T D Q V E I R A C T K  
901 ACAGCAGAACCAGTGTGTGCTTGGCAAGCTGGCAGGTACTGCGCCTTGAAACCCATTCTGGCAGTGTGACAGTGCATGAGGCTGAGCAAGTGGCGC  
113 Q Q N R V C A C E A G R Y C A L K T H S G S C R Q C M R L S K C G  
1001 CCTGGCTTCGAGTGGCCAGTCAAGAGCCCAATGGAAATGTGCTATGCAAGGCTGTGCCCGAGGACGTTCTCTGACACCACATCATCCACTGATG  
147 P G F G V A S S R A P N G N V L C K A C A P G T F S D T T S S T D  
1101 TGTGACGCCCCACCGCATCTGTAGCATCTGGCTATCCCGAAATGCAAGCACAGATGCAGTCTGTGCGCCGAGTCCCAACTCTAAGTGCCATCCC  
180 V C R P H R I C S I L A I P G N A S T D A V C A P E S P T L S A I P  
1201 AAGGACTCTACGTATCTCAGCCAGAGCCACAAGATCCCAACCCCTGGATCAAGAGCCAGGGCCAGCCAAACTCCAAGCATCCTTACATCGTTGGGT  
213 R T L Y V S Q P E P T R S Q P L D Q E P G P S Q T P S I L T S L G  
1301 TCAACCCCATTTATGAACAAAGTACCAAGGTGGCAGTCTCTCCAATTGGTCTGATTGTTGGAGTGACATCACTGGGTCTGCTGATGTTAGGATGG  
247 S T P I I E Q S T K G G I S L P I G L I V G V T S L G L L M L G L  
1401 TGAAGTGCATCATCTGGTGCAGAGGAAAAAGGCCCTCTGCTACAAAGAGATGCCAAGTGCCTCATGTGCCTGATGAGAAATCCAGGATGCAGT  
280 V N C I I L V Q R K K K P S C L Q R D A K V P H V P D E K S Q D A V  
1501 AGGCCTTGAGCAGCAGACCTGTTGACCACAGCACCCAGTTCAGCAGCAGTCCCTAGAGAGCTCAGCCAGCGCTGGGGACCGAAGGGCGCCCCCTGGG  
313 G L E Q Q H L L T T A P S S S S S S L E S S A S A G D R R A P P G  
1601 GGGCATCCCCAAGCAAGATCATGGCGGAGGCCAAGGTTTCAGGAGGCCGCTGCCAGTCCAGGATTTAGATTCTTCCCAGGAAGCCACGGGACCC  
347 G H P Q A R V M A E A Q G F Q E A R A S S R I S D S S H G S H G T  
1701 ACGTCAACGTCACTGCATCGTGAACGCTGTAGCAGCTGTGACCACAGTTCAGTGTCTTCCCAAGCCAGCCACAGTGGGAGACCCAGATGCCAA  
380 H V N V T C I V N V C S S S D H S S Q C S S Q A S A T V G D P D A K  
1801 GCCCTCAGCGTCCCCAAGGATGAGCAGGTCCCCTTCTCAGGAGGAGTGTCCGTCTCAGTCCCCGTGTGAGACTACAGAGACACTGCAGAGCCATGAG  
413 P S A S P K D E Q V P F S Q E E C P S Q S P C E T T E T L Q S H E  
1901 AAGCCCTTGCCCTTGGTGTGCCGATATGGGCATGAAGCCAGCCAAAGCTGGCTGGTTTGATCAGATTGCAGTCAAAGTGGCCTGACCCCTGACAGGGG  
447 K P L P L G V P D M G M K P S Q A G W F D Q I A V K V A •

**NheI (2022)**

2001 TAACACCCTGCAAAGGGACCCCGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTT  
2101 ATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGTGCATAAACAAGTTAACAACAACAAATGCATTATTTTTATGTTTCAGGTTCC  
2201 AGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACTCCAATCAA  
2301 GCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGG  
2401 AGTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCAGTGCCTCCACATTCCCTTTTTAGTAAAATAT  
2501 TCAGAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTA  
2601 GTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTTCTGGTGTACTTGAGGGGGATGAGTTCCT  
141 • N R T Y K L P I L E E  
2701 CAATGGTGGTTTTGACCAGTTCGCAATTCATCTCAATGAGCACAAGCAGTCCAGGACATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGAC  
129 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 G C P S V  
2801 CACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCCCTTCTGCCCGTGTCTCAGCAGCAGCCCAATGGCA  
96 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 A S G I A  
2901 ATGGCTTCAGCACAGACAGTGCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGT  
62 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 A G V H H  
3001 GCTTGTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCCT  
29 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 M  
3101 ATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACG

3201 AGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTACT  
3301 AGTCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTTGATGTACTGCCAAAACCGCA  
3401 TCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACC  
3501 GTCATTGACGTCAATAGGGGGCTACTTGCCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGG  
3601 AAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTT  
3701 ATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC  
3801 CGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCCTGGAAGCTCCC  
3901 TCGTGGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTA  
4001 TCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAG  
4101 TCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT  
4201 GGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGG  
4301 CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG  
4401 GGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGATTTTGGTCATGGCTAGTTAATTAACATTTAATCAGCGGCCGAATAAAATATCTTTATT  
4501 TTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTC  
4601 CCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA