



1 GGACCTGCAGGGCCTGAAATAACCTCTGAAAGAGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTAGGTTAGGGTGTG
101 GAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCCAGGCTCCCCAGCAGGCAG
201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCACTAGTCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGA
301 GAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGTACTGGCTCCGCCTTTTTCC
401 GAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTC
501 GCATCTCTCTTACGCGCCCGCCGCCCTACCTGAGGCGCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAACTGC
601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGC

701 CTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCTGTTTCTGCTGCGCAGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGTCAA
801 CATGTTTGTGTTCTTGGTGTGCTTCCACTGGTCAAGTTCCTCAATGCGTTAATCTCACCCCGAACTCAACTCCCACCCGCATATACAAATTCCTTACC
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 AGAGGAGTGTACTATCCTGACAAAGTGTTCGGTCAAGTGTCTCCACTCTACTCAGGACCTCTTCTGCCTTTCTTTCTAACGTTACATGGTTTCATG
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
Δ H69+V70 (1004) T951 (1077)
1001 TGATCTCTGGGACAAACGGCACAAACGCTTCGACAACCCTGTATTGCCATTCAATGATGGGTGTACTTTGCCTCCATCGAGAAATCCAACATCATTCC
67 V I 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 I R
1101 AGGATGGATTTTCGGGACTACTCTGGACTCAAAGACACAGAGCCTGTGATCGTTAAACAACGCCACAAACGTTGTATCAAAGTGTGCGAATTCCAGTTT
100 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 Q F

1201 TGCAATGATCCCTTCTCTGGACACAAGAATAACAAGTCTGGATGGAGAGCGAATTCGGGTCTACAGCAGCGCAAACAACGACCTTCGAGTACGTGA
134 C N D P F L D H K N N K S W M E S E F R V Y S S A N N C T F E Y V
1301 GTCAACCTTTCTGATGGACCTGGAAGGGAACAGGGAAACTTCAAGAACCTGAGAGAGTTTGTCTTTAAGAATCATCGACGGCTATTTTAAGATCTATAG
167 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 F K I Y S
ins214EPE (1425)
ΔN211,L212I (1416)

1401 TAAGCATACGCCTATCATTGTAAGGGAGCCCGAGGATCTTCCCAGGGCTTTTCAGCCCTGGAACCTTTGGTTGACTTGCCCTATTGGTATCAATATCACC
200 K H T P I I V R E P E D L P Q G F S A L E P L V D L P I G I N I T
1501 AGATTTAGACCTTCTGGCATTGCATCGGTCTTACTTCCAGGTGATTCCTCTCCGGGTGGACTGCCGGCGCCGCTGCCTACTATGTCGGCTATC
234 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 V G Y
1601 TGCAACCAAGACGTTCTGCTCAAGTACAACGAAACGGCACTATTACGGATGCTGTTGATTGTGCCCTGGACCTCTGTCTGAGACTAAATGCACCT
267 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 C T L
1701 CAAGAGCTTTACCGTTGAGAAGGGGATTTACCAAACAGTAATTTCCGGGTCCAACCCAGGAAAGCATTGTGCGGTTCCCAAATATCACCATCTGTGT
300 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 N L C

1801 CCCTTTGATGAAGTGTTCATGCTACAAGTTTGTCTTCTGTGTACGCATGGAATAGGAAACGCATCTCCAATTGTGTGCTGATTACTCCGTGCTGTACA
334 P F D 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 V L Y

S373P (1908)
S371L (1902) S375F (1914)
1901 ATCTGGCCCAATCTTACCTTCAAGTGTATGGCGTTTACCTACCAAACCTAACGACCTGTGCTTCACTAATGTGTATGCCGACTTTTTGTGATACG
367 N L A P F F 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 V I R

K417N (2040)
2001 AGGCGATGAAGTGAACAGATTGCACCAGGGCAGACCGGCACATTGCCGACTACAACCTCAAGCTTCCAGATGACTTTACCGGATGTGTTATTGCATGG
400 G D E V R Q I A P G Q T G N I A D Y N Y K L P D D F T G C V I A W

N440K (2109) G446S (2127)
2101 AACTCAAACAGCTGGATTCCAAGGTGAGCGCAACTATAACTACCTGTATAGACTGTTCCAGGAAATCCAACCTGAAACCTTCGAGCGAGATATAAGCA
434 N S N K L D S K V S G N Y N Y L Y R L F R K S N L K P F E R D I S

S477N, T478K (2220) E484A (2241) G496S (2277) N501Y (2292)
Q493R (2268) Q498R (2283)
2201 CAGAAATCTACCAGGCTGGAACAAAACCTGCAACGGCGTGGCTGGGTTCAACTGCTACTTCCATTGCGCAGTTACAGCTTCACTACATACGGGGT
467 T E I Y Q A G N K P C N G V A G F N C Y F P L R S Y S F R P T Y G V

Y505H (2304)

2301 GGGTCAACCAACCTATCGTGTCTAGTCTGAGTTTTGAGCTCTCCATGCCCCAGCCACAGTCTGTGGCCCAAGAAAAGCACCAATCTGGTGAAGAAC
500▶ G H 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 V K N

T547K (2430)

2401 AAATGCGTGAACCTTTAACTTTAACGGACTCAAGGGAACCGGCGTATTGACGGAGAGTAACAAGAAGTTCCTGCCATTCCAGCAGTTCGGTCGCGATATTG
534▶ K C V N F N F N G L K G T G V L T E S N K K F L P F Q Q F G R D I

2501 CCGACTACCGACGCTGTCCGAGATCCCAGACATTGGAGATTCTTGATATCACACCTGTAGTTTCGGCGGAGTGAGCGTGATTACGCCCGAACCA
567▶ 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 G T N

D614G (2631)

2601 TACCAGCAATCAGGTTGCCGCTGTATCAGGGCGTGAATTGACCCGAGGTACCTGTGCCATCCACGCTGACCAACTTACACCCACATGGCGAGTATAT
600▶ 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 W R V Y

H655Y (2754)

2701 TCCACCGCTCCAACGTCTTTCAGACAGTGTCTGATCGGTGCAGAAATGTTAATAATAGCTACGAGTGTGATATCCCATCGGTGCTGGAA
634▶ S T G S N V F Q T R A G C L I G A E Y V N N S Y E C D I P I G A G

P681H (2832)

N679K (2826) R683A + R685A (2838)

2801 TATGCGCTCTTATCAAACCTCAAACCAATCTCACAGGCGGCGAGCTAGTGTAGCATCCCAAAGTATCATTGCCTACACAATGAGCCTCGGTGCTGAGAA
667▶ I C A S Y Q T Q T K S H R A A A S V A S Q S I I A Y T M S L G A E N

2901 TTCTGTGCGCTACAGCAACAACCTCATTGCTATCCCTACTAACTTACAATCAGTGTGACAACCTGAAATTCTGCCGCTATCTATGACCAAAAACAGCGTT
700▶ 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 T S V

N764K (3081)

3001 GACTGCACCATGTACATCTGTGGCATTCTACCGAATGTAGCAATCTCTCTGCAATACGGATCATTCTGCACTCAGCTGAAGCGTGCCTCACAGGTA
734▶ 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 K R A L T G

D796Y (3177)

3101 TTGCAGTTGAGCAGGACAAGAATACGCAGGAAGTGTGGCCAGGTGAAGCAAACTACAAAACTCCACCCATAAAAATCTTTGGCGGATTCAATTTCTC
767▶ 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 Y F G G F N F S

3201 ACAGATCCTGCCGATCCCTCAAACCTCAAGCGTAGCTTTATCGAGGATCTGCTCTTCAACAAGGTAACCTCGCAGATGCCGTTTTCATCAAGCAG
800▶ 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 I K Q

N856K (3357)

3301 TATGGCGATTGTCTGGGAGACATCGCCGCTCGGGACCTGATCTGTGCACAGAAGTTCAAAAGGACTGACCGTGTGCCTCCCTTGTGACCGACGAGATGA
834▶ Y G D C L G D I A A R D L I C A Q K F K G L T V L P P L L T D E M

3401 TAGCCCAATACTAGCGCCCTGCTGGCCGCGCACCATCACTTCTGGGTGGACATTCGGAGCTGGCGTGCCTTCAGATTCTTTTGTATGCGATGGC
867▶ 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 Q M A

3501 CTACCGCTTTAACGGCATCGGTGTGACACAAAACGTTCTGTATGAAAACAGAAACTCATCGCAACCAAGTCAACAGTGTATCGGTAAGATACAGGAT
900▶ 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 I Q D

Q954H (3651)

N969K (3696)

3601 AGCCTGTCCACTGCCAGCGCATTGGGAAAGTGCAGGATGTAGTGAACCAATGCCAGGCACTTAACACCCTGGTGAACAGCTCTCTTCAAGT
934▶ S L S S T A S A L G K L Q D V V N H N A Q A L N T L V K Q L S S K

L981F (3732)

3701 TTGGTGCCATTTCTAGCGTGTGAATGACATAATAGCCGGTTGGACAAGGTGGAGGCTGAAGTGCAGATTGATAGGCTGATAACTGGGCGCCTTCACTC
967▶ F G A I S S V L N D I F S R L D K V E A E V Q I D R L I T G R L Q S

3801 TCTTCAGACCTATGTGACCCAGCAGCTCATCCGCGTGTGAAATTCGCGCATCCGTAACCTGGCAGCAACCAAAATGTCGAGTGTGTGCTGGGTGAG
1000▶ 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 L G Q

3901 TCTAAGAGAGTGGACTTTTGGCGGAAGGGTATCACCTGATGCTTTTCTCAGTCTGCACCCATGGTGTGGTCTTTCTGCACGTACTTATGTTCCAG
1034▶ 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 Y V P

4001 CTCAGGAAAAGAACTTACTACAGCCCGCAGCCATCTGCCAGATGGGAAAGCCACTTTCCAGGGAAGGCGTATTCTGTCCAATGGTACTCATTGGTT
1067▶ 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 H W F

4101 CGTCACTCAGAGAAATTTCTACGAGCCAGATTATAACCACTGACAATACATTTGTATCCGCAATTGTGATGTGGTTATCGGGATTGTGAATAACT
1100▶ 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 N N T

4201 GTTTACGATCCTTTCAGCCAGAGCTGGACTCCTTCAAGGAGGAGCTTGACAAAATTTTTAAGAATCACACATCACCTGACGTGCAGCTCGGAGATATTT
1134▶ 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 G D I

4301 CAGGAATCAATGCTTCCGTGGTCAATATTGAGAAGGAGATAGACAGGCTGAATGAGGTTGCCAAGAACCTCAACGAGTCTCTGATCGATCTGCAGGAGTT
1167▶ 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 Q E L

4401 GGGCAAGTACGAACAGTATATAAATGGCTTGGTACATTTGGCTTGGGTTCAATTGCTGGGCTGATAGCTATCGTCATGGTGAACATATGTTGTGTTGC
1200▶ 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 L C C

NheI (4558)

4501 ATGACATCTGCTGTAGTTGTCTGAAGGGCTGCTGCTCATGCGGAGCTGTTGCTAAAGCTAGTGGCCAGACATGATAAGATACATTGATGAGTTTGG
1234▶ M T S C C S C L K G C C S C G S C C •

4601 CAAACCAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTA

4701 ACAACAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTA

4801 AAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATG

4901 TGCATTAGTGTGTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATAGTGTATTTTCCCAAGGTTTGAACCTAGCTCTTCAATTTCTTATGTTTAAAT

5001 GCACTGACCTCCACATTCCCTTTTTAGTAAATATTGAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAATCCAGA

5101 TGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAATAGAAATTGGACAGCAAGAAAAGCGAGCTTCTAGC

5201 TTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTC
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 Y D

5301 AGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAG
108 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 D F

5401 TCCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCT
74 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 I I E

5501 CCCAGTCTTGGTCTGTATGGCCGCCCGACATGGTCTTGTTCCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTG
41 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 D Q

5601 CTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACGCGTGG
8 Q S I N F T K M

5701 ATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCCATTTGCGTCAATGGGGCGGAG

5801 TTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACC

5901 GCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAAGT

6001 CATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCA

6101 GTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGT

6201 TGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT
BspLU11l (6262)

6301 AAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC

6401 TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGG

6501 AAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCC

6601 GACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG

6701 CGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT

6801 TACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAA

6901 AAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAA

7001 CATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAC

7101 AAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA