



200

1 GAGCTTGGCCCATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACGCCATGTTGACATTGATTATTGACTA
101 GTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGGTTACATAACTACGGTAAATGGCCCGCTGGCTGACCGCC
201 CAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAA
301 ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACA
401 TGACCTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATA
501 GCGGTTTGACTCACGGGGATTTC AAGTCTCCACCCATTGACGTCAATGGGAGTTTGT TTTGGCACCAAATCAACGGGACTTTCCAAAATGTCGTAAC
601 AACTCCGCCCCATTGACGCAAATGGGCGGTAGCGGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCC
701 ATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCTCCGGTGCACCGATCCTGAGAACTCAGGgtgagtttggggacccttgattg
801 ttctttctttttcgtattgtaaaattcatgttatatggagggggcaaagtttcagggtgttgttagaatgggaagatgtcccttgatcacctatgga
901 ccctcatgataatttgtttctttcactttctactctgttgacaaccattgtctcctcttattttcttttcattttctgtaactttttcgttaaacctta
1001 gcttgcatttgtaacgaattttaattcacttttgtttattttgtcagattgtaagtactttctctaatcactttttttcaaggcaatcagggtatatt
1101 atattgtacttcagcacagtttagagaacaattgttataattaaatgataaggtagaatattttctgcatataaattctggctggcgtggaatatctt
1201 attggtagaacaactacaccctggtcatcatcctgcctttctctttatggttacaatgatatacactgtttgagatgaggataaaaatactctgagtcca
1301 aaccgggcccctctgtaaccatgttcatgcctttctctttctacagCTCCTGGGCAACGTGCTGGTTGTGTGCTGTCTCATCATTTTGGCAAAGA

Agel (1416)

1401 ATTCTCGACGGATC¹ACCGGTCAACATGTTTGTGTTCTTGGTGTGCTTCCACTGGTCAGTCCCAATGCGTTAATCTCACACCCGAACTCAACTCCC
1501 ACCCGCATATACAAATTCCTTACCAGAGGAGTACTATCCTGACAAAGTGTTCGGTCAAGTGTCTCCACTCTACTCAGGACCTCTTTCTGCCTTTCC
25▶ P A Y T N S F T 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
1601 TTTTCTAACGTTACATGGTTTCATGCAATCCATGTGTCTGGGACAAACGGCACAAACGCTTCGACAAACCCTGATTGCCATTCAATGATGGGGTGTACT
59▶ F S N V T W F H 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
1701 TTGCTCCACAGAGAAATCCAACATCATTGAGGATGGATTTTCGGGACTACTCTGGACTCAAAGACACAGAGCCTGCTGATCGTTAAACAACGCCACAAA
92▶ F A S T E K S N I 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
1801 CGTTGTATCAAAGTGTGCGAATTCAGTTCGCAATGATCCCTTCTGGGAGTGTACTATACAAGAATAACAAGTCTGGATGGAGAGCGAATTTCCGG
125▶ V V I K V C E F 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
1901 GTCTACAGCAGCGAAACAATGACCTTCGAGTACGTGAGTCAACCTTTCTGATGGACCTGGAAGGAAACAGGAAACTCAAGAACCTGAGAGAGT
159▶ V Y S S A N N C 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
2001 TTGCTTTAAGAACATCGACGGCTATTTAAGATCTATAGTAAGCATAACGCTATCAACCTGGTAAGGGATCTTCCCCAGGGCTTTTCAGCCCTGGAACC
192▶ F V F K N I D G Y F K I Y S K H T P I N L V R D L P Q G G F S A L E P
2101 TTTGGTTGACTTGCTATTGGTATCAATACCAGATTTCCAGACCTTCTGGCATTGCAIGGGTCTTACTTACTCCAGGTATTCTCTCCCTCCGGGTGG
225▶ L V D L P I G I 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
2201 ACTGCCGGCGCCGCTGCCTACTATGTGCGGTATCTGCAACCAAGAAGTTCCTGCTCAAGTACAACGAAAACGGCACTATTACGGATGCTGTTGATTGTG
259▶ T A G A A A Y Y 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
2301 CCCTGGACCTCTGTCTGAGACTAAATGACCCCTCAAGAGCTTACCCTTGGAGGGGATTTACCAAACAGTAATTTCCGGGTCCAACCCACCGAAAG
292▶ A L D P L S E T K 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
2401 CATTGTGCGGTTCCCAAATACCAATCTGTGTCCCTTTGGCGAAGTGTCAATGTACAAGTTTGTCTGTGTACGCATGGAATAGGAAACGCATC
325▶ I V R F P N I T 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
2501 TCCAATTGTGCTGATTACTCCGTGCTGTACAATTCGGCTCTTTCTCAACCTTCAAGTGTATGGCGTTACCTACCAAATTAACGACCTGTGCT
359▶ S N C V A D Y S V L Y N S A S F S T F K C Y G V S P T K L N D L C
2601 TCACTAATGTGTATGCCGACTCTTTGTGATACGAGGCATGAAGTGAACAGATTGCACCAGGGCAGACCGGCAAAATTTGCCGACTACAACATAAGCT
392▶ F T N V Y A D S F 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
2701 TCCAGATGACTTTACCGGATGTGTTATTGCATGGAACCAACATCTGGATTCCAAGTGGGTGGCAACTATAACTACCTGTATAGACTGTTCCAGGAAA
425▶ P D D F T G C V 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
2801 TCCAACCTGAAACCATTCGAGCGAGATATAAGCACAGAAATCTACCAGGTGGAAGTACGCCCTGCAACGGCGTGGAAAGGTTCAACTGCTACTTCCCAT
459▶ S N L K P F E R D I S T E I Y Q A G S T P C N G V E G F N C Y G
2901 TGCAGATTACGGATTCAGCTACAAACGGGTGGGTACCAACCTATCGTGTCTAGTCTGAGTTTTGAGCTCCTCCATGCCCCAGCCACAGTCTG
492▶ L Q S Y G F Q P T 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
3001 TGGCCCCAAGAAAAGCACCAATCTGGTGAAGAACAATGCGTGAACCTTAACTTTAACGGACTCACAGGAACCGGCGTATTGACGGAGAGTAACAAGAAG
525▶ G P K K S T N L 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
3101 TTCTGCCATTCCAGCAGTTCGTCGCGATATTGCCGACACTACCGACGCTGTCCGAGATCCCAGACATTGGAGATTCTTGATATCACACCTGTAGTT
559▶ F L P F Q Q F G 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

D614G (3266)

3201 TCGGCGGAGTGAGCGTGATTACGCCCGGAACCAATACCAGCAATCAGGTTGCCGTCCTGTATCAGGCGTGAATTGCACCGAGGTACCTGTGCGCCATCCA
592 F G G V S V I T P G T N T S N Q V A V L Y Q G V N C T E V P V A I H
3301 CGTGACCAACTTACACCCACATGGCGAGTATATCCACCGGCTCCAACGCTTTTCAGACACGTGCTGGATGTCTGATCGGTGCGAACACGTTAATAAT
625 A D Q L T P T W 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
3401 AGCTACGAGTGTGATATCCCATCGGTGCTGGAATATGCCCTCTTATCAAACCTCAAACCAACTCTCCTAGGCGGGCACGTAGTGTAGCATCCCAAAGTA
659 S Y E C D I P I 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
3501 TCATTGCCTACACAATGAGCCTCGGTGCTGAGAATTCTGTGCGCTACAGCAACAACCTCATTGCTATCCCTACTAACTTCACAATCAGTGTGACAACCTGA
692 I I A Y T M S L G 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
3601 AATTCTGCCCGTATCTATGACCAAAAACAGCGTTGACTGCACCATGTACATCTGTGGCGATTCTACCGAATGTAGCAATCTCCTCCTGCAATACGGATCA
725 I L P V S M T K 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
3701 TTCTGCACTCAGTGAATCGTGCCCTCACAGTATTGCAGTTGAGCAGGACAAGAATACGCAGGAAGTGTTCGCCAGGTGAAGCAAACTACAAAACCTC
759 F C T Q L N R A 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
3801 CACCCATAAAAAGACTTTGGCGGATTCAATTTCTCACAGATCCCTGCCGATCCCTCAAACCTCCAAGCGTAGCTTTATCGAGGATCTGCTCTTCAACAA
792 P P I K D F G G F 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
3901 GGTAAACCTCGCAGATGCCGTTTCATCAAGCAGTATGGCGATTGTCTGGGAGACATCGCCGCTCGGGACCTGATCTGTGCACAGAAGTTCATGGACTG
825 V T L A D A G F 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
4001 ACCGTGCTGCCTCCCTTGTGACCGACGAGATGATAGCCCAATACACTAGCGCCCTGTGCGCCGACCATCACTTCTGGGTGACATTCCGGAGCTGGCG
859 T V L P P L L T 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
4101 CTGCCCTCAGATTCCTTTGCTATGCAGATGGCCTACCGCTTAAACGGCATCGGTGTGACACAAAACGTTCTGTATGAAAACAGAACTCATGCCCAA
892 A A L Q I P F A M 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
4201 CCAAGTTCAACAGTGTATCGGTAAGATACAGGATAGCCTGTATCCACTGCCAGCGCATTGGGAAAAGTTGCAGGATGTAGTGAACCAGAATGCCAGGCA
925 Q F N S A I G K 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
4301 CTTAACACCTGGTGAACAGCTCTCTTCAAATTTGGTGCCATTTCTAGCGTGTGAATGACATACTGAGCCGTTGGACAAGGTGGAGGCTGAAGTGC
959 L N T L V K Q L 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
4401 AGATTGATAGGCTGATAACTGGGCGCCTCAGTCTCTCAGACCTATGTGACCAGCAGCTCATCCGCGCTGTGAAATTCGCGCATCCGCTAACCTGGC
992 Q I D R L I T G R 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
4501 AGCAACCAAAATGCCGAGTGTGTGCTGGGTGAGTCTAAGAGAGTGGACTTTTGGCGGAAGGGGTATCACCTGATGTCTTTTCTCAGTCTGCACCCCAT
1025 A T K M S E C V L G Q S K R V D F C G K G Y H L M S F P Q S A P H
4601 GGTGTGGTCTTTCTGCACGTGACTTATGTCCAGCTCAGGAAAAGAACTTCACTACAGCCCGACCATCTGCCACGATGGGAAAAGCCACTTCCAGGG
1059 G V V F L H V T 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
4701 AAGCGTATTCGTGTCCAATGGTACTCATTGGTTCGCTACTCAGAGAAATTTCTACGAGCCCGAGATTATAACCACTGACAATACATTTGTATCCGGCAA
1092 E G V F V S N G T 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
4801 TTGTGATGTGGTTATCGGGATTGTGAATAAATACTGTTTACGATCCTTTGCAGCCAGAGCTGGACTCCTTCAAGGAGGAGCTTGACAAAATTTTAAAGAT
1125 C D V V I G I V N N T V Y D P L Q P E L D C S F K E E L D K Y F K N
4901 CACATACCTGACGTGCACCTCGGAGATATTTAGGAATCAATGCTTCCGTGCTCAATATTGAGGAGATAGACAGGCTGAATGAGGTTGCCAAGA
1159 H T S P D V D L 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
5001 ACCTCAACGAGTCTGTATCGATCTGCAGGAGTTGGGCAAGTACGAACAGTATATCAAATGGCCATGGTACATTTGGCTTGGGTTTATTGCTGGGCTGAT
1192 N L N E S L I D L 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

Nhel (5192)

5101 AGCTATCGTCATGGTGACAATTATGTTGTGTTGCATGACATCTGCTGTAGTTGTCTGAAGGGCTGCTGCTCATGCGGCAGCTGTTGCTAAAGCTAGCCT
1225 A I V M V T I M L C C M T S C C S C L K G C C S C G S C C •
5201 CGAGGGATCCGTCGAGGAATCACTCCTCAGGTGCAGGCTGCCTATCAGAAGTGGTGGCTGGTGGCCAATGCCCTGGCTCACAATACCACTGAGAT
5301 CTTTTTCCCTCTGCCAAAATTATGGGACATCATGAAGCCCTTGAGCATCTGACTTCTGGCTAATAAAGGAAATTTATTTTCATTGCAATAGTGTGT
5401 GGAATTTTTTGTGCTCTCACTCGGAAGGACATATGGGAGGGCAAATCATTAAAACATCAGAATGAGTATTTGGTTTAGAGTTTGGCAACATATGCCCA
5501 TATGCTGGCTGCCATGAACAAAGTTGGCTATAAAGAGGTCATCAGTATATGAAACAGCCCTGCTGTCCATTCTTATTCCATAGAAAAGCCTTGACT
5601 TGAGGTTAGATTTTTTTATATTTTGTGTTATTTTTTCTTTAACATCCCTAAAATTTTCTTACATGTTTTACTAGCCAGATTTTTCTCCTCTCT
5701 CCTGACTACTCCAGTCATAGCTGTCCCTTCTCTTATGGAGATCCCTCGACGGATCGGCCCAATTCGTAATCATGTCATAGCTGTTTCTGTGTGAA
5801 ATTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAAGTAACTCACATTAATTGCGTT
5901 GCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCTGCCAGCTGCATTAATGAATCGCCAACCGCGGGGAGAGGCGGTTTGCATTGGGCGCTCT
6001 TCCGTTCTCGTCACTGACTCGCTGCGCTCGGTCTGCTGCGGCGAGCGGTATCAGCTCACTCAAAGCGGTAATACGGTTATCCACAGAATCAG
6101 GGGATAACGCAGGAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCC
6201 TGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCC
6301 TCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTT
6401 CGGTGTAGGTCGTTTCGCTCAAGCTGGGCTGTGTGCACGAACCCCGCTTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCAACCC
6501 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCT

6601 AACTACGGCTACTACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAA
6701 CCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACCGCGAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA
6801 CGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTAAATTAATAAATGAAGTTTTAAA
6901 TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTCCATCAG
287 • W H K I L S A G I E A I Q R N R E D M T
7001 TTGCTGACTCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC
266 A Q S G T T Y I V V I R S P K G D P G L A A I I G R S G R E G A G
7101 AGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGAA
233 S K D A I F W G A P L A S R L L P G A V K D A E M W D I L Q Q R S
7201 GCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGGCGAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTTTGGTATGGCTTCATTCA
199 A L T L L E G T L L K R L T T A M A V P M T T D R E D N P I A E N L
7301 GCTCCGGTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGTTAGTCTTCGGTCTCCGATCGTTGTGCAAGAAGTAAAGTT
166 E P E W R D L R T V H D G M N H L F A T L E K P G G I T T L L L N
7401 GGCCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACC
133 A A T N D S M T I A A S C L E R V T M G D T L H K E T V P S Y E V
7501 AAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGGCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCA
99 L D N Q S Y H I R R G L Q E Q G A D I R S L V A G C L L V K F T S M
7601 TCATTGGAAAACGTTCTTCGGGGCGAAAACCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAGTATCTTCAGC
66 M P F R E E P R F S E L I K G S N L D L E I Y G V R A G L Q D E A
7701 ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATA
33 D K V K V L T E P H A F V P L C F A A F F P I L A V R F H Q I S M
7801 CTCTTCTTTTTCAATATTATTGAAGCATTATCAGGGTATTGTCTCATGAGCGGATACATATTTGAATGATTTAGAAAAATAAACAAATAGGGGTTCC
7901 CGCGCACATTTCCCGAAAAGTGCCACCTAAATTTGTAAGCGTTAATATTTTGTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAA
8001 TAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGAACG
8101 TGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAA
8201 AGCACTAAATCGAACCCTAAAGGGAGCCCCGATTAGAGCTTGACGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGAAAGCGAAAGGAGCG
8301 GGCCTAGGGCGCTGGCAAGTGTAGCGTACGCTGCGCGTAACCACACACCCGCCGCGTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCA
8401 GGCTGCGCAACTGTTGGAAAGGGCGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAAC
8501 GCCAGGTTTTCCAGTCACGACGTTGAAAACGACGGCCAGTGAGCGCGGTAATACGACTCACTATAGGGCGAATTGGAGCTCACCGCGGTGGCGGC
8601 CGCTCTAGA