



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

**BspHI (560)**  
**AgeI (552)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCATCATGAAGATCTTCTTCCAGTGCCTGGCTGCCCTTCT  
131 M K I F L P V L L A A L L  
601 GGTGTGGAGCGAGCCAGCTCGCTGATGTGCTTCTCTGCTTGAACAGAAAGCAATCTGTACTGCCTGAAGCCGACCATCTGCTCCGACCAGGACAAAC  
131 G V E R A S S L M C F S C L N Q K S N L Y C L K P T I C S D Q D N  
701 TACTGCGTGACTGTGTCTGCTAGTGCCGGCATTGGGAATCTCGTGACATTTGGCCACAGCCTGAGCAAGACCTGTTCCCGGCTGCCCATCCAGAAG  
471 Y C V T V S A S A G I G N L V T F G H S L S K T C S P A C P I P E  
801 GCGTCAATGTTGGTGTGGCTTCCATGGGCATCAGCTGCTGCCAGAGCTTTCTGTGCAATTTAGTGCGGCCGATGGCGGGCTCGGGCAAGCGTCACCT  
801 G V N V G V A S M G I S C C Q S F L C N F S A A D G G L R A S V T L

**NheI (963)**

901 GCTGGGTGCCGGGCTGCTGCTGAGCCTGCTGCCGGCCTGCTGCGGTTTGGCCCTGACCGCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGT  
1131 L G A G L L L S L L P A L L R F G P •  
1001 TTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACA  
1101 AGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTAAAGCAAGTAAAACCTTACAAATGTTGATGGAA  
1201 TTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGC  
1301 CAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGTGTGAAGTACTGCTTTCATTTCTTTATGTTT  
1401 TAAATGCACTGACCTCCACATTCCTTTTTAGTAAAAATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAATGTTTTTTATTAGGCAGAAT  
1501 CCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTT  
1601 CTAGCTTTAGTCTCTGGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAAGCAGTCAGGAGCA  
1411 • 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  
1701 TAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGT  
1091 Y 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  
1801 CAAAGCTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCAATGTAGGCTCAATGTGGACAGCAGAGAT  
761 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  
1901 GATCTCCCGAGTCTTGGTCTGATGGCCGCCGACATGGTGTGTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGA  
431 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  
2001 TCCTGTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGTGATTACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAG  
911 D Q Q S I N F T K M  
2101 CGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTGCGTCAATGGGG  
2201 CGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTC  
2301 AAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCATGTTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCCAT  
2401 AAGGTCATGACTGGCATAATGCCAGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTGCAAGT  
2501 GGGCAGTTTACCATAACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGCGGGG  
2601 GTCGTTGGCGGTGAGCCAGCGGGCCATTTACCGTAAGTTATGTAACGCTGACGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGA  
2701 ACCGTA AAAAGGCCGCTTGTGGCTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGAC  
2801 AGGACTATAAAGATACCAGCGTTTCCCGTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCT  
2901 TCGGAAGCGTGGCGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTGCTCCAAGTGGGTGTGTGCACGAACCCCGTTT  
3001 AGCCCGACCGCTGCGCTTATCCGTAACCTATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG  
3101 CAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTGCTGAAAG

3201 CCAGTTACCTTCGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGGCA  
3301 GAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTA  
3401 ATTAACATTTAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATC  
3501 AAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA