



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGCTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGAGCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

AgeI (552) BspHI (568)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTAGGAGGGCCATGAACAAGTGGCTGTGCTGCGCACTCCTG
1 M N K W L C C A L L
601 GTGCTCCTGGACATCATTGAATGGACAACCCAGAAACCCCTTCTCCAAAGTACTTGCAATTATGACCCAGAAACTGGTCATCAGCTCCTGTGTGACAAT
11 V L L D I I E W T T Q E T L P P K Y L H Y D P E T G H Q L L C D K
701 GTGCTCCTGGCACCTACCTAAAACAGCACTGCACAGTGAAGGAGAACATTTGTGTCCCTGACCACTCTTATACGGACAGCTGGCACACCAG
44 C A P G T Y L K Q H C T V R R K T L C V P C P D H S Y T D S W H T S
801 TGATGAGTGTGTATTGACGCCAGTGTGCAAGAACTGCAGTCCGTGAAGCAGGAGTGCAACCGCACCCACAACCGAGTGTGTGAGTGTGAGGAAGGG
77 D E C V Y C S P V C K E L Q S V K Q E C N R T H N R V C E C E E G
901 CGTTACCTGGAGATCGAATTCGCTTGAAGCACCGAGCTGCCCGGGCTCCGGCTGGTGAAGTGAACCCCAAGCAAGCAACACAGTTTGC AAA
111 R Y L E I E F C L K H R S C P P G S G V V Q A G T P E Q N T V C K
1001 AATGTCCAGATGGGTTCTTCTCAGTGAGACTTCATCGAAAGCACCTGTAGAAAACACACGAACTGCAGCACATTTGGCCTCCTGTAATTCAGAAAAG
144 K C P D G F F S G E T S S K A P C R K H T N C S T F G L L L I Q K G
1101 AAATGCAACACATGACAACGTGTGTTCCGGAACAGAGAAGCCACGCAAAAAGTGTGGAATAGATGTCACCCTGTGTGAAGAGGCTTCTTCCAGTTTGT
177 N A T H D N V C S G N R E A T Q K C G I D V T L C E E A F F R F A
1201 GTTCTACCAAGATTATACCAAAATGGCTGAGTGTGTTTGGTGGACAGTTTGCCTGGGACCAAAGTGAATGCCGAGAGTGTAGAGAGGATAAAACGGAGAC
211 V P T K I I P N W L S V L V D S L P G T K V N A E S V E R I K R R
1301 ACAGTCAAGAGCAAACTTCCAGCTGTGAGCTGTGAAACATCAAAACAGAGACCAGGAAATGGTGAAGAAGATCATCCAAGACATTGACCTCTG
244 H S S Q E Q T F Q L L K L W K H Q N R D Q E M V K K I I Q D I D L C
1401 TGAAAGCAGCGTGACGCGGCATCTCGGCCACCGCAACCTCACACAGAGCAGCTTCTGCTGCTGATGGAGAGCCTGCCTGGGAAGAAGATCAGCCAGAA
277 E S S V Q R H L G H A N L T T E Q L R A L M E S L P G K K I S P E
1501 GAATTGAGAGAACGAGAAAGACCTGCAAACTGAGCGAGCAGCTCCTGAAGTACTCAGTTTATGGAGGATCAAAAATGGTGACCAAGACACCTTGAAGG
311 E I E R T R K T C K S S E Q L L K L L S L W R I K N G D Q D T L K
1601 GCCTGATGTATGCCCTCAAGCACTTGA AAAACATCCCACTTTCCAAAAGTGTACCCACAGCTGTGAGGAAGACCATGAGGTTCTGCACAGCTTACAAT
344 G L M Y A L K H L K T S H F P K T V T H S L R K T M R F L H S F T M
1701 GTACAGACTGTATCAGAAGCTTTTTAGAAAATGATAGGGAATCAGGTTCAATCTGTGAAAATAAGCTGCTTATAACTAGGAATGGTCACTGGGCTGTT
377 Y R L Y Q K L F L E M I G N Q V Q S V K I S C L •

NheI (1838)

1801 CTTCAGGACGGGCCAAAACCTGACTGAGGAGAAGAAGACCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGCAAAACCACAACCTAGAAATGCA
1901 GTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAT
2001 TTTATGTTTCAGGTTTCCGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACT
2101 TTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCC
2201 TCACCTTCTTTCATGAGTTTAAAGATATAGTGTATTTCCCAAGTGTGAACTAGCTCTTCATTTCTTATGTTTTAAATGCACTGACCTCCACATTTCC
2301 CTTTTTAGTAAAATTCAGAAAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAAT
2401 ATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTG
141 • N R T Y K
2501 AGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACA
134 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 M
2601 TGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTGGGGTGCCTGACAGCCACAATGGTGTCAAAGCTTCTGCCCGTTGCTCAC
101 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
2701 AGCAGACCCAATGGCAATGGCTTCCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATG
68 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 I E G T K T R I
2801 GCCGCCCCGACATGGTCTTGTTCCTCATAGAGCATGGTGTATCTTCCAGTGGCGACCTCCACAGCTCCAGATCCTGTGAGAGATGTTGAAGGTTCT
34 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 K
2901 TCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCT
1 M
3001 GACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTGTACGACATTTTGGAAA
3101 GTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATG

3201 TACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAAGTCATGTACTGGGCATAATGCC
3301 AGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCAC
3401 CCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGG
3501 CCATTTACCGTAAGTTATGTAACCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGC
3601 GTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT
3701 CCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA
3801 GCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGTGCCTTATCCGG
3901 TAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
4001 ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTG
4101 GTAGCTTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC
4201 TTTGATCTTTTACGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCA
4301 ATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACT
4401 AGCAAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA