



1 GGACCTGCAGGGCTGAATAACCTCTGAAAGGGAAGCTTGGTTAGGTACCTCTGAGCGGAAAGAACAGCTGTGGAATGTGTGCAGTTAGGGTGTG
101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCAGGCTCCCAGCAGGCAG
201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTCTCCGGTGGCGTCTAGTGGCAGAGCGCACATCGCCACAGTCCCCGA
301 GAAGTTGGGGGAGGGGTGCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACCTGGGAAAGTGTGCTGTACTGGCTCCGCTTTTTCCC
401 GAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTC
501 GCATCTCTCTTACGCGCCCGCCCTACCTGAGGCGCCATCCACGCGGTTGAGTCTGCGGTTCTGCCGCTCCGCGCTGTGGTGCCTCCTGAACTGC
601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGC

Agel (791)

701 CTGACCCTGCTTGCTCAACTCTACGTCTTTGTTTCTGTTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGTCCAC
801 CATGTTTGTGTTTCTGTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAAGCTCAATTACCCCTGCATACACTAATTCTTTCACA
1 M F V F L V L L P L V S S Q C V N L T T R T Q L P P A Y T N S F T
901 CGTGGTGTGTTTACCCTGACAAAGTTTTCAGATCCTCAGTCTTACATCAACTCAGGACTTGTCTTACCTTTCTTTTCAATGTTACTTGGTCCATG
34 R G V Y Y P D K V F R S S V L H S T Q D L F L P F F S N V T W F H
1001 CTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGCTCCTACCATTAAATGATGGTGTGTTATTTTCTTCCACTGAGAAGTCTAACAT
67 A I H V S G T N G T K R F D N P V L P F N D G V Y F A S T E K S N I
1101 AATAAGAGGCTGGATTTTGGTACTACTTTAGATTGCAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAGTCTGTGAATTT
100 I R G W I F G T T L D S K T Q S L L I V N N A T N V V I K V C E F
1201 CAATTTTGAATGATCCATTTTGGGTGTTTATTACCACAAAAACAACAAAGTTGGATGGAAAGTGAAGTTCAGAGTTTATTCTAGTGCGAATAATTGCA
134 Q F C N D P F L G V Y Y H K N N K S W M E S E F R V Y S S A N N C
1301 CTTTTGAATATGCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAAACAGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTTA
167 T F E Y V S Q P F L M D L E G K Q G N F K N L R E F V F K N I D G Y
1401 TTTTAAATATATTCTAAGCACAGCCTATTAATTTAGTGCCTGATCTCCCTCAGGTTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATT
200 F K I Y S K H T P I N L V R D L P Q G F S A L E P L V D L P I G I
1501 AACATCACTAGGTTTCAAACCTTACTTGTCTTACATAGAAGTTATTTGACTCCTGGTATTCTTCTCAGGTTGGACAGCTGGTGTGACGCTTATTATG
234 N I T R F Q T L L A L H R S Y L T P G D S S S G W T A G A A A Y Y
1601 TGGGTTATCTTCAACCTAGGACTTTTCTATTAATAATAATGAAAAAGAACATTACAGATGCTGTAGACTGTGCACTTGCACCTCTCTCAGAAACAAA
267 V G Y L Q P R T F L L K Y N E N G T I T D A V D C A L D P L S E T K
1701 GTGTACGTTGAAATCCTTCACTGTAGAAAAAGGAATCTATCAAACTTCAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCTTAATATTACA
300 C T L K S F T V E K G I Y Q T S N F R V Q P T E S I V R F P N I T
1801 AACCTGTGCCCTTTTGGTGAAGTTTTAACGCCACCAGATTTGCATCTGTTTATGCTTGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTG
334 N L C P F G E V F N A T R F A S V Y A W N R K R I S N C V A D Y S
1901 TCCTATATAATCCGCATCATTTTCCACTTTTAAAGTGTATGGAGTGTCTCCTACTAAATTAATGATCTCTGCTTTACTAATGTCTATGCAGATTCAAT
367 V L Y N S A S F S G T F K C Y G V S P T K L N D L C F T N V Y A D S F
2001 TGTAATTAGAGGTGAAGTCAGACAAATCGCTCCAGGGCAAAGTGGAAAGTGTGCTGATTATAATTATAAATTACCAGATGATTTTACAGGTCGCGTT
400 V I R G D E V R Q I A P G Q T G K I A D Y N Y K L P D D T A G T G C V
2101 ATAGCTTGAATCTAACCAATCTGATTCTAAGGTTGGTGAATTAATTAATCACTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGGAGAGAG
434 I A W N S N N L D S K V G G N Y N Y L Y R L F R K S N L K P F E R
2201 ATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGAATGGTGTGAAAGTTTTAATGTTACTTTCTTTACAATCATATGGTTTCCAACCCAC
467 D I S T E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T
2301 TAATGGTGTGGTTACCAACCATACAGAGTAGTAGTCTTTTCTTTGAACTTCTACATGCACCAGCAACTGTTTGGGACCTAAAAAGTCTACTAATTTG
500 N G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K K S T N L
2401 GTTAAAAACAAATGTGCAATTTCAACTTCAATGGTTTAAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCA
534 V K N K C V N F N F N G L T G T G V L T E S N K K F L P F Q Q F G
2501 GAGACATTGCTGACACTACTGATGCTGTCCGTGATCCACAGACACTTGAAGTCTTGAACATTACACCATGTTCTTTTGGTGGTGTGAGTGTATAACACC
567 R D I A D T T D A V R D P Q T L E I L D I T P C S F G G V S V I T P
2601 AGGAACAAATACTTCAACCAGGTTGCTGTTCTTTATCAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTGATGCAGATCAACTACTCTACTTGG
600 G T N T S N Q V A V L Y Q D V N C T E V P V A I H A D Q L T P T W
2701 CGTGTGTTTCTACAGGTTCTAATGTTTCTTCAACACGTGCAGGCTGTTAATAGGGGCTGAACATGTCAACAACCTCATATGAGTGTGACATACCCATTG
634 R V Y S T G S N V F Q T R A G C L I G A E H V N N S Y E C D I P I
2801 GTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTTCTCCTCGGGGACAGTAGTGTAGCTAGTCAATCCATCATTGCTTACTACTATGCTACTTGG
667 G A G I C A S Y Q T Q T N S P R R A R S V A S Q S I I A Y T M S L G
2901 TGCAGAAAATTGAGTTGCTTACTCTAATAACTCTATTGCCATACCCACAAATTTTACTATTAGTGTACCACAGAAATTTACCAGTGTCTATGACCAAG
700 A E N S V A Y S N N S I A I P T N F T I S V T T E I L P V S M T K
3001 ACATCAGTAGATTGTACAATGTACATTTGTTGGTATTCAACTGAATGCAGCAATCTTTTGTGCAATATGGCAGTTTTTGTACACAATTAACCGTGCTT
734 T S V D C T M Y I C G D S T E C S N L L L Q Y G S F C T Q L N R A
3101 TAACTGGAATAGCTGTTGAACAAGACAAAAACCCCAAGAAGTTTTGCAAGTCAAACAAATTTACAAAACACCACCAATTAAGATTTTGGTGGTTT
767 L T G I A V E Q D K N T Q E V F A Q V K Q I Y K T P P I K D F G G F
3201 TAATTTTTCACAAATATTACCAGATCCATCAAAACCAAGCAAGAGGTCATTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTC
800 N F S Q I L P D P S K P S K R S F I E D L L F N K V T L A D A G F

3301 ATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCTCATTGTGCACAAAAGTTTAAACGGCCTTACTGTTTTGCCACCTTTGCTCACAG
834▶ I K Q Y G D C L G D I A A R D L I C A Q K F N G L T V L P P L L T
3401 ATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCTGGTTGGACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTAT
867▶ D E M I A Q Y T S A L L A G T I T S G W T F G A G A A L Q I P F A M
3501 GCAAAATGGCTTATAGGTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAGAACCAAAATGATTGCCAACCAATTTAATAGTGCTATTGGCAAA
900▶ Q M A Y R F N G I G V T Q N V L Y E N Q K L I A N Q F N S A I G K
3601 ATCAAGACTCACTTTCTCCACAGCAAGTGCCTTGGAAAACCTCAAGATGTGGTCAACCAAAATGCACAAGCTTTAAACACGCTTGTAAACAACCTTA
934▶ I Q D S L S S T A S A L G K L Q D V V N Q N A Q A L N T L V K Q L
3701 GCTCCAATTTTGGTGCAATTTCAAGTGTAAAATGATATCCTTTACGTCTTGACAAAGTTGAGGCTGAAGTCAAATTTGATAGGTTGATCAGGCGAG
967▶ S S N F G A I S S V L N D I L S R L D K V E A E V Q I D R L I T G R
3801 ACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCTTCTGCTAATCTTCTGCTACTAAAATGTCAGAGTGTGTA
1000▶ L Q S L Q T Y V T Q Q L I R A A E I R A S A N L A A T K M S E C V
3901 CTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTATGTCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTT
1034▶ L G Q Q S K R V D F C G K G Y H L M S F P Q S A P H G V V F L H V T
4001 ATGTCCTGCACAAGAAAAGAACTTCAACAAGTCTCCTGCCATTTGTCATGATGAAAAGCACACTTCTCCTGTAAGGTGTCTTTGTTTCAAATGGCAC
1067▶ Y V P A Q E K N F T T A P A I C H D G K A H F P R E G V F V S N G T
4101 AACTGGTTTGTAAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACACATTTGTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTC
1100▶ H W F V T Q R N F Y E P Q I I T T D N T F V S G N C D V V I G I V
4201 AACACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATTTTTAAGAATCATACATCACCAGATGTTGATTTAG
1134▶ N N T V Y D P L Q P E L D S F K E E L D K Y F K N H T S P D V D L
4301 GTGACTCTCTGGCATTAACTTCAGTTGTAACACTTCAAAAAGAAATGACCGCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTCATCGATCT
1167▶ G D I S G I N A S V V N I Q K E I D R L N E V A K N L N E S L I D L
4401 CCAAGAAGTTGAAAAGTATGAGCAGTATATAAATGGCCATGGTACATTTGGCTAGGTTTTATAGCTGGCTTGATTGCCATAGTAATGGTGACAATATG
1200▶ Q E L G K Y E Q Y I K W P W Y I W L G F I A G L I A I V M V T I M

NheI (4567)

4501 CTTTGTGTATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGAT
1234▶ L C C M T S C C S C L K G C C S C G S C C •
4601 GAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATA
4701 AACAAAGTTAACACAACAATTCATTATTTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTAT
4801 GGAATTTAAAATACAGCATAGCAAACTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTG
4901 TTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAACACTGCTTTCATTTCTTTAT
5001 GTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCA
5101 GAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTTGGACAGCAAGAAAGCGA
5201 GCTTCTAGCTTTAGTTCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGG
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
5301 AGCATAGTCAGAGATGAGCTCTGTCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATG
111▶ 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
5401 GTGTCAAAGTCTTCCCGTTGCTCAGCAGACCCAAATGGCAATGGCTTCAGCAGACAGTAGGCTGACCCTGCAATGAGGCTCAATGTGGACAGCAG
77▶ 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 S
5501 AGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGGACCTCCACCAGCTC
44▶ 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
5601 CAGATCCTGCTGAGAGATGTTGAAGTCTTATGATGGCCCTCTATAGTGTGCTATTATACTATGCGGATATACTATGCCGATGATTAATTGTCAA
11▶ L D Q Q S I N F T K M
5701 ACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTGCGTCAAT
5801 GGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGTGGAGACTTGGAAATCCCCGTG
5901 AGTCAAACCGCTATCCACGCCATTGATGTAAGTCCGATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTC
6001 CCATAAGTTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCC
6101 AAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTATTATTGACGTCAATGGG
6201 GGGGTCGTTGGGCGTCCAGCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCTGACGTTAATTAAGAACATGTGAGCAAAAGCCAGCAAAAGGCC
6301 AGGAACCGTAAAAAGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCAGCGCTCAAGTCAAGGTTGGCGAAACC
6401 CGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCT
6501 CCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGACGAACCCCC
6601 GTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA

6701 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCT
6801 GAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTGCAAGCAGCAGATTACG
6901 CGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTGGTCATGGCTA
7001 GTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACTAACATACGCTCTC
7101 CATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA