



1 GGACCTGCAGGGCCTGAAATAACCTCTGAAAGAGGAAGCTTGGTTAGGTACCTTCTGAGCGGAAAGAACAGCTGTGGAATGTGTGTCAGTTAGGGTGTG
101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCAGGCTCCCAGCAGGCAG
201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTCTCCGGTGGCCGTGAGTGGGAGAGCGCACATCGCCACAGTCCCCGA
301 GAAGTTGGGGGAGGGGTGCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACCTGGGAAAGTGTGCTGTACTGGTCCGCTTTTTCC
401 GAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGGCCGAGAACAGCTGAAGCTTCGAGGGGCTC
501 GCATCTCTCTTACGCGCCCGCCCTACCTGAGGCCGATCCACGCGGTTGAGTGCAGTCTGCGCCTCCGCTGTGGTGCCTCTGAACTGC
601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGC

BspLU11(799)

Agel (791)

701 CTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCTGTTTCTGTTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCAA

801 CATGTTTGTGTTCTTGGTGTGCTTCCACTGGTCAGTCCCAATGCGTTAATCTCAATCACCCGAACTCAATCCCTATACAAATTCCTTACCAGAGGAGTG
1 M F V F L V L L P L V S S Q C V N L I T R T Q S Y T N S F T R G V
901 TACTATCCTGACAAAGTGTTCGGTCAAGTGTCTCCACTCTACTCAGGACCTCTTCTGCCTTCTTTCTAACGTTACATGTTTCATGCAATCCATG
34 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 A I H

1001 TGTCTGGGACAAACGGCACAAACGCTTCGACAACCCTGCAATGCCATTCAATGATGGGTGTACTTTGCCTCCACAGAGAAATCCAACATCATTGAGG
67 V S G T N G T K R F D N P A L P F N D G V Y F A S T E K S N I I R G
1101 ATGGATTTTCGGGACTACTCTGGACTCAAAGACACAGAGCCTGCTGATCGTTAAACACGCCACAAACGTTGTCATCAAAGTGTGCAATTCCAGTTTTC
100 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 C
Δ Y144 + H146Q (1223)

1201 AATGATCCCTTCTGACGCTGTACCCAGAAAGAATAACAAGTCTGGATGGAGAGCGAATTTTCGGGTCTACAGCAGCGCAAACTGCACCTTCGAGTACG
134 N D P F L D V Y Q K N N K S W M E S E F R V Y S S A N N C T F E Y
Q183E (1335)

1301 TGAGTCAACCCTTCTGATGGACCTGGAAGGGAAAAGAGGGAACTTCAAGAACCTGAGAGAGTTTGTCTTTAAGAATCGACGGCTATTTTAAAGATCTA
167 V S Q P F L M D L E G K E G N F K N L R E F V F K N I D G Y F K I Y

1401 TAGTAAGCATACGCTATCAACCTGGAAGGGATCTTCCCAGGGCTTTTCAGCCCTGGAACCTTTGGTTGACTTGCCTATTGGTATCAATATCACCAGA
200 S K H T P I N L E R D L P Q G F S A L E P L V D L P I G I N I T R
G252V (1542)

1501 TTTAGACCCTTCTGGCATTGCAICGGTCTTATCTTACTCCAGTGGATTCTCTCCTCCGGGTGGACTGCCGGCGCCGCTGCCTACTATGTCGGCTATCTGC
234 F Q T L L A L H R S Y L T P V D S S S G W T A G A A A Y Y V G Y L
1601 AACCAAGAAGTCTCTGCTCAAGTACAACGAAACGGCACTATTACGGATGCTGTTGATTGTCGCCCTGGACCTCTGCTGAGACTAAATGCACCCTCAA
267 Q P R T 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 K
1701 GAGCTTTACCGTTGAGAAGGGGATTTACCAAACAGTAATTTCCGGGTCCAACCCACCGAAAGCATTGTGCGGTTCCCAAATATCACCAATCTGTGTCC
300 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 P

S371F (1899)

1801 TTTCAITGAAGTGTTCATGCTACAACCTTTGCTTCTGTGTACGCATGGAATAGGAAACGCATCTCCAATTGTGTCGCTGATTACTCCGTGATCTACAATI
334 F H E V F N A T T F A S V Y A W N R K R I S N C V A D Y S V I Y N

1901 TIGCCCAATTCITCGCTTCAAGTGTATGGCGTTTCCACTACCAAATTAACGACCTGTGCTTCACTAATGTGTATGCCGACTCTTTGTGATACGAGG
367 F A P F F A 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 G
T376A (1914)
S375F (1911)
S373P (1905)

2001 CAATGAAGTGAAGCAGATTGCACCAGGGCAGACCGGCACATTGCCGACTACAACCTACAAGCTTCCAGATGACTTTACCGGATGTGTTATTGCATGGAAC
400 N E V S 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 N
R408S (2010)
D405N (2001)
K417N (2037)

2101 TCAAACAAAGCTGGATTCCAAGCCATCCGGCAACTATAACTACCTGTATAGACTGTTACGAAATCCAAAGCTGAAACCAATTCGAGCGAGATATAAGCACAG
434 S N K L D S K P S G N Y N Y L Y R L F R K S K L K P F E R D I S T
N440K (2106)
V445P+G446S (2121)
N460K (2166)

2201 AAATCTACCAGGCTGGAACAACCCCTGCAACGGCGTGGCTGGGCTTAACTGCTACTTCCCATTTGCAGAGTTACGGATTACAGACCTACATACGGGGTGGG
467 E I Y Q A G N K P C N G V A G P N C Y S P L Q S Y G F R P T Y G V G
S477N, T478K (2217)
E484A (2238)
F490S (2256)
Q498R (2280)
F486P (2244)
N501Y (2289)

Y505H (2301)

2301 TCACCAACCTATCGTGTCTAGTCCTGAGTTTTGAGCTCCTCCATGCCCCAGCCACAGTCTGTGGCCCCAAGAAAAGCACCAATCTGGTGAAGAACAAA
500▶ 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 K

2401 TGCCTGAACTTTAACTTTAACGGACTCACAGGAACCGGCGTATTGACGGAGAGTAACAAGAAGTTCCTGCCATTCCAGCAGTTCGGTCGCGATATTGCCG
534▶ 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 R D I A

2501 AACTACCAGCTGTCCGAGATCCCCAGACATTGGAGATTCTTGATATCACACCCTGTAGTTTCGGCGGAGTGAGCGTGATTACGCCCGGAACCAATAC
567▶ 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 T

D614G (2628)

2601 CAGCAATCAGTTGCCGTCTGTATCAGGGCGTGAATTGCACCGAGGTACCTGTCGCATCCACGCTGACCAACTACACCCACATGGCGAGTATATTCC
600▶ 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 S

H655Y (2751)

2701 ACCGGCTCCAACGTCTTTTCCAGACACGTGCTGGATGTCTGATCGGTGCAGAATATGTTAATAATAGCTACGAGTGTGATATCCCCATCGGTGCTGGAATAT
634▶ 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 I

P681H (2829) R685A (2841)

2801 GGCCTCTTATCAAACCTCAAACCAAATCTCACAGGGCGGCAGCTAGTGTAGCATCCCAAAGTATCATTGCCTACACAAATGAGCCTCGGTGCTGAGAATTC
667▶ 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 S

2901 TGTGCGCTACAGCAAACTCCATTGCTATCCCTACTAACTTACAATCAGTGTGACAACTGAAATTCGCCCGTATCTATGACCAAAAACAAGCGTTGAC
700▶ 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 D

N764K (3078)

3001 TGCACCATGTACATCTGTGGCGATTCTACCGAATGTAGCAATCTCCTCCTGCAATACGGATCATTCTGCACTCAGCTGAAGCGTGCCTCACAGGATTG
734▶ 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 I

D796Y (3174)

3101 CAGTTGAGCAGGACAAGAATACGCAGGAAGTGTGGCCAGGTGAAGCAAATCTACAAAACCTCCACCATAAAATACTTTGGCGGATTCAATTTCTCACA
767▶ 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 F N F S Q

3201 GATCCTGCCCGATCCCTCAAACCTCAAGCGTAGCTTTATCGAGGATCTGCTCTTCAACAAGGTAACCCTCGCAGATGCCGTTTCATCAAGCAGTAT
800▶ 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 Y

3301 GGCATTGTCTGGGAGACATCGCCGCTCGGGACCTGATCTGTGCACAGAAGTCAATGGACTGACCGTGTGCCTCCCTTGTGACCGACGAGATGATAG
834▶ 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 D E M I

3401 CCCAATACTAGCGCCCTGCTGGCCGACCATCACTTCTGGGTGGACATTCGGAGCTGGCGCTGCCCTTACAGATCCTTTTGTATGCAGATGGCCTA
867▶ 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 Y

3501 CCGCTTTAACGGCATCGGTGTGACACAAAACGTTCTGTATGAAAACAGAACTCATCGCAACCAGTTCAACAGTGTATCGGTAAGATACAGGATAGC
900▶ 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 S

Q954H (3648)

N969K (3693)

3601 CTGTCATCCACTGCCAGCGCATTGGGAAAGTTGCAGGATGTAGTGAACCACAATGCCAGGCCTTAACACCCTGGTGAACAGCTCTCTTCAAAGTTTG
934▶ 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 F

3701 GTGCCATTTCTAGCGTGTGAATGACATACTGAGCCGTTGGACAAGTGGAGGCTGAAGTGCAAGTGTATAGGCTGATAACTGGGCGCTTCAAGTCTCT
967▶ 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 L Q S L

3801 TCAGACCTATGTGACCCAGCAGCTCATCCGCGTCTGCTGAAATTCGCGCATCCGCTAACCTGGCAGCAACCAAAATGTCGAGTGTGTGCTGGTCAAGTCT
1000▶ 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 D I S

3901 AAGAGAGTGGACTTTTGGGGAAGGGTATCACCTGATGTCTTTTCTCAGTCTGCACCCCATGGTGTGGTCTTTCTGCACGTGACTTATGTCCAGCTC
1034▶ 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 A

4001 AGGAAAAGAACTTCACTACAGCCCCAGCCATCTGCCACGATGGGAAAGCCACTTTCCAGGGAAGGCGTATTCTGTCCAATGGTACTCATTGGTTCGT
1067▶ 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 V

4101 CACTCAGAGAAATTTCTACGAGCCCCAGATTATAACCACTGACAATACATTTGTATCCGGCAATTGTGATGTGGTTATCGGGATTGTGAATAACTGTT
1100▶ 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 V

4201 TACGATCCTTTGACGCCAGAGCTGGACTCCTTCAAGGAGGAGCTTGACAAATATTTAAGAATCACACATCACCTGACGTGACCTCGGAGATATTTTCA
1134▶ 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 S

4301 GAATCAATGCTTCCGTGGTCAATATTAGAAGGAGATAGACAGGCTGAATGAGGTTGCCAAGAACCTCAACGAGTCTCTGATCGATCTGCAGGAGTTGGG
1167▶ 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 G

4401 CAAGTACGAACAGTATATCAAATGGCCTGGTACATTTGGCTTGGGTTTATTGCTGGGCTGATAGCTATCGTCATGGTACAATTATGTTGTGTTGCATG
1200▶ 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 M

NheI (4555)

4501 ACATCCTGCTGATGTTGTCTGAAGGGCTGCTGCTCATGCGGCAGCTGTTGCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAA
1234▶ T S C C S C L K G C C S C G S C C •

4601 ACCACAAC TAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACA

4701 ACAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAA
→ ←

4801 TACAGCATAGCAAAACTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGC

4901 ATTAGCTGTTGCAGCCTCACCTCTTTTATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTGCTCTTCAATTTCTTTATGTTTTAAATGCA

5001 CTGACCTCCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGC

5101 TCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAATTTGGACAGCAAGAAAGCGAGCTTCTAGCTTT

5201 AGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCACTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGA
140 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 S
5301 GATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCC
107 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 D
5401 TTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCC
73 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 G
5501 CAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGTCTG
40 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 Q
5601 AGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATG
7 S I N F T K M
5701 GCGTCTCCAGCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTGCGTCAATGGGGCGGAGTTG
5801 TTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGAGACTTGAAATCCCGTGAGTCAAACCGCT
5901 ATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCAT
6001 GTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACTTGATGTACTGCCAAGTGGGCAGTT
6101 TACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGGTCTGTTGG
6201 GCGGTCAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAA
6301 AAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT
6401 AAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAG
6501 CGTGGCGCTTTTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACAGCCGAC
6601 CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA
6701 GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC
6801 CTTGGAAGAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAA
6901 GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACAT
7001 TTAATCAGCGGCCCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTTGTAATCGTAACTAACATACGCTCTCCATCAAAACAAA
7101 ACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCGAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA