



100

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

**BspHI (560)**

**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTATCATGAACCCACAACGGGAGGCAGCACCCAAATCCTATGC  
1 M N P Q R E A A P K S Y A  
601 TATTCGTGATTCTCGACAGATGGTGTGGGTCTGAGTGGAAATCTTTAATAGCAGCTCCTCTTAGCCGAGCATTAAAGCCTGCACTCTTCAATTAATA  
13 I R D S R Q M V W V L S G N S L I A A P L S R S I K P V T L H L I  
701 GCCTGTAGAGACACAGAATTCAGTGACAAGGAAAGGGTAATATGGTTTACCTGGGAATCAAGGGAAAAGATCTGTCTCTTCTGTGCAGAAATTCAGG  
47 A C R D T E F S D K E K G N M V Y L G I K G K D L C L F C A E I Q  
801 GCAAGCCTACTTTGCAGCTTAAGCTTCAGGGCTCCCAAGATAACATAGGGAAGGACACTTGTGAAAAGTGTGAAATTCACACATGCATAAACCTGGA  
80 G K P T L Q L K L Q G S Q D N I G K D T C W K L V G I H T C I N L D  
901 TGTGAGAGAGAGCTGCTTCATGGGAACCTTGACCAATGGGAATAGGAGTGGGTAGAAAGAAGTGGAAAGAGTTCCTTTCAACATCACCATCTCAGGAAG  
113 V R E S C F M G T L D Q W G I G V G R K K W K S S F Q H H H L R K

**NheI (1057)**

1001 AAGGACAAAGATTTCTCATCCATGCGGACCAACATAGGAATGCCAGGAAGGATGTAGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGAC  
147 K D K D F S S M R T N I G M P G R M •  
1101 AAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAA  
1201 CAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATCTAA  
1301 AATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGT  
1401 GCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATG  
1501 CACTGACCTCCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAAATCCAGAT  
1601 GCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTTGGACAGCAAGAAAGCGAGCTTCTAGCT  
1701 TTAGTTCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCA  
141 • 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  
1801 GAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGT  
107 S 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  
1901 CCTTCTGCCCCTGCTCAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTC  
74 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  
2001 CCCAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTCTTGTCTCATAGAGCAGTGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCCTGC  
41 G 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  
2101 TGAGAGATGTTGAAGGCTTCATGGTGGCCCTCTATAGTGTGATGCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGAA  
7 Q S I N F T K M  
2201 TGGCGTCTCCAGTTATCTGACGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTGCGTCAATGGGGCGGAGT  
2301 TGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCG  
2401 CTATCCACGCCATTGATGTAAGTCCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCATAAGGTC  
2501 ATGTAAGTGGGCATAATGCCAGGCGGGCCATTTACCCTGATGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTGTAAGTCCCAAGTGGGCAG  
2601 TTTACCCTAAATACTCCACCCATTGACGTCAATGGAAGTCCCTATTGGCGTACTATGGAAACATACGTCAATATTGACGTCAATGGGCGGGGGTCTGTT  
2701 GGGCGGTCAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACCGCTGACGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA  
2801 AAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGGACT  
2901 ATAAAGATACCAGGCTTTCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCAGGATACCTGTCGCTTTCTCCCTTCGGGA  
3001 AGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCG  
3101 ACCGCTGCGCTTATCCGGTAATATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC

3201 GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT  
3301 ACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGCAAGCAGCAGATTACGCGCAGAAAAA  
3401 AAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAAC  
3501 ATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACA  
3601 AAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA