



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspHI (557)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTATGATGCCTGGGAGCCCTCGGCCGCGCAAGCTGGGTGCT
601 GTTGCTGCGGCTGCTGCGGTTGTGCGGCCCGGGGCTGGGTGAGGCATGCAGCTGCGCCCCGGCGCACCCCTCAGCAGCACATCTGCCACTCGGCACTT
13▶ L L R L L A L L R P P G L G E A C S C A P A H P Q Q H I C H S A L
701 GTGATTCGGGCCAAAATCTCCAGTGAGAAGGTAGTTCGGCCAGTGCAGACCCTGCTGACACTGAAAAATGCTCCGGTATGAAATCAAACAGATAAAGA
47▶ V I R A K I S S E K V V P A S A D P A D T E K M L R Y E I K Q I K
801 TGTTCAAAGGTTTGAAGAAAGTCAAGGATGTTCAAGTATCTATACGCCTTTTACTCTTCCCTCTGTGGTGTGAAACTAGAAGCCAACAGCCAGAAGCA
80▶ M F K G F E K V K D V Q Y I Y T P F D S S L C G V K L E A N S Q K Q
901 GTATCTTTGACTGGTCAGGTCCTCAGTGTGAAAAGTCTTCATCCATCTGTGCAACTACATCGAGCCCTGGGAGGACCTGTCTTGGTGCAGAGGGAA
113▶ Y L L T G Q V L S D G K V F I H L C N Y I E P W E D L S L V Q R E
1001 AGTCTGAATCATCACTACCATCTGAACTGTGGCTGCCAAATACCACCTGCTACACAGTACCCTGTACCATCTCGGCCCTAACGAGTGCCTCTGGACAG
147▶ S L N H H Y H L N C G C Q I T T C Y T V P C T I S A P N E C L W T
1101 ACTGGCTGTTGGAACGAAAGCTCTATGGTTACCAGGCTCAGCATTATGTCTGTATGAAGCATGTTGACGGCACCTGCAGCTGGTACCGGGGCCACCTGCC
180▶ D W L L E R K L Y G Y Q A Q H Y V C M K H V D G T C S W Y R G H L P

NheI (1236)

1201 TCTCAGGAAGGAGTTTGTGACATCGTTCAGCCCTAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGT
213▶ L R K E F V D I V Q P •
1301 GAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATT
1401 TATGTTTCAGGTTAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTT
1501 AACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCCTC
1601 ACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGTGTGAACTAGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTC
1701 TTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATAT
1801 CCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAG
1901 GGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATG
134▶ 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 I L E R C M
2001 CCACAGGGGCTGACCACCTGATGGATCTGCCACCTCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAG
100▶ 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 K Q G N S V A
2101 CAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGC
67▶ 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 T K T R I A
2201 CGCCCCGACATGGTGTGTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTC
34▶ 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 S I N F T K
2301 ATGGTGGCCCTCTATAGTGAGTTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGA
0▶ M
2401 CGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGT
2501 CCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTA
2601 CTGCCAAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGTAAGTACGCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGCATAATGCCAG
2701 GCGGGCATTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCGAGTTTACCCTAAATACTCCACCC
2801 ATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTACGCCAGGCGGGCC
2901 ATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGT
3001 TTTTCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGGACTATAAAGTACCAGGCGTTTCC
3101 CCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGC

3201 TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTA
3301 ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC
3401 AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGT
3501 AGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTT
3601 TGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAATCAGCGGCCGCAAT
3701 AAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAG
3801 CAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA