



100

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGTGCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGAAACTCTTTCTAATGCAAGTGGTACTTTTGCAT
1 M E T L S N A S G T F A I
601 ACGCCTTTTAAAGATACTGTGTCAAGATAACCTTCGCACAACGTGTTCTGTTCTCCTGTGAGCATCTCCTTGCCTGGCCATGGTCTCCTAGGGGCA
13 R L L K I L C Q D N P S H N V F C S P V S I S S A L A M V L L G A
701 AAGGGAACACCCGAACCCAGATGGCCAGGCACTGTCTTTAAACACAGAGGAAGACATTCATCGGGCTTCCAGTCGCTTCTCACTGAAGTGAACAAGG
47 K G N T A T Q M A Q A L S L N T E E D I H R A F Q S L L T E V N K
801 CTGGCACACAGTACCTGTGAGAACGGCAACAGGCTCTTTGGAGAGAAAATGTGAGTCTCCTCAACGTTTAAAGGAATCCTGTCTCAATTCTACCA
80 A G T Q Y L L R T A N R L F G E K T C Q F L S T F K E S C L Q F Y H
901 TGCTGAGCTGAAGGAGCTTCTCTTATCAGAGCTGCAGAAGAGTCCAGGAAACACATCAACACCTGGGTCTCAAAAAAGACCGAAGGTAAAATTGAAGAG
113 A E L K E L S F I R A A E E S R K H I N T W V S K K T E G K I E E
1001 TTGTTGCCGGTAGCTCAATTGATGCAGAAACAGGCTGGTCTTGTCAATGCCATCTACTTCAAAGGAAAGTGAATGAACGTTTACGAAACATACA
147 L L P G S S I D A E T R L V L V N A I Y F K G K W N E P F D E T Y
1101 CAAGGAAATGCCCTTAAAAATAAACAGGAGGAGCAAAGGCCAGTGCAGATGATGTATCAGGAGGCCACGTTAAGCTCGCCACGTTGGCGAGGTGCG
180 T R E M P F K I N Q E E Q R P V Q M M Y Q E A T F K L A H V G E V R
1201 CGCGCAGCTGTGAGCTGCCTTACGCCAGGAAGGAGCTGAGCCTGCTGGTGTCTGCTGACGACGGCGTGGAGCTCAGCACGGTGAAAAAAGTCTC
213 A Q L L E L P Y A R K E L S L L V L L P D D G V E L S T V E K S L
1301 ACTTTTGAGAAACTCACAGCCTGGACCAAGCCAGACTGTATGAAGAGTACTGAGGTTGAAGTCTCCTTCCAAAATTTAAACTACAAGAGGATTATGACA
247 T F E K L T A W T K P D C M K S T E V E V L L P K F K L Q E D Y D
1401 TGAATCTGTGCTTCGGCATTGGGAATTGTTGATGCCTTCCAACAGGGAAGGCTGACTTGTGGCAATGTCAGCGGAGAGAGACCTGTGTCTGTCCAA
280 M E S V L R H L G I V D A F Q Q G K A D L S A M S A E R D L C L S K
1501 GTTCGTGCACAAGAGTTTTGTGGAGGTGAATGAAGAAGGCACCGAGGCAGCGGAGCTGAGCTGCTTTGTAGTTGCAGAGTGTGCATGGAATCTGGC
313 F V H K S F V E V N E E G T E A A A A S S C F V V A E C C M E S G
1601 CCCAGTTCTGTGCTGACCACCTTTCTTTTCTTATCAGGCACAACAGAGCAACAGCATTCTGTTCTGTGGCAGGTTCTCATGCCATAAAGGGTGC
347 P R F C A D H P F L F F I R H N R A N S I L F C G R F S S P •

NheI (1741)

1701 ACTTACCGTGCACTCGGCCATTTCCCTTCTCTGTGTCGCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAACCAACTAGAAT
1801 GCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATT
1901 CATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTTGGTATGGAATTTAAAATACAGCATAGCAAA
2001 ACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGA
2101 GCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTCCCAAGTGTGAAGTACTCTTCATTTCTTTATGTTTAAATGCACTGACCTCCACAT
2201 TCCCTTTTATGAAAAATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCAT
2301 AATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTAC
141 • N R T Y
2401 TTGAGGGGATGAGTTTCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGC
135 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 S I L E R C
2501 ACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCT
102 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 K Q G N S
2601 CACAGCAGACCAATGGCAATGGCTTCAAGCAGACAGTACCTGCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTG
69 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 E I G T K T R
2701 ATGGCCGCCCCGACATGGTGTGTTGCTCATAGACATGGTGTGATCTTCTAGTGGCAGCTCCAGCAGCTCAGATCCTGCTGAGAGATGTTGAAGG
35 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 Q S I N F T
2801 TCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTA
2 K M
2901 TCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGG
3001 AAAGTCCCCTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTG
3101 ATGTAAGTCCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCAAGTAGGAAAGTCCCATAAAGTCATGTAAGTGGCATAAT

3201 GCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTC
3301 CACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGACCCAGGC
3401 GGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCT
3501 GCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
3601 TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTC
3701 ATAGCTCACGCTGTAGGTATCTCAGTTCGGGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTAGCCCGACCGCTGCGCCTTATC
3801 CGGTAACATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
3901 GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
4001 TTGGTAGCTCTTGATCCGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA
4101 TCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCC
4201 GCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAA
4301 ACTAGCAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA