



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTTCGGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTATCATGAAAGCCTCTAGTCTTGCCTTCAGCCTTCTCTCTGC
1 M K A S S L A F S L L S A
601 TCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACTCAATTTGGGAAGCTGTGTGATGCCACAAAACCTTCAGGAAATACGAAATGGATTT
13 A F Y L L W T P S T G L K T L N L G S C V I A T N L Q E I R N G F
701 TCTGAGATACGGGCGAGTGTGCAAGCCAAAGATGGAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACAAAAGCCTGCGAATCGAT
47 S E I R G S V Q A K D G N I D I R I L R R T E S L Q D T K P A N R
801 GCTGCCTCCTGCGCATTGCTAAGACTCTATCTGGACAGGATTTAAAACTACCAGACCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGC
80 C C L L R H L L R L Y L D R V F K N Y Q T P D H Y T L R K I S S L A
901 CAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTGTGCATGCCACATGACATGCCATTGTGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTG
113 N S F L T I K K D L R L C H A H M T C H C G E E A M K K Y S Q I L

NheI (1092)

1001 AGTCACTTTGAAAAGCTGGAACCTCAGGAGCAGTGTGAAGGCTTTGGGGAACTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGCTAGCTG
147 S H F E K L E P Q A A V V K A L G E L D I L L Q W M E E T E •
1101 GCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTAAAAAATGCTTTATTGTGAAATTTGTGATGCTATTGCTTTA
1201 TTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCA
1301 AGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATG
1401 AATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGT
1501 TTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTAGAAATAATTTAAATACATCATTGCAAT
1601 GAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAA
1701 TAGAAATTTGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGTTGCCATTCA
141 • N R T Y K L P I L E E I T T K V L K G N M
1801 TCTCAATGAGCACAAGCAGTCAAGGAGCATAGTCAAGATGAGTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGA
119 E 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
1901 GTAGGGGTGCCTGACAGCCAAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCA
86 Y 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 V T V R G
2001 ATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCGGACATGGTGTCTTGTCTCATAGAGCATGGTGTATCT
52 I 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
2101 TCTCAGTGGCAGCTCCACCAGTCCAGATCTGCTGAGAGATGTTGAAGTCTTATGGTGGCCCTCTATAGTGTGCTATTATACTATGCCGATAT
19 E T A V E V L E L D Q Q S I N F T K M
2201 ACTATGCCGATGATTAATTTGCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTAC
2301 ACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATG
2401 GGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATGTTAATAGCGATGACTAATACGT
2501 AGATGTAAGTCCCAAGTAGGAAAGTCCCATAGGTCATGTAAGTGGGCATAATGCCAGGCGGGCCATTTACCCTCATTGACGTCAATAGGGGGCGTACTTGG
2601 CATATGATACACTTGTACTGCCAAGTGGGAGTTCACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACA
2701 TACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCTGAGGTTAATAAGAACAT
2801 GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGA
2901 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACCAGGCTTTCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCCTGCCG
3001 TTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCAA
3101 GCTGGGCTGTGTGCAGAACCCCGTTCCAGCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAAACCGTAAGACACGACTTATCGCCA

3201 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA
3301 CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGATTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT
3401 TTTTGTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACTCA
3501 CGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTG
3601 TGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTT
3701 CTCTATCGAA