



100

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

NcoI (560)  
BstEII (555)  
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCCGCTACCTGAGATCACCGGTACCATGGATCCCAAGGGGCTCCTCTCCTTACCTTCGTGCT  
601 GTTTCTCTCCCTGGCTTTTGGGCAAGCTACGGAACAGGTGGGCGCATGATGAACTGCCCAAAGATTCTCCGGCAGTTGGGAAGCAAAGTGTGCTGCC  
13▶ F L S L A F G A S Y G T G G R M M N C P K I L R Q L G S K V L L P  
701 CTGACATATGAAAGATAAATAAGAGCATGAACAAAAGCATCCACATTGTCGTACAATGGCAAATCACTGGAGAACAGTGTGAGAACAAAATAGTGT  
47▶ L T Y E R I N K S M N K S I H I V V T M A K S L E N S V E N K I V  
801 CTCTTGATCCATCCGAAGCAGGCCCTCCACGTTATCTAGGAGATCGTACAAGTTTTATCTGGAGAATCTCACCTGGGGATACGGGAAAGCAGGAAGGA  
80▶ S L D P S E A G P P R Y L G D R Y K F Y L E N L T L G I R E S R K E  
901 GGATGAGGATGTACCTTATGACCCTGGAGAAAAATGTTTCAGTTCCAGCGCTTTGCTGCAAGTTGAGGCTTATGAGCAGTCTCCACTCCAGAAATT  
113▶ D E G W Y L M T L E K N V S V Q R F C L Q L R L Y E Q V S T P E I  
1001 AAAGTTTTAAACAAGACCAGGAGAACGGGACCTGCACCTTGATACTGGGCTGCACAGTGGAGAAGGGGGACCATGTGGCTTACAGCTGGAGTGAAAAGG  
147▶ K V L N K T Q E N G T C T L I L G C T V E K G D H V A Y S W S E K  
1101 CGGGCACCCACCCACTGAACCCAGCCAAACAGCTCCACCTCTGTCCCTCACCTCGGCCCCAGCATGCTGACAATATCTACATCTGCACCGTGAGCAA  
180▶ A G T H P L N P A N S S H L L S L T L G P Q H A D N I Y I C T V S N

NcoI (1270)

1201 CCCTATCAGCAACAATTCCAGACCTTCAGCCCGTGGCCCGGATGCAGGACAGACCCCTCAGAAAACAAACCATGGGCAGTGTATGCTGGGCTGTAGGG  
213▶ P I S N N S Q T F S P W P G C R T D P S E T K P W A V Y A G L L G  
1301 GGTGTCATCATGATTCTCATCATGGTGGTAATACTACAGTTGAGAAGAAGAGGTAAAACGAACATTACCAGACAACAGTGGAAAAAAAAGCCTTACGA  
247▶ G V I M I L I M V V I L Q L R R R G K T N H Y Q T T V E K K S L T  
1401 TCTATGCCAAGTCCAGAAACCAGGTCCTCTTCAGAAGAACTTGACTCCTTCCAGCTCAGGACCCTTGACCACCATATATGTTGCTGCCACAGGCC  
280▶ I Y A Q V Q K P G P L Q K K L D S F P A Q D P C T T I Y V A A T E P

NheI (1589)

1501 TGCCAGAGTCTGTCCAGGAACAAATTCATCACAGTCTATGCTAGTGTGACACTCCAGAGAGCTGACACCAGAGACCAACAAAGGGCTAGCTGGCC  
313▶ V P E S V Q E T N S I T V Y A S V T L P E S •  
1601 AGACATGATAAGATACATTGATGAGTTTGGACAACCACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTT

1701 GTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTCCAGGGGAGGTGGGGAGTTTTTAAAGCAAGT  
1801 AAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGAAAACCTTAAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAAT  
1901 AAGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTCATGGAGTTTAAAGATATAGTATTTTCCCAAGGTTTG

2001 AACTAGCTCTTCAATTTCTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAA  
2101 AATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAG

2201 AAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCT  
141▶ • N R T Y K L P I L E E I T T K V L K G N M E  
2301 CAATGAGCACAAGCAGTCAGGAGCATAGTCTGACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCACCTCATCAGAGTA  
118▶ I L V F C D P A 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  
2401 GGGTGCCTGACAGCCAAATGGTGTCAAAGTCCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTACAGCAGACAGTACCCTGCCAATG  
85▶ P H R V A V I T D F D K Q G N S V A S G I A I A E A C C V T V R G I  
2501 TAGGCCTCAATGTGACAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTGTCCTCATAGAGCATGGTATCTTCT  
51▶ Y A E I H V A S 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  
2601 CAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCAAGTGGCCCTCTATAGTGAAGTCTTATGCTGATGCTGATGCTGAT  
18▶ T A V E V L E L D Q Q S I N F T K M  
2701 ATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACAG

2801 CCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATGACGTCAATGGGG  
2901 TGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCCAAACCGCATCATGTAATAGCGATGACTAATACGTAGA  
3001 TGTAAGTCCAAAGTAGGAAAGTCCATAAGTTCATGTAAGTGGGCGTAAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCAT

3101 ATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATAC  
3201 GTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTG  
3301 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGC  
3401 TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATAACCAGGCGTTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTCCGACCCTGCCGCTTA  
3501 CCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCT  
3601 GGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG  
3701 GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAG  
3801 TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTT  
3901 TGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACTCACGT  
4001 TAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGA  
4101 ATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTC  
4201 TATCGAA