



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

**BspHI (560)**

**AgeI (552)**

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCATCATGAAGTGCTTGTGATCTCCCTAGCCCTATGGCTGGG  
1 M K C L L I S L A L W L G  
601 CACAGTGGGCACACGTGGGACAGAGCCGAAGTCCAGCGAGACCCAGCGCAGGAGCCTACAGGTGGCTCTGGAGGAGTCCACAAAACCCACCTGTGCAG  
13 T V G T R G T E P E L S E T Q R R S L Q V A L E E F H K H P P V Q  
701 TTGGCCTTCCAAGAGATCGGTGTGGACAGAGCTGAAGAAGTGTCTTCTCAGCTGGCACCTTTGTGAGGTTGGAATTAAGCTCCAGCAGACCAACTGCC  
47 L A F Q E I G V D R A E E V L F S A G T F V R L E F K L Q Q T N C  
801 CCAAGAAGGACTGAAAAAGCCGGAGTGCACAAATAAACCAACCGGAGAAGGCGAAATGCCTGGCTGCATTAAAAATGGACCCCAAGGGTAAAAATCT  
80 P K K D W K K P E C T I K P N G R R R K C L A C I K M D P K G K I L  
901 AGGCCGGATAGTCCACTGCCAATTCTGAAGCAAGGCCTCAGGATCCTCAGGAGTTGCAATGCATTAAGATAGCACAGGCTGGCGAAGACCCCCACGGC  
113 G R I V H C P I L K Q G P Q D P Q E L Q C I K I A Q A G E D P H G

**NheI (1062)**

1001 TACTTCTACCTGGACAGTTTGCCTTCTCCAGGGCCCTGAGAACCAATAAGCCCTAGACAGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTT  
147 Y F L P G Q F A F S R A L R T K •  
1101 TGGACAAACCACAACCTAGAATGCAGTAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAA  
1201 GTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAAT  
1301 TCTAAATACAGCATAGCAAACTTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCC  
1401 AATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGTTTGAAGTACTGCTTTCATTTCTTTATGTTTT  
1501 AAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTATTAGGCAGAATC  
1601 CAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTC  
1701 TAGCTTTAGTTCTGGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCAT  
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  
1801 AGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGT  
109 D 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  
1901 AAAGTCTTCTGCCCCTGCTCAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATG  
76 F D 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  
2001 ATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTCTCATAGAGCAGTGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGAT  
42 I E 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  
2101 CCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCCTATAGTGTGCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGC  
9 Q Q S I N F T K M  
2201 GTGGATGGCGTCTCCAGCTTATCTGACGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTGCGTCAATGGGGC  
2301 GGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCA  
2401 AACCGCTATCCACGCCATTGATGTAAGTCCGATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATA  
2501 AGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCCTGATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTG  
2601 GGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAGTCCCTATTGGCGTACTATGGAAACATACGTCAATATTGACGTCAATGGGCGGGG  
2701 TCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCTGCAGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA  
2801 CCGTAAAAAGGCCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAGCAAAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACA  
2901 GGACTATAAGATACCAGGCTTTCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTT  
3001 CGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTC  
3101 GCGGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC

3201 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGC  
3301 CAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGCAAGCAGCAGATTACGCGCAG  
3401 AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTGGTCATGGCTAGTTAA  
3501 TTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCA  
3601 AAACAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA