



1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

**AgeI (552) SphI (568)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTAGGAGGGCCAGCATGCTACGAGCTCTTTGGCTCTTCTGGATC  
1▶ M L R A L W L F W I  
601 TTGGTGGCCATAACAGTCTCTCAGCAAACGCTGTTCTGTCTCAGGAGTCTCTGTCTATGTGATGCTTCTGGGGTGTGTGATGGCCGCTCCAGTCTTTCA  
11▶ L V A I T V L F S K R C S A Q E S L S C D A S G V C D G R S R S F  
701 CCTCTATCCCTCCGACTCACAGCAGCATGAAAAGCCTTGACCTGTCTTCAACAAGTACCTACATTGGCCATGGTACCTCCGAGCGTGTGCGAA  
44▶ T S I P S G L T A A M K S L D L S F N K I T Y I G H G D L R A C A N  
801 CCTCCAGTTCTGATTTTGAAGTCCAGCAGAATCAATACAATAGAGGGAGACGCTTTTATTCTCTGGGCAGTCTTGAACATTTGGATTTGTCTGATAAT  
77▶ L Q V L I L K S S R I N T I E G D A F Y S L G S L E H L D L S D N  
901 CACCTACTAGTTTATCTCCTCCTGGTTCGGGCCCCCTTCTCTTTGAAATACTTAAACTAATGGGAAATCCTTACCAGACTGGGGTAACTACGCG  
111▶ H L S S L S S W F G P L S S L K Y L N L M G N P Y Q T L G V T S  
1001 TTTTCCCAATCTCACAATTTACAAACCTCAGGATAGGAAATGTAGAGACTTTCAGTGAGATAAGGAGAATAGATTTTGTGGGCTGACTTCTCTCAA  
144▶ L F P N L T N L Q T L R I G N V E T F S E I R R I D F A G L T S L N  
1101 TGAAGTGAATTAAGGCATTAAGTCTCCGGAATTATCAGTCCCAAAGTCTAAAGTCGATCCGCGACATCCATCACCTGACTTTCACCTAAGCGAGTCT  
177▶ E L E I K A L S L R N Y Q S Q S L K S I R D I H H L T L H L S E S  
1201 GCTTCTGCTGGAGATTTTGCAGATATTCTGAGTCTGTGAGATATTTAGAATAAGAGATACTAACTGGCCAGGTTCCAGTTTCCACTGCCCCG  
211▶ A F L L E I F A D I L S S V R Y L E L R D T N L A R F Q F S P L P  
1301 TAGATGAAGTCAGCTCACCGATGAAGAAGCTGGCATTCCGAGGCTCGGTTCTACTGATGAAAGCTTAAACGAGCTCCTGAAGCTGTTGCGTTACATCTT  
244▶ V D E V S S P M K K L A F R G S V L T D E S F N E L L K L L R Y I L  
1401 GGAAGTGTGCGAGGTAGAGTTCGACGACTGTACCCTCAATGGGCTCGGCGATTTCAACCCCTCGGAGTCAAGCAGTGTGAGCGAGCTGGGTAAGTAGAA  
277▶ E L S E V E F D D C T L N G L G D F N P S E S D V V S E L G K V E  
1501 ACAGTCACTATCCGAGGTTGCATATCCCCAGTCTATTTGTTTTATGACCTGAGTACTGTCTATCCCTCTGGAGAAGGTGAAGCGAATCACAGTAG  
311▶ T V T I R R L H I P Q F Y L F Y D L S T V Y S L L E K V K R I T V  
1601 AGAACAGCAAGTCTTCTGGTTCCTGCTGTTCTCCAGCATTTAAATCATTAGAATCTTAGACCTCAGCGAAAATCTGATGGTTGAAGAATATTT  
344▶ E N S K V F L V P C S F S Q H L K S L E F L D L S E N L M V E E Y L  
1701 GAAGAAGTCAAGCTGAAGGGAGCCTGGCCTTCTCAAAACCTTAGTTTGGAGCAGAATCATTGAGATCAATGCAAAAAACAGGAGAGATTTTGTG  
377▶ K N S A C K G A W P S L Q T L V L S Q N H L R S M Q K T G E I L L  
1801 ACTCTGAAAACTGACCTCCCTTGACATCAGCAGGAACACTTTTTCATCCGATGCCGACAGTGTGAGTGGCCAGAAAAGATGCGCTTCTGAAATTTGT  
411▶ T L K N L T S L D I S R N T F H P M P D S C Q W P E K M R F L N L  
1901 CCAGTACAGGGATCCGGGTGGTAAAAACGTGCATTCTCAGACGCTGGAGGTGTTGGATGTTAGTAACAACAATCTTGACTCATTTTCTTTGTTCTTGCC  
444▶ S S T G I R V V K T C I P Q T L E V L D V S N N N L D S F S L F L P  
2001 TCGGTCGAAGACTCTATATTTCCAGAAATAAGTCAAAAACACTCCAGATGCTTCCGTTGTTCCCTGTGTTGCTGGTCAAAAATCAGAGAGAATGCA  
477▶ R L Q E L Y I S R N K L K T L P D A S L F P V L L V M K I R E N A  
2101 GTAAGTACTTTCTCAAAGACCAACTGGTCTTTTCCAACTGGAGACTCTGGAAGCAGGCGACAACCCTTTGTTTGTCTCTGCAACTCCTATCCT  
511▶ V S T F S K D Q L G S F P K L E T L E A G D N H F V C S C E L L S  
2201 TTACTATGGAGACCCAGCTCTGGCTCAAATCCTGGTTGACTGGCCAGACAGTACCTGTGTGACTCTCCGCTCGCTGCACGGCCACAGGCTTCCAGGA  
544▶ F T M E T P A L A Q I L V D W P D S Y L C D S P P R L H G H R L Q D  
2301 TGCCCGCCCTCCGCTTGGAAATGTCACCAGGCTGCACTGGTGTCTGGAGTCTGCTGTGCCCTTCTCCTGTTGATCTTGTCTGAGGTGCCCTGTGCCAC  
577▶ A R P S V L E C H Q A A L V S G V C C A L L L L I L L V G A L C H  
2401 CATTTCCACGGACTGTGGTACCTGAGAATGATGTGGGCGTGGCTCCAGGCCAAGGAAGCCAAAGAAAGTCCCTGCAGGACGTTGGATCTACCCCT  
611▶ H F H G L W Y L R M M W A W L Q A K R K P K K A P C R D V G S Y P

**NheI (2524)**

2501 ATGATGTGCCAGACTACGCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTAAAAAATGCT  
644▶ Y D V P D Y A •  
2601 TTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAAC AATTGCATTATTTATGTTTCAGGT  
2701 TCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGG AATTCTAAAATACAGCATAGCAAAACTTTAACCTCAAATC  
2801 AAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGACGCTCACCTCTTTTCAT  
2901 GGAGTTTAAAGATATAGTGTATTTTCCAAAGTTTGAAGTACTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATAGTAAAT  
3001 ATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAG  
3101 TAGTTGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCT  
141▶ N R T Y K L P I L E

3201 CTCAATGGTGGTTTTGACCAGCTTGCCATTTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTG  
130 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 M G C P S  
3301 ACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGG  
96 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 A S G I A  
3401 CAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGTATGGCCGCCCGACATG  
63 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  
3501 GTGCTTGTGCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTC  
30 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 K M  
3601 CTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAA  
CGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTA  
3701 CTAGTCAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCG  
3801 CATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCACTGTACTGGGCATAATGCCAGGCGGGCCATTTA  
3901 CCGTCATTGACGTCAATAGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAAT  
4001 GGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAG  
4101 TTATGTAACGCCTGCAGGTTAATAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGC  
4201 TCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTTCCCCTGGAAGCTC  
4301 CCTCGTGCCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG  
4401 TATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGTGGCCTTATCCGGTAACTATCGTCTTG  
4501 AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAA  
4601 GTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTAGCTCTTGATCC  
4701 GGCAAACAACACCAGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA  
4801 CGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAATCAGCGGCCCAATAAAATATCTTTA  
4901 TTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAACGAAACAAAACAACTAGCAAATAGGCTG  
5001 TCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTATCGAA  
5101