



1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGGTCATCATGATGGATCAAGCCAGATCAGATTCTCTAACTTGT
1 M M D Q A R S A F S N L F
601 TGGTGGGAACCATTTGTCATACACCCGTTTAGCCTTGCTCGCAAGTAGATGGAGATAACAGTCATGTGGAGATGAAACTGGCTGCAGATGAAGAAGAA
13 G G E P L S Y T R F S L A R Q V D G D N S H V E M K L A A D E E E
701 AATGCCGACAATAACATGAAGGCTAGTGTGAGAAAACCAAGAGTTAATGGGAGACTCTGCTTTCAGCTATTGCACTAGTCATTTTCTTCTTGATTG
47 N A D N N M K A S V R K P K R F N G R L C F A A I A L V I F F L I
801 GATTCATGAGTGGCTACCTGGGCTATTGTAAGCGTGTAGAACAAAAGAGGAGTGTGTGAAACTGGCTGAAACGGAGGAGACAGACAAGTCAGAAACCAT
80 G F M S G Y L G Y C K R V E Q K E E C V K L A E T E E T D K S E T M
901 GGAACAGAGGATGTTCTACATCATCTCGTTATATTGGGCAGACCTCAAACTGTGTCAGAGAAGTTGAACTCCATAGAGTTTGCCGACACCATC
113 E T E D V P T S S R L Y W A D L K T L L S E K L N S I E F A D T I
1001 AAGCAGCTGAGCCAGAATACATACACTCCTCGTGAGGCTGGATCTCAAAAAGATGAAAGTCTTGCTTATTATATTGAAAATCAGTTCATGAATTTAAAT
147 K Q L S Q N T Y T P R E A G S Q K D E S L A Y Y I E N Q F H E F K
1101 TCAGCAAAGTCTGGCGAGATGAACACTATGTGAAGATTCAAGTAAAAAGCAGCATTGGTCAAAACATGGTGACCATAGTGCAGTCAAATGGTAACTTAGA
180 F S K V W R D E H Y V K I Q V K S S I G Q N M V T I V Q S N G N L D
1201 CCCAGTGGAGTCTCCCGAGGTTATGTGGCATTGAGTAAACCTACAGAAGTTTCTGGTAAACTGGTCCATGCTAATTTTGGCACTAAAAAGGACTTTGAA
213 P V E S P E G Y V A F S K P T E V S G K L V H A N F G T K K D F E
1301 GAACTAAGTTATTCTGTGAATGGATCTTTAGTATTGTAGAGCAGGGGAAATTAATTTGCAGAAAAGTTGCAAATGCCCAAAGCTTTAATGCAATTG
247 E L S Y S V N G S L V I V R A G E I T F A E K V A N A Q S F N A I
1401 GTGTCCTCATATACATGGACAAGAATAAATCCCGTGTGGAGGAGACCTGCACTCTTTGGACATGCTCATCTAGGAAGTGGTATCCATACACACC
280 G V L I Y M D K N K F P V V E A D L A L F G H A H L G T G D P Y T P
1501 TGGCTTCTCTTTCAATCATACTCAGTTTCCGCCATCTCAGTCATCAGGGTGCCTAATATACCTGTGCAAAACATCTCAAGAGCTGCTGCAGAAAAG
313 G F P S F N H T Q F P P S Q S S G L P N I P V Q T I S R A A A E K
1601 CTATTTGGAATAAAGGAAGCTGTCTGTAGATGGAACATAGATTCTTCATGTAAGCTGGAACCTTTCACAGAATCAAATGTGAAGCTCATTGTGA
347 L F G K M E G S C P A R W N I D S S C K L E L S Q N Q N V K L I V
1701 AAAACGTACTGAAAGAAAGAATAACTTAACATCTTTGGAGTTAATAAGGTTATGAGGAACAGACCGTTATGTTGTAGTAGGACCCAGAGACCGC
380 K N V L K E R R I L N I F G V I K G Y E E P D R Y V V V G A Q R D A
1801 TTTGGGTGCTGGTGTGCGGCGAAGTCCAGTGTGGGAACAGGCTTTCTGTTGAAACTGCCAAGTATTCTCAGATATGATTTCAAAGATGGATTTAGA
413 L G A G V A A K S S V G T G L L L K L A Q V F S D M I S K D G F R
1901 CCCAGCAGAAGTATAATCTTGGCAGCTGGACTGCAGGCGACTTTGGAGCTGTTGGTGCCACTGAGTGGTGGAGGGATACCTTTCATCTTTGCATTTAA
447 P S R S I I F A S W T A G D F G A V G A T E W L E G Y L S S L H L
2001 AAGCTTCACTTATTAATTTGGATAAAGTTGTCCTTGGTACTAGTAACTCAAAGTTTCTGCCAGCCCCTATTATATACACTTATGGGAAAGATAA
480 K A F T Y I N L D K V V L G T S N F K V S A S P L L Y T L M G K I M
2101 GCAAAGTAAAGCATCCAGTTGATGGAAAATCTATATAGAGACGCAATTGGATTAGCAAAGTTGAGAAAATTTCTTTGACAATGCTGCATATCCT
513 Q D V K H P V D G K S L Y R D S N W I S K V E K L S F D N A A Y P
2201 TTCCTGCAATTTCTGGAATCCAGCAGTTTCTTTTGTGTTTGTGAGGATGCAGACTATCCTTATTTGGGCACTAGATTGGATACCTATGAGGCATTGA
547 F L A Y S G I P A V S F C F C E D A D Y P Y L G T R L D T Y E A L
2301 CTCAGAAAAGTTCTCAGCTCAACCAAATGGTTCGTACAGCAGCGGAAGTGGCTGGTCACTCATTATTAAGTACCCATGACGTTGAATTGAACCTGGA
580 T Q K V P Q L N Q M V R T A A E V A G Q L I I K L T H D V E L N L D
2401 CTATGAGATGTATAACAGCAAACACTGTCATTTATGAAGGATCTGAACAGTTCAAAACAGATATCAGGGATATGGTCTAAGTCTACAGTGGCTGTAT
613 Y E M Y N S K L L S F M K D L N Q F K T D I R D M G L S L Q W L Y
2501 TCCGCTCGTGAGACTACTTCCGTGCTACTTCTAGACTAACAAGTATTTTCAATGCTGAGAAAACAAACAGATTTGTCATGAGGGAAATCAATGATC
647 S A R G D Y F R A T S R L T T D F H N A E K T N R F V M R E I N D
2601 GTATTATGAAAGTGGAGTATCACTTCTGTGCGCCTATGTATCTCCAAGAGAGTCTCCTTCCGACATATCTTCTGGGCTCTGGCTCTCACACTCTCTC
680 R I M K V E Y H F L S P Y V S P R E S P F R H I F W G S G S H T L S
2701 AGCTTTAGTGGAGAAGTGAAGCTTCTGCAAAAAATATTACTGCTTTAATGAAACCTCTCAGAAAACAGTTGGCCCTGGCTACTTGGACTATTGAG
713 A L V E N L K L R Q K N I T A F N E T L F R N Q L A L A T W T I Q

NheI (2865)

2801 GGAGTCGCAAATGCCCTCTCTGGTGACATTTGGAATATTGACAATGAGTTTTAAATTAACATGCAGTAGCTGGCCAGACATGATAAGATACATTGATGA
747 G V A N A L S G D I W N I D N E F •
2901 GTTTGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAA
3001 CAAGTTAAACAACA AATTCATTCTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGG
3101 AATTCTAAAATACAGCATAGCAAAC TTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTT

3201 GCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTCTTTCATTCTTTATGT
3301 TTAAATGCACCTGACCTCCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGA
3401 ATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGC
3501 TTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCACTCAATGAGCACAAAGCAGTCAGGAG
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 A
3601 CATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGT
110 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 T
3701 GTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAG
77 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
3801 ATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTTCCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCA
43 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 L
3901 GATCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAAC
10 D Q Q S I N F T K M
4001 AGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTCGCTCAATGG
4101 GCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCTGAG
4201 TCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCC
4301 ATAAGGTCATGTACTGGGCATAATGCCAGGCGGCCATTTACCCTGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAA
4401 GTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGG
4501 GGGTCGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAG
4601 GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCGG
4701 ACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTTCTCC
4801 CTTCCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGT
4901 TCAGCCCAGCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATT
5001 AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGA
5101 AGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCA
5201 CAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGT
5301 TAATTAACATTTAAATCAGCGGCCAATAAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCA
5401 TCAAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA