



1 GGACCTGCAGGGCTGAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACAGCTGTGGAATGTGTGTCAGTTAGGGTGTG
 101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCAGGCTCCCAGCAGGCAG
 201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTCTCCGGTCCCGTCACTGGGAGAGCGCACATCGCCACAGTCCCCGA
 301 GAAAGTGGGGGAGGGGTGCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACCTGGGAAAGTGTGTCGTGACTGGTCCGCTTTTTCCC
 401 GAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTC
 501 GCATCTCTCTTACGCGCCCGCCGCCCTACCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGC
 601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCCTCACGCTTTGC

Agel (791)

701 CTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCTGTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCAA
 801 CATGTTTGTGTTCTTGGTGTGCTTCCACTGGTCAAGTCCCAATGCGTTAATCTCACCACCCGAACTCACTCCACCCGCATATACAAATTCCTTACC
 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 AGAGGAGTGTACTATCCTGACAAAGTGTTCGGTCAAGTGTCTCCACTCTACTCAGGACCTCTTTCTGCCTTTCTTTCTAACGTTACATGGTTTCATG
 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
 T951 (1083)
 1001 CAATCCATGTGTCTGGGACAAACGGCACCAACGCTTCGACAACCCTGTATTGCCATTCAATGATGGGGTGTACTTTGCCTCCATCGAGAAATCCAACAT
 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 I E K S N I
 1101 CATTGAGGATGGATTTTCGGGACTACTCTGGACTCAAAGACACAGAGCCTGTGATCGTTAAACAACGCCACAAACGTTGTATCAAAAGTGTGCGGAATTC
 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

ins145N (1236)
Y145S (1233)
Y144T (1230)

1201 CAGTTTTGCAATGATCCCTTCTGGGAGTGCACCAAGAATAACAAGTCTGGATGGAGAGCGAATTTCCGGTCTACAGCAGCGCAAACA
 134 Q F C N D P F L G V T S N H K N N K S W M E S E F R V Y S S A N N
 1301 GCACCTTCGAGTACGTGAGTCAACCTTTCTGATGGACCTGGAAGGAAACAGGAAACTTCAAGAACCTGAGAGAGTTTGTCTTTAAGAATCGACGG
 167 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 F V F K N I D G
 1401 CTATTTAAGATCTATAGTAAGCATACGCCATCAACCTGGTAAGGGATCTTCCCGAGGGCTTTCCAGCCCTGGAACCTTTGGTTGACTTGCCATTGGT
 200 Y 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
 1501 ATCAATATCACCAGATTTAGACCTTCTGGCATTGCAICGGTCTTATCTTACTCCAGGTGATTCCTCCTCCGGGTGGACTGCCGGCGCCGCTGCCTACT
 234 I N I T R F Q T L L A L H R S Y L T P G D S S G W T A G A A A Y
 1601 ATGTCGGTATCTGCAACGAAGACCTTCTGCTCAAGTACAACGAAACGGCACTATTACGGATGCTGTTGATTGTGCCCTGGACCTCTGTGAGAC
 267 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 A L D P L S E T
 1701 TAAATGCACCCTCAAGAGCTTTACCGTTGAGAAGGGGATTTACCAAACAGTAATTTCCGGGTCACCCACCGAAAGCATTGTGCGGTTCCCAAATATC
 300 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 I V R F P N I

R246K (1839)

1801 ACCAATCTGTGTCCCTTTGGCGAAGTGTCAATGCTACAAGTTTGTCTGTGTACGCATGGAATAGGAAACGCATCTCCAATTGTGTGCTGATTACT
 334 T N L C P F G E V F N A T K F A S V Y A W N R K R I S N C V A D Y
 1901 CCGTGTGTACAATTCCGCTCTTTCTCAACCTTCAAGTGTATGGCGTTTACCTACCAAACCTAACGACCTGTGCTTCACTAATGTGTATGCCGACTC
 367 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 F T N V Y A D S
 2001 TTTTGTGATACGAGGCGATGAAGTGAAGTGTGACAGATTTGACCCAGGGCAGACCGGCAAAATGCCGACTACAACCTACAAGCTTCCAGATGACTTTACCGGATGT
 400 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 P D D F T G C
 2101 GTTATTGCATGGAACCTCAACAACTCTGGATTCCAAGTGGGTGGCAACTATAACTACCTGTATAGACTGTTCAAGAAATCCAACCTGAAACCATTCGAGC
 434 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 S N L K P F E

E484K (2253)

2201 GAGATATAAGCACAGAAATCTACCAGGCTGGAAGTACGCCCTGCAACGGCGTGAGGGGTTCAACTGCTACTTCCCATTGCAGAGTTACGGATTCCAGCC
 467 R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P

N501Y (2304)

2301 TACAATGGGGTGGGATACCAACCTATCGTGTGCTAGTCTGAGTTTGTGCTCCTCCATGCCCCAGCCACAGTCTGTGCCCAAGAAAAGCACCAAT
 500 T Y G V G Y Q P Y R V V L S F E L L H A P A T V C G P K K S T N
 2401 CTGGTGAAGAACAATGCGTGAACCTTAACTTTAACGACTCACAGGAACCGCGTATTGACGGAGAGTAACAAGAAGTTCTGCCATTCCAGCAGTTCCG
 534 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 F L P F Q Q F
 2501 GTCGCGATATTGCCGACACTACCGACGCTGTCCGAGATCCCCAGACATTGGAGATTCTTGATATCACACCCTGTAGTTTCCGGCGAGTGAGCGTGATTAC
 567 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 F G G V S V I T

D614G (2643)

2601 GCCCGAACCAATACCAGCAATCAGGTTGCCGTCCTGTATCAGGGCTGAATTGCACCGAGGTACCTGTGCCATCCACGCTGACCAACTACACCCACA
600 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 A D Q L T P T
2701 TGGCGAGTATATCCACCGGCTCCAACGCTTTTCAGACACGTGCTGGATGTCTGATCGGTGCAGAACCGTTAATAATAGCTACGAGTGTGATATCCCCA
634 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 S Y E C D I P

R683A (2850)

P681H (2844) R685A (2856)

2801 TCGGTGCTGGAATATGCGCCTCTTATCAAACCTCAAACCAACTCTCACAGGGCGGCAAGCTAGTGTAGCATCCCAAAGTATCATTGCCTACACAATGAGCCT
667 I G A G I C A S Y Q T Q T N S H R A A A S V A S Q S I I A Y T M S L
2901 CGGTGCTGAGAATTCTGTCGCTACAGCAACAACTCCATTGCTATCCCTACTAACTTACAATCAGTGTGACAACTGAAATTTGCCGTATCTATGACC
700 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 I L P V S M T
3001 AAAACAAGCGTTGACTGCACCATGTACATCTGTGGCGATTCTACCGAATGTAGCAATCTCCTCCTGCAATACGGATCATTCTGCACCTCAGTGAATCGTG
734 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 F C T Q L N R
3101 CCCTCACAGGTATTGCAGTTGAGCAGGACAAGAATACGAGGAAGTGTGGCCAGGTGAAGCAAATCTACAAAACCTCCACCATAAAAGACTTTGGCGG
767 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 P P I K D F G G
3201 ATTCAATTTCTCACAGATCCTGCCCGATCCCTCAAACCTCCAAGCGTAGCTTTATCGAGGATCTGCTTTCAACAAGGTAACCCCTGCAGATGCCGGT
800 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 V T L L A D A G
3301 TTCATCAAGCAGTATGGCGATTGTCTGGGAGACATCGCCGCTCGGACCTGATCTGTGCACAGAAAGTTCAATGGACTGACCTGCTCCCTCCCTTGGGA
834 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 T V L P P L L
3401 CCGACGAGATGATAGCCCAATACACTAGCGCCTGCTGGCGGACCATCACTTCTGGGTGGACATTCGGAGCTGGCGCTGCCCTTCAGATTCCTTTTGC
867 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 A A L Q I P F A
3501 TATGCAGATGGCTACCCTTTAACGGCATCGGTGTGACACAAAACGTTCTGTATGAAAACAGAAACTCATCGCCAACAGTTCAACAGTGTATCGGT
900 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 Q F N S A I G

D950N (3651)

3601 AAGATACAGGATAGCCTGTCATCCACTGCCAGCGCATTGGGAAAGTTGCAGAACGTAGTGAACCAAGATGCCAGGCACTTAACACCTGGTGAACAGC
934 K I Q D S L S S T A S A L G K L Q N V V N Q N A Q A L N T L V K Q
3701 TCTCTTCAAATTTTGGTGCCATTTCTAGCGTGTGAATGACATACTGAGCCGTTGGACAAGTGGAGGCTGAAGTGCAGATTGATAGGCTGATAACTGG
967 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 Q I D R L I T G
3801 GCGCCTTCAGTCTCTTACAGCCTATGTGACCCAGCAGCTCATCCGCGCTGCTGAAATTCGCGCATCCGCTAACCTGGCAGCAACCAAAATGTCGAGTGT
1000 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 A T K M S E C
3901 GTGCTGGTTCAGTCTAAGAGAGTGGACTTTTGGGGAAGGGTATCACCTGATGTCTTTCTCAGTCTGCACCCCATGGTGTGGTCTTTCTGCACGTGA
1034 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 G V V F L H V
4001 CTTATGTCCCAGCTCAGGAAAAGAACTTCACTACAGCCCCAGCCATCTGCCACGATGGGAAAGCCACTTTCCAGGGAAAGCGTATTCTGTCCAATGG
1067 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 E G V F V S N G
4101 TACTCATTGGTTCGCTCACTCAGAGAAATTTCTACGAGCCCCAGATTATAACCACTGACAATACATTTGTATCCGGCAATTGTGATGTGGTTATCGGGATT
1100 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 C D V V I G I
4201 GTGAATAACTGTTTACGATCCTTTCAGCCAGAGCTGGACTCCTTCAAGGAGGAGCTTGACAAATATTTTAAAGAACACACATCACCTGACGTGACCC
1134 V 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
4301 TCGGAGATATTTCAGGAATCAATGCTTCCGTGGTCAATATTCAGAAGGAGATAGACAGGCTGAATGAGGTTGCCAAGAACCTCAACGAGTCTCTGATCGA
1167 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 N L N E S L I D
4401 TCTGCAGGAGTTGGGCAAGTACGAACAGTATATCAAATGGCCTGGTACATTTGGCTTGGGTTTATTGCTGGGCTGATAGCTATCGTCATGGTGAACAAT
1200 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 A I V M V T I

NheI (4570)

4501 ATGTTGTGTTGCATGACATCCTGCTGTAGTTGTCTGAAGGGCTGCTGCTCATCGGACAGTGTGCTAAAGCTAGTGGCCAGACATGATAAGATACATT
1234 M L C C M T S C C S C L K G C C S C G S C C •
4601 GATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAGAAAATGCTTTATTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCA
4701 ATAAACAAGTTAAACAACAACAATTGCATTTCATTTTATGTTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGG
4801 TATGGAATTCTAAATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGG
4901 CTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAACCTAGCTCTTCATTTCTT
5001 TATGTTTTAAATGCACTGACCTCCACATTCCTTTTTCAGTAAATATTCAGAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAG
5101 GCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTTGGACAGCAAGAAAG
5201 CGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTTCATCTCAATGAGCACAAGCAGTC
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
5301 AGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACA
112 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
5401 ATGGTGTCAAAGCTTCTGCCGTTGCTCACAGCAGACCCAAATGGCAATGGCTTCAGCAGACAGCAGTACCCTGCCAATGTAGGCCCTCAATGTGGACG
78 I 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
5501 CAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCGACATGGTGTGTTGCTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAG
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
5601 CTCGAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTC
12 E L D Q Q S I N F T K M
5701 AAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCCTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTC

5801 AATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCC GTT GATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCC
5901 GTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAA
6001 GTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACT
6101 GCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATG
6201 GGC GGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGCAGGTTAATTAAAGAACATGTGAGCAAAGGCCAGCAAAG
6301 GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAA
6401 ACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTT
6501 TCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCC
6601 CCCGTT CAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA
6701 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCT
6801 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGCAAGCAGCAGATT
6901 ACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGG
7001 CTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACTAACATACGCT
7101 CTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA