



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTGTGCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGTTGAGTTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCGGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

AgeI (552) NcoI (568)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTAGGAGGGCCACCATGGTGTGGAAGGAAACCTGAAGTGGGG  
1 M V L E G N P E V G  
601 TCCCCGAAACCTCAGACCTCCAGCACCGGGGAAACAAGGGCTTTCGCTTCTCTCTCTCCCGTGAAGATGCGCAGCCAGGCGAGGAGCCATCAAGT  
11 S P R T S D L Q H R G N K G S C V L S S P G E D A Q P G E E P I K  
701 ATGGTGAACCTCATCGTCTGGGATGCTGTGAGGAAGGAGGTGAGGAAACCGAGGCTCAGAGAGGGGAAGTACTGGCCCAAGGGCACACAGCTGCTACAA  
44 Y G E L I V L G C C E E G G E E T E A Q R G E V T G P R A H S C Y N  
801 TGTTGTCTGGCAAGTGGGACAAAGGGCCCGGGCAAGCCGCTGGCACTGAGCCGCGGTGCGACGCCAACGGGGTGAAGCCAGACGTCATGCACCAC  
77 G C L A S G D K G R R R S R L A L S R R S H A N G V K P D V M H H  
901 ATCTCCAGCCGCTCTCAAGGCACTGAGTAACCGTGGTGCAGCAGCATCTCGTATACACTGTCCCGAGCCACTCGGTATAGTGGAGTATACAC  
111 I S T P L V S K A L S N R G Q H S I S Y T L S R S H S V I V E Y T  
1001 ATGATAGCGACACAGACATGTTCCAGATTGGCCGCTCCACAGAGAACATGATTGACTTCTGTGTAACAGACACGTCCTGAGGAGGGGCTGCCAGGG  
144 H D S D T D M F Q I G R S T E N M I D F V V T D T S P G G G A A E G  
1101 CCCTTCTGCCAGAGCACCATCTCCCGTATGCTGCCGATCTCTGTGACCGCCGCCACCCTATACTGCCCGCATCTATGCCGCTGGCTTCGATGCC  
177 P S A Q S T I S R Y A C R I L C D R R P P Y T A R I Y A A G F D A  
1201 TCTAGCAACATCTTCTTGAGAGCGAGCGGCCAAATGGCGGACCCAGATGGCTGATGGATGGACTGACCACCAATGGAGTCTGGTGTGCACCCGG  
211 S S N I F L G E R A A K W R T P D G L M D G L T T N G V L V M H P  
1301 CAGCGGCTTCTCCGAGGACTCAGCCCCGGGTGTGGCGGAGATCTCGTGTGGGAATGTGTACACATTGCGGGACAGCCGCTCAGCCAGCAGCG  
244 A G G F S E D S A P G V W R E I S V C G N V Y T L R D S R S A Q Q R  
1401 GGGCAAGCTGGTGGAAAACGAGTCCAACGTGTGCAGGACGGCTCTCTCATCGACCTGTGTGGGGCCACACTGCTGTGGCGCACACCGGGGGGCTGTG  
277 G K L V E N E S N V L Q D G S L I D L C G A T L L W R T P A G L L  
1501 CGGCTCCACACTGAAGCAACTGGAGGCCAGCGGAGGAGGAAATGCAGCGCGGCCAGTGCCCGTGGGCTCAGCACTCTGGCTTCCCGAGCC  
311 R A P T L K Q L E A Q R Q E A N A A R P Q C P V G L S T L A F P S

NcoI (1664)

1601 CAGCCCGTGGCCGACAGCGCCGACAAACAGCAGCCCTGGGTCTACGTCCGCTGCGGGCACGTCCATGGCTACCACGGCTGGGGCTGCCGGCGGGAGCG  
344 P A R G R T A P D K Q Q P W V Y V R C G H V H G Y H G W G C R R E R  
1701 GGGCCCCAGGAGCGGAATGCTCTCTGCGCCCTTGTGGGCTTATGTGCTCTATGGCTTGGCCAGGAGCGCGCTCTGCCTGGACCTGGGCGG  
377 G P Q E R E C P L C R L V G P Y V P L W L G Q E A G L C L D P G P  
1801 CCTAGCCATGCCTTTCACCTTGCAGCCGACGTCGCTCTGAGAAGACTGCCGCTACTGGGCCAGACACCACTGCCACCGCACCCATGCTTCCATG  
411 P S H A F A P C G H V C S E K T A R Y W A Q T P L P H G T H A F H  
1901 CCGCTGCCCTTTTGGGGGCTGGCTTACCGGCGAGCATGGCTGCTCCGCTCATTTCAGGGCCCGCTGGATTAGGCTCCCTGGGGCCCCCTGCT  
444 A A C P F C G A W L T G E H G C V R L I F Q G P L D •

NheI (2039)

2001 GCTGTGCCACCTGCCACCCAGGTCCCCACCTCCTGCAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGC  
2101 AGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACATTCATTCA  
2201 TTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAAAC  
2301 TTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGC  
2401 CTCACCTTCTTTCATGGAGTTTAAAGATATAGTATTTTCCCAAGGTTTGAACCTAGCTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTC  
2501 CCTTTTATAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAA  
2601 TATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTT  
141 • N R T Y K  
2701 GAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCAC  
135 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 S I L E R C  
2801 ATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCCTTCTGCCGTTGCTCA  
101 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 K Q G N S V  
2901 CAGCAGCCCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTGGTCTGAT  
68 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 G T K T R I  
3001 GGCCGCCCCGACATGGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAAGTCT  
35 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 Q S I N F T  
3101 TTCATGTTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATC  
1 K M

3201 TGACGGTTCATAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTCGCTCAATGGGGCGGAGTTGTTACGACATTTTGAA  
3301 AGTCCCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGAT  
3401 GTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAAGTCATGTACTGGGCATAATGC  
3501 CAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCA  
3601 CCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGG  
3701 GCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAAITAAAGACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGG  
3801 CGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTT  
3901 TCCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT  
4001 AGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCG  
4101 GTAACATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC  
4201 TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTT  
4301 GGTAGCTCTTGATCCGGCAAACAACCCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATC  
4401 CTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGC  
4501 AATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACATAACGCTCTCCATCAAACAAAACGAAACAAAACAAC  
4601 TAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA