



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTTCGCCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGGTCCACATGGAGCCCGCCGGCCCGCCCTGGCCGCTAGGGCC
1 M E P A G P A P G R L G P
601 GCTGCTGCTCTGCCTGCTGCTCTCCGCTCCTGTTTCTGTACAGGAGCCACGGGGACAGAAAGTGAAGGTGACTCAGCCTGAGAAATCAGTGTCTGTTGCT
13 L L L C L L L S A S C F C T G A T G T E V K V T Q P E K S V S V A
701 GCTGGGATTGACCATCTCAACTGCACTGTGACCTCCTGTTGCCGGTGGGACCCATTAGGTGGTACAGAGGAGTAGGGCAAAGCCGGCTGTTGATCT
47 A G D S T I L N C T V T S L L P V G P I R W Y R G V G Q S R L L I
801 ACAGTTTTCACAGGAGAACAACCTTCTCGAGTTAGAAATGTTTTCAGATACTACTAAGAGAAACAATATGGACTTTTCCATCCGTATCAGTAATGTCACCCC
80 Y S F T G E H F P R V R N V S D T T K R N N M D F S I R I S N V T P
901 AGAAGATGCCGGTACCTACTACTGTGTGAAGTCCAGAGAGGATCATCAGAGCTGACACAGAAATACAATCTGGAGGGGGAACAGAGGTCTATGACTC
113 E D A G T Y Y C V K F Q R G S S E P D T E I Q S G G G T E V Y V L
1001 GCCAAACCTTCTCCACGGAGGTATCCGGCCAGCAGACAGGGGCATACCTGACCAGAAAGTGAACCTCACCTGCAAGTCTCATGGCTTCTCTCCCGGA
147 A K P S P P E V S G P A D R G I P D Q K V N F T C K S H G F S P R
1101 ATATCACCTGAAGTGGTCAAAGATGGGCAAGAACTCCACCCCTTGAGACCACCGTGAACCTAGTGGAAAGAATGTCTCTACAACATCTCCAGCAC
180 N I T L K W F K D G Q E L H P L E T T V N P S G K N V S Y N I S S T
1201 AGTCAGGGTGGTACTAACTCCATGGATGTTAATTCTAAGGTGATCTGCGAGGTAGCCACATCACCTGGATAGAAGCCCTCTTCTGTTGGATTGCTAAC
213 V R V V L N S M D V N S K V I C E V A H I T L D R S P L R G I A N
1301 CTGTCTAACTTCAGGTTTCCACCCGTAAGGTCAACACAGCCCGGACATCAATGAACAGGTGAACCTCACCTGCCGGCTGAGAGGTTCT
247 L S N F I R V S P T V K V T Q Q P P T S M N Q V N L T C R A E R F
1401 ACCCGAGGATCTCCAGCTGATCTGGCTGGAGAATGAAACGTATCACGGAATGACACGCCAAGAATCTCACAAAGAACACGGATGGGACCTATAATTA
280 Y P E D L Q L I W L E N G N V S R N D T P K N L T K N T D G T Y N Y
1501 CACAAGCTTGTCTGTTGAACCTCATCTGCTCATAGAGAGGACGTGGTGTTCACGTCCAGGTGAAGCAGCAACAGCCAGCGATACCCGAAACCAT
313 T S L F L V N S S A H R E D V V F T C Q V K H D Q Q P A I T R N H
1601 ACCGTGCTGGGACTTGCCCACTCGAGTATCAAGGGAGCATGCAAACCTTCCCTGGTAATAATGCTACCCACAACCTGGAATGTCTTCATCGGTGTGGGGC
347 T V L G L A H S S D Q G S M Q T F P G N N A T H N W N V F I G V G
1701 TGGCGTGTCTTGTCTAGTCTGCTGATGGCTGCTCTCTCCTCCGGATCAAACAGAAAGCAAGGGGTCAACATCTTCCACACGGTTGCA
380 V A C A L L V L V L L M A A L Y L L R I K Q K K A K G S T S S T R L H
1801 CGAGCCCGAGAAGATGCCAGGGAAATAACCCAGATCCAGGACACAAATGACATCAACGACATCACATACGACAGCTGAATCTGCCAAAGAGAAGAAG
413 E P E K N A R E I T Q I Q D T N D I N D I T Y A D L N L P K E K K
1901 CCCGACCCCGGCCCTGAGCCTAACCAACACAGAATATGCAAGCATTGAGACAGGCAAAGTGCCTAGGCCAGAGGATACCTCACCTATGCTGACC
447 P A P R A P E P N N H T E Y A S I E T G K V P R P E D T L T Y A D
2001 TGGACATGGTCCACCTCAACCGGCACAGCCAGCCCCAAGCCTGAGCCATCTTCTCAGAGTATGCTAGTGTCCAGGTCCAGAGGAAGTGAATGGGGCT
480 L D M V H L N R A Q P A P K P E P S F S E Y A S V Q V Q R K •

NheI (211)

2101 GTGGTCTGTACGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTAAAT
2201 TTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCATTTTATGTTTCAGGTTAGGGGGAGGTG
2301 TGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTTGGTATGGAATCTAAAATACAGCATAGCAAACTTAACTCCAATCAAGCCTCTACTTG
2401 AATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATA
2501 TAGTGTATTTTCCAAGGTTTGAAGTACTCTTCTATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTAGAAAATAAT
2601 TTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAG
2701 GGAACAAAGGAACCTTTAATAGAAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTT
141 • N R T Y K L P I L E E I T T
2801 TTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCCAGGAGCATAGTCCAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGG
125 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 G C P S V V R I S
2901 ATCTGTCCACCTCATCAGAGTAGGGGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGC
92 R D V E D S 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
3001 ACAGACAGTACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTGGTCTGATGGCCGCCCCGACATGGTGTCTGTTGTC
59 C V T V R G I Y A E I H V A S I I E G T K T R I A A G V H H K N D
3101 TCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACAGCTCCAGATCTGCTGAGAGATGTTGAAGGTCTTATGGTGGCCCTCTATAGTGAGTGC
25 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 M

3201 TATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTA
3301 TATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTGATTTACTAGTCAAAACAA
3401 ACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTA
3501 ATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGT
3601 CAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTAT
3701 TGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTT
3801 GCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTCCATAGGCTCCGCCCCCTGA
3901 CGAGCATCAGAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAAGTCCCTCGTGCGCTCT
4001 CCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG
4101 TGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGT
4201 AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC
4301 TACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCA
4401 CCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGC
4501 TCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTCATTACATC
4601 TGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGAAG
4701 TGCAGGTGCCAGAACATTTCTCTATCGAA