



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGCTGCGTCCGCGCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCATCATGATGAGACAGAGGCAGAGCCATTATTGTTCCGTGCT
1 M M R Q R Q S H Y C S V L
601 GTTCTGAGTGTCAACTATCTGGGGGGACATTCCAGGAGACATTTGCTCAGAAGAGAATCAAATAGTTTCTCTTATGCTTCTAAAGTCTGTTTTGAG
13 F L S V N Y L G G T F P G D I C S E E N Q I V S S Y A S K V C F E
701 ATCGAAGAAGATTATAAAAATCGTCAGTTTCTGGGCTGAAGGAAATGTGGATGTTGAGTTGATTGATAAGAGCACAAACAGATACAGCGTTTGGTTCC
47 I E E D Y K N R Q F L G P E G N V D V E L I D K S T N R Y S V W F
801 CCACTGCTGGCTGGTATCTGTGGTCAGCCACAGGCCTCGGCTTCTGGTAAGGGATGAGGTCACAGTACGATTGCGTTTGGTTCTGGAGTCAGCACCT
80 P T A G W Y L W S A T G L G F L V R D E V T V T I A F G S W S Q H L
901 GGCCCTGGACCTGCAGCACCATGAACAGTGGCTGGTGGGCGGCCCTTGTGGTACTGCAGAGCCAGAGGAGGCTGTGCGCGAAATCCACCTCCCC
113 A L D L Q H H E Q W L V G G P L F D V T A E P E E A V A E I H L P
1001 CACTTCATCTCCCTCAAAGTGAGGTGGACGTCTCTGGTTTCTCGTTGCCATTTTAAAGATGAAGGGATGGTCTGGAGCATCCAGCCCGGTGGAGC
147 H F I S L Q G E V D V S W F L V A H F K N E G M V L E H P A R V E
1101 CTTTCTATGCTGCTGGAAAGCCCAGCTTCTCTGATGGGCATCTGCTGCGGATCGCCAGTGGGACTCGCTCTCCATCCCATCACTTCCAACAC
180 P F Y A V L E S P S F S L M G I L L R I A S G T R L S I P I T S N T
1201 ATTGATCTATTATCACCCACCCGAAGATATTAAGTTCCACTTGTACCTTGTCCCGAGCAGCCTTGCTAACAAAGGCGATAGATGATGAGGAAGAT
213 L I Y Y H P H P E D I K F H L Y L V P S D A L L T K A I D D E E D
1301 CGTTCCATGGTGTGCGCTGCAGACTTCGCCCCCAATGGAACCCCTGAACCTTGGTTCAGTTATATTGTCTAATTCTGCTAACCTGAAAGTAAATGC
247 R F H G V R L Q T S P P M E P L N F G S S Y I V S N S A N L K V M
1401 CCAAGGAGTTGAAATTGCTACAGGAGCCCTGGAGAAATCAGCACTTCTCAAATCTATGCTGGGAGATGAAGGAACCCATTCAACTTGAGATTAC
280 P K E L K L S Y R S P G E I Q H F S K F Y A G Q M K E P I Q L E I T
1501 TGAAAAAGACATGGGACTTTGGTGTGGGATCTGAGGTGAAGCCAGTGGATCTCCAGCTTGTAGCTGCATAGCCCTCCTCTTTCTCAGTGCAGCC
313 E K R H G T L V W D T E V K P V D L Q L V A A S A P P P F S G A A
1601 TTTGTAAGGAGAACCACCGCAACTCAAGCCAGGATGGGGACCTGAAAGGGTGTCTGATGATCTCCAGGACAATGAGGTTCTTACTGAGAATGAGA
347 F V K E N H R Q L Q A R M G D L K G V L D D L Q D N E V L T E N E
1701 AGGACTGGTGGAGCAGGAAAAGACAGCAGCAAGAAATGAGGCCCTTGTCTGAGCATGGTGGAGAAGAAAGGGGACCTGGCCCTGGACCTGCTCTTACG
380 K E L V E Q E K T R Q S K N E A L L S M V E K K G D L A L D V L F R

NheI (1867)

1801 AAGCATTAGTAAAGGGACCCTTACCTCGTGTCTATCTTAGACAGCAGAATTTGAAAATGAGTCAGCTAGCTGGCCAGACATGATAAGATACATTGAT
413 S I S E R D P Y L V S Y L R Q Q N L •
1901 GAGTTTGGACAAACCACAACCTAGAATGCAGTGAAGAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATA
2001 AACAAAGTTAACAAACAACAAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTAT
2101 GGAATTCTAAATACAGCATAGCAAACTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTG
2201 TTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTACTCTTCATTCTTTAT
2301 GTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTAGAAAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCA
2401 GAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGA
2501 GCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGG
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 C D P
2601 AGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATG
111 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 V I
2701 GTGTCAAAGCTCTTCTGCCGTTGCTCAGCAGACCAATGGCAATGGCTTACAGCAGACAGTACCCCTGCAATGTAGGCTCAATGTGGACAGCAG
77 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 V A S
2801 AGATGATCTCCCACTTGGTCTGATGGCCCGCCAGCATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACAGCTC
44 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 V L E
2901 CAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGTGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAA
11 L D Q Q S I N F T K M
3001 ACAGCGTGGATGGCTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTGCGTCAAT
3101 GGGGCGGAGTTGTTACGACATTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACTCCATTGACGTCAATGGGTGGAGACTTGAAATCCCGGTG

3201 AGTCAAACCGCTATCCACGCCATTGATGTAAGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTACTGCCAAGTAGGAAAGTC
3301 CCATAAGGTCATGTAAGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTAAGTACTGCC
3401 AAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGC
3501 GGGGGTCTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC
3601 AGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC
3701 CGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCT
3801 CCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCAGAACCCCC
3901 GTTCAGCCGACCGCTGCGCTTATCCGTAATATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA
4001 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCT
4101 GAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACG
4201 CGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTA
4301 GTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTC
4401 CATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCGAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA