



1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTAGGAGGGCCACCATGGCGTGGAGCCTTGGGAGCTGGTGGGT
GGTGCCTGCTGGTGTACAGATTGGGAATGGTACCACCTCCCGAAAATGTCAGAATGAATTCTGTTAATTTCAAGAACATTCTACAGTGGGAGTCACCTG
111 G C L L V S A L G M V P P P E N V R M N S V N F K N I L Q W E S P
701 CTTTTGCCAAGGGAACCTGACTTTCACAGCTCAGTACCTAAGTTATAGGATATTCGAATAAATGCATGAATACTACCTTGACGGAATGTGATTTCTC
44 A F A E G N L T F T A Q Y L S Y R I F Q D K C M N T T L T E C D F S
801 AAGCTTTTCAAGTATGGTGACCACACCTTGAGAGTCAAGGCTGAATTTGAGATGAGCATTGAGACTGGGTAACATCACCTTCTGTCTGTGGATGAC
77 S L S K Y G D H T L R V R A E F A D E H S D W V N I T F C P V D D
901 ACCATTATTGGACCCCTGGAATGCAAGTAGAAGTACTTGTGATTCTTTACATATGCGTTTCTTAGCCCTAAAATGAGAATGAATACGAAACTGGGA
111 T I I G P P G M Q V E V L A D S L H M R F L A P K I E N E Y E T W
1001 CTATGAAGAATGTGTATAACTCATGGACTTATAATGTGCAACTACTGAAAAACGGTACTGATGAAAAGTTTCAAATTAATCCCAAGTATGACTTTGAGGT
144 T M K N V Y N S W T Y N V Q Y W K N G T D E K F Q I T P Q Y D F E V
1101 CCTCAGAAACCTGGAGCCATGGACAACCTATTGTGTTCAAGTTCGAGGGTTTCTTCTGATCGGAACAAAGCTGGGGAATGGAGTGAAGCTGTCTGTGAG
177 L R N L E P W T T Y C V Q V R G F L P D R N K A G E W S E P V C E
1201 CAAACAACCCATGACGAAACGGTCCCCTCCTGGATGGTGGCGTCCATCCTCATGGCTCGGTCTTCTGCTGCTGGCCTCCTCGGCTGCTTCGCT
211 Q T T H D E T V P S W M V A V I L M A S V F M V C L A L L G C F A
1301 TGCTGGTGGCTTTACAAGAAGACAAAGTACGCCCTTCTCCCTAGGAATTCCTTCCACAGCACCTGAAAGAGTTTTGGCCATCCTCATCATAACAC
244 L L W C V Y K K T K Y A F S P R N S L P Q H L K E F L G H P H H N T
1401 ACTTCTGTTTTCTCCTTTCCATTGTGCGGATGAGAATGATGTTTTGACAAGTAAGTGTATTGCGAAGACTCTGAGAGCGGCAAGCAGAATCCTGGT
277 L L F F S F P L S D E N D V F D K L S V I A E D S E S G K Q N P G

NheI (1569)

1501 GACAGCTGCAGCCTCGGGACCCCGCTGGGAGGGGCCCAAAGCTAGGCTCTGAGAAGGAAACACACTGCTAGCTGGCCAGACATGATAAGATACATTG
311 D S C S L G T P P G Q G P Q S •
1601 ATGAGTTTGGACAAACCAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAA
1701 TAAACAAGTTAACAACAACAATTGCATTATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGT
1801 ATGGAATTCTAAAATACAGATAGCAAACTTAACTCCAATCAAGCTCTACTGAACTCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGCC
1901 TGTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTCATGGAGTTTAAAGATATAGTGTATTTCCCAAGGTTTGAAGTAGCTCTTCAATTTCTT
2001 ATGTTTTAAATGCACTGACCTCCACATTCCCTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGG
2101 CAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGC
2201 GAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCA
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
2301 GGAGCATAGTCAGAGATGAGCTCTGTCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAA
111 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 V I
2401 TGGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCAATGTAGGCCTCAATGTGGACAGC
78 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
2501 AGAGATGATCTCCCAAGTCTGGTCTGATGGCCGCCGACATGGTCTTGTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGC
45 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 V L
2601 TCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAAGTATTATACTATGCGGATATACTATGCGGATGATTAATTGTCA
11 E L D Q Q S I N F T K M
2701 AAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACAGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCA
2801 ATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGG
2901 TGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAG
3001 TCCATAAAGTCTATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTG
3101 CCAAGTGGGAGTTTACCCTAAATACTCCACCATTGACGTCAATGGAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGG

3201 GCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAAITTAAGAACATGTGAGCAAAGGCCAGCAAAGG
3301 CCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
3401 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCGACCCCTGCCGTTACCGGATACCTGTCCGCCTTT
3501 CTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCC
3601 CCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
3701 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCG
3801 CTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTA
3901 CGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACACTCACGTTAAGGGATTTTGGTCATGGC
4001 TAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTC
4101 TCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA