



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTGTGCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

**SphI (560)**

**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCAGCATGCCCTGCCCGCCCGGAGGTGAAAGCGGGTGTGA  
13▶ M P P A P P G G E S G C E  
601 GGAGCGCGGCGCGCAGGTCATATTGAACATCCAGATACCTATCATTACTCGATGCTGTTGATAACAGCAAGATGGCTTTGAACTCAGGGTCACCACCA  
13▶ E R G A A G H I E H S R Y L S L L D A V D N S K M A L N S G S P P  
701 GCTATTGGACCTTACTATGAAAACCATGGATACCAACCGGAAAACCCCTATCCCGCACAGCCACTGTGGTCCCCTGTCTACGAGGTGCATCCGGCTC  
47▶ A I G P Y Y E N H G Y Q P E N P Y P A Q P T V V P T V Y E V H P A  
801 AGTACTACCGTCCCGTGGCCAGTACGCCCCGAGGGTCTGACGAGGCTTCAACCCCGTCTGTCACGCAGCCAAATCCCCATCCGGGACAT  
80▶ Q Y Y P S P V P Q Y A P R V L T Q A S N P V V C T Q P K S P S G T V  
901 GTGCACCTCAAAGACTAAGAAAGCACTGTGCATCACCTTGACCTGGGACCTTCTCGTGGGAGCTGCGCTGGCCGCTGGCCTACTCTGGAAGTTTCATG  
113▶ C T S K T K K A L C I T L T L G T F L V G A A L A A G L L W K F M  
1001 GGCAGCAAGTGTCCAACCTGGGATAGAGTGCAGTCTCAGTACCTGCATCAACCCCTAACTGGTGTGATGGCGTGCACACTGCCCGGGGGG  
147▶ G S K C S N S G I E C D S S G T C I N P S N W C D G V S H C P G G  
1101 AGGACGAGAATCGGTGTGCTCGCCTCTACGGACCAAATTCATCCTTCAGGTGACTCATCTCAGAGGAAGTCTGGCACCTGTGTGCCAAGACGACTG  
180▶ E D E N R C V R L Y G P N F I L Q V Y S S Q R K S W H P V C Q D D W  
1201 GAACGAACTACGGGCGGGCGGCTGCAGGGACATGGGCTATAAGAATAATTTTTACTCTAGCCAAGGAATAGTGGATGACAGCGGATCCACCAGCTTT  
213▶ N E N Y G R A A C R D M G Y K N N F Y S S Q G I V D D S G S T S F  
1301 ATGAAACTGAACACAAGTCCCGCAATGTCGATATCTATAAAAACTGTACCACAGTATGCTGTTCTTCAAAGCAGTGGTTTCTTTACGCTGTATAG  
247▶ M K L N T S A G N V D I Y K K L Y H S D A C S S K A V V S L R C I  
1401 CCTGCGGGTCAACTGAACTCAAGCCGCCAGAGCAGGATTGTGGGCGGAGAGCGCTCCCGGGGCTGGCCCTGGCAGGTGACCTGCACGTCCA  
280▶ A C G V N L N S S R Q S R I V G G E S A L P G A W P W Q V S L H V Q  
1501 GAACGTCCACGTGTGCGGAGGCTCCATCATACCCCGAGTGGATCGTGCAGCCGCCACTGCGTGGAAAACTCTTAACAATCCATGGCATTGGACG  
313▶ N V H V C G G S I I T P E W I V T A A H C V E K P L N N P W H W T  
1601 GCATTTGCGGGATTTTGAGACAATCTTTCATGTTCTATGGACCCGATACCAAGTAGAAAAAGTATTCTCATCCAATTAATGACTCCAAGACCAAGA  
347▶ A F A G I L R Q S F M F Y G A G Y Q V E K V I S H P N Y D S K T K  
1701 ACAATGACATTGCGTGTGAAGCTGCAGAAGCCTCTGACTTTCAACGACCTAGTGAACAGTGTGTCTGCCAACCAGGCATGATGCTGCAGCCAGA  
380▶ N N D I A L M K L Q K P L T F N D L V K P V C L P N P G M M L Q P E  
1801 ACAGCTCTGCTGATTTCCGGTGGGGGCCACCGAGGAGAAAGGAAAGACCTCAGAAGTGTGAACGCTGCCAAGGTGCTTCTCATTGAGACACAGAGA  
413▶ Q L C W I S G W G A T E E K G K T S E V L N A A K V L L I E T Q R  
1901 TGCAACAGCAGATGTCTATGACAACCTGATCACACCAGCCATGATCTGCGCGCTTCTGCGAGGGAACTGCGATTCTTCCAGGGTGCAGTGGAG  
447▶ C N S R Y V Y D N L I T P A M I C A G F L Q G N V D S C Q G D S G  
2001 GGCCTCTGGTCACTTCGAAGAACAATATCTGGTGGCTGATAGGGGATACAAGCTGGGGTCTGGCTGTGCCAAAGCTTACAGACCAGGATGTACGGGAA  
480▶ G P L V T S K N N I W W L I G D T S W G S G C A K A Y R P G V Y G N

**NheI (2155)**

2101 TGTGATGGTATTCACGGACTGGATTTATCGACAAATGAGGGCAGACGGCTAATCCGCTAGTGGCCAGACATGATAAGATACATTGATGAGTTGGACAA  
513▶ V M V F T D W I Y R Q M R A D G •  
2201 ACCACAAC TAGAATGCAGTGA AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACA  
2301 ACAACAATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATCTAAAA  
2401 TACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGC  
2501 ATTAGCTGTTGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTCTCTCATTCTTTATGTTTTAAATGCA  
2601 CTGACCTCCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGC  
2701 TCAAGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTT  
141 •  
2801 AGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGA  
140▶ N R T Y K 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  
2901 GATGAGCTCTCTGCATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCC  
107▶ I L E R C 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  
3001 TTCTGCCGTTGCTCACAGCAGCCAAATGGCAATGGCTTCAGCACAGACAGTGCACCTGCAATGTAGGCCCTCAATGTGGACAGCAGAGATGATCTCCC  
73▶ K Q G N S V 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  
3101 CAGTCTGGTCTGATGGCCGCCCCGACATGGTGTGTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGSCTCCAGATCCTGCTG  
40▶ T K T R I 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

3201 AGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATG  
7 S I N F T K M  
3301 GCGTCTCCAGCTTATCTGACGGTTCACATAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTG  
3401 TTACGACATTTTGAAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAAATCCCGTGAGTCAAACCGCT  
3501 ATCCACGCCCATTTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAAGTCAT  
3601 GTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTT  
3701 TACCGTAAATACTCCACCCATTGACGTCAATGGAAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGG  
3801 GCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAA  
3901 AAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGACTAT  
4001 AAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGAAG  
4101 CGTGGCGCTTTTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCCGAC  
4201 CGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA  
4301 GGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC  
4401 CTTGGAAGAAAGAGTTGGTAGCTCTTGTATCCGGCAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAA  
4501 GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACAT  
4601 TTAATCAGCGGCCGAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAA  
4701 ACGAAAACAAAACAACTAGCAAAATAGGCTGTCCCGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA