



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCATCATGACTCCTCCCTGCCTCAATTGTTCCATCTTTCCTGG
1 M T P P C L N C S I F P G
601 AGCGCTCTCCCAAAATGTTCAAGGAGCCCTGGTCTGCAATATCAGTGAGTTCAAGTGGCCCTATCAACCTGAGGATCTGAACCTTACCGATGAGGCC
13 A L S P N A S R S P L V C N I S E F K W P Y Q P E D L N L T D E A
701 CTGAGGCTGAAGTATTTGGGGCCACAGCAGATGAAACAGTTTGTCCCATCTGTGTCACGTACCTGCTGATCTTCGTGGTGGGCACTCTGGCAACGGGC
47 L R L K Y L G P Q Q M K Q F V P I C V T Y L L I F V V G T L G N G
801 TGACCTGCACCGTCATCTGCGCAACAAGACTATGCGCAGCCCAACTTCTACCTCTTCAAGCCTCGTGTGTCGATATGCTGGTCTCCTGGTGGG
80 L T C T V I L R N K T M R T P T N F Y L F S L A V S D M L V L L V G
901 CTTGCCTTGGAGCTTTATGAGATGCAGCAAAATTACCCGTTCCAGCTGGGTGCGAGTGCCTGCTACTCCGAATACTGCTTTAGAGACCGTCTGCCTA
113 L P L E L Y E M Q Q N Y P F Q L G A S A C Y F R I L L L E T V C L
1001 GCTTCAGTGCTCAATGTCACAGCCCTGAGTGTGGAGCGTTATGTGGCCGTGGTGCGCCACTCCAAGCAAGTCTGTGATGACACGGGCCATGTGCGCC
147 A S V L N V T A L S V E R Y V A V V R P L Q A K S V M T R A H V R
1101 GCATGGTGGGGCCATCTGGGTCTCGTACTCTTCTCTCTGCCCCAACACCAGCCTGCATGGCCTCAGTCAACTAAGTGTGCCCTGCCGGGGCCGGT
180 R M V G A I W V L A T L F S L P N T S L H G L S Q L T V P C R G P V
1201 GCCCGACTCAGCTATATGTTGCTGGTGGTCCCATGGACTTCTACAAGTGGTGGTACTGACTACCGCACTGCTCTTCTTGTCTGCCATGGTCACC
213 P D S A I C S L V G P M D F Y K L V V L T T A L L F F C L P M V T
1301 ATCAGTGTCTGTATCTGCTCATTGGGCTGCGGCTGCGGAGGAGAGATGTTGCTCCAAGTGGAGTCAAGGGCAGGAAAAACCAGCAACCCAGGAGA
247 I S V L Y L L I G L R L R R E R M L L Q V E V K G R K T A A T Q E
1401 CCTCCACAGAAGGATTGAGTGTGCAAGATAGGGGACGGAGACAGGTGACCAAGATGCTGTTTGCAGTGGTGTGGTATTGGCATCTGCTGGGCTCCATT
280 T S H R R I Q L Q D R G R R Q V T K M L F A L V V V F G I C W A P F
1501 CCATGCTGACCGTATCATGTGGAGCCTGGTGTATGGACTCAACGGAAGGCTGCACCTGGCCTACCAGTGTGCCAATTGCCTCTGGCATCTTCTTC
313 H A D R I M W S L V Y G H S T E G L H L A Y Q C V H I A S G I F F
1601 TATCTCGGCTCAGCAGCAACCCGGTCTACAGCCTCATGTCTACTCGTTCGAGAGACCTTCTGCAAGCCCTGGGCTTGGAAACCCAGTGTGTC
347 Y L G S A A N P V L Y S L M S T R F R E T F L Q A L G L G T Q C C C
1701 ATCGCCGCAACCTATAGGCTCCATAACCAATCAGGTTGACCAACAGGACCCCTGTGTGACGTGGGCGCACAGGAACAGCAGGGACCACTCT
380 H R R Q P Y H G S H N H I R L T T G S T L C D V G H R N S R D E P L

NheI (1875)

1801 GGCTGTGAATGAGGATCCAGGGTGTGAGCAAGAGACAGACCCCTCCTGAATAAAATTTGGAAGGAGCCTTACCAGTAGTGGCCAGACATGATAAGAT
413 A V N E D P G C Q Q E T D P S •
1901 ACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATATAAG
2001 CTGCAATAAACAAGTTAAACAACAACATTCATTCTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAA
2101 TGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATC
2201 AGGGGCTGTTGCCAATGTGATTAGCTTTGACGCTCACCTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAGTACTGCTTTCAT
2301 TTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAATGTTTTTT
2401 ATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAA
2501 GAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAG
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
2601 CAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAG
113 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 P H R V A
2701 CCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAAGCAGACAGTACCTGCAATGTAGGCCCTCAATGTG
80 V I T D 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
2801 GACAGCAGATGATCTCCAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTGTTGTCTCATAGAGCATGGTGTCTCTCAGTGGCGACCTCC
47 V A S I 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
2901 ACCAGCTCAGATCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGTGATTATACTATGCCGATATACTATGCCGATGATTA
13 V L E L D Q Q S I N F T K M
3001 TTGTCAAACAGCGTGGATGGCTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATT
3101 GCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAACAACCTCCATTGACGTCATGGGGTGGAGACTTGAAAA

3201 TCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTA
3301 GGAAAGTCCCATAAAGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGAT
3401 GTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGT
3501 CAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGC
3601 AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTG
3701 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC
3801 GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCAGC
3901 AACCCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG
4001 TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGC
4101 GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGC
4201 AGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGT
4301 CATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACAT
4401 ACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA