



1 GGACCTGCAGGGCTGAAATAACCTCTGAAAGAGGAACCTGGTTAGGTACCTCTGAGCGGAAAGAACAGCTGTGGAATGTGTGCAGTTAGGGTGTG
101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCAGGCTCCCAGCAGGCAG
201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTCTCCGGTGGCCGTGAGTGGGAGAGCGCACATCGCCACAGTCCCGA
301 GAAGTTGGGGGAGGGGTGCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACCTGGGAAAGTGTGCTGTACTGGCTCCGCTTTTCC
401 GAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTC
501 GCATCTCTCTTACGCGCCCGCCCTACCTGAGGCGCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGC
601 GTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCCTCACGCTTTC

BspLU11 I

Agel (791)

701 CTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCTGTTTCTGTTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGGCCTACCTGAGATCACCGGTCAA

T19I (855)

Δ L24P25P26 + A27S (870)

801 CATGTTTGTGTTCTTGGTGTGCTTCCACTGGTCAGTCCCAATGCGTTAATCTCATCACCCGAACTCAATTCCTATACAAATTCCTTCACCAGAGGAGTG
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

Δ H69V70 (995)

901 TACTATCCTGACAAAGTGTTCGGTCAAGTGTCTCCACTCTACTCAGGACCTCTTCTGCCTTTCTTTTCTAACGTTACATGGTTTCATGCAATCTCTG
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 S
1001 GGACAAACGGCACAAACGCTTCGACAACCTGTATTGCCATTCAATGATGGGGTGTACTTTGCCTCCACAGAGAAATCCAACATCATTGAGGATGGAT
67▶ 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 I R G W I
1101 TTTCCGGACTACTCTGGACTCAAAGACACAGAGCCTGCTGATCGTTAAACAACGCCACAAACGTTGTATCAAAGTGTGCGAATTCCAGTTTTGAATGAT
100▶ 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 N D

G142D (1209)

1201 CCCTTCCTGACGTGTACTATCACAAGAATAACAAGTCTGGATGGAGAGCGAATTTCCGGTCTACAGCAGCGCAACAACACTGCACCTTCGAGTACGTGA
134▶ P F L D V Y Y 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 GTCAACCTTTCTGATGGACCTGGAAGGAAACAGGAAACTTCAAGAACCTGAGAGAGTTGTCTTTAAGAATCGACGGCTATTTAAGATCTATAG
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

V213G (1422)

1401 TAAGCATACGCCATCAACCTGGAGGATCTTCCCAGGGCTTTTCCAGCCCTGGAACCTTTGGTTGACTTGCCATTTGGTATCAATATCACAGATTT
200▶ K H T P I N L G R D L P Q G F S A L E P L V D L P I G I N I T R F
1501 CAGACCCTTCTGGCATTGCATCGGTCTTACTTCCAGGTGATTCCTCCTCCGGTGGACTGCCGGCGCCGCTGCCTACTATGTCGGTATCTGCAAC
234▶ 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 V G Y L Q
1601 CAAGAACCTTCTGCTCAAGTACAACGAAACGGCACTATTACGGATGCTGTTGATTGTCCTGGACCTCTGCTGACTAAATGACACCTCAAGAG
267▶ 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 K S
1701 CTTTACCCTTGAAGGGGATTTACCAAACAGTAATTTCCGGTCCAACCCACGAAAGCATTGTGCGGTTCCCAAATATACCAATCTGTGTCCCTTT
300▶ 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 F

G339D (1800)

S371F (1896)

1801 GATGAAGTGTTCATGCTACAAGTTTCTTCTGTGTACGCATGGAATAGGAAACGCATCTCCAATTGTGCTGCTGATTACTCCGTGCTGATAATTTTG
334▶ 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 N F

T376A (1911)

S375F (1908)

S373P (1902)

D405N (1998)

1901 CCCAATTCCTCGCTTCAAGTGTATGGCGTTTCCACTACCAAACCTAACGACCTGTGCTTCACTAATGTGTATGCCGACTCTTTTGTATACGAGGCA
367▶ 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 N

R408S (2007)

K417N (2034)

2001 TGAAGTGAAGCAGATTGCACCAGGGCAGACCGGCATTTGCCGACTACAACCTACAAGCTTCCAGATGACTTTACCGGATGTGTTATTGCATGGAACCTCA
400▶ 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 S

N440K (2103)

L452R (2139)

2101 AACAGCTGGATTCCAAGTGGTGGCAACTATAACTACAGATATAGACTGTTCCAGGAAATCCAACCTGAAACCATTGAGCGGAGATATAAGCACAGAAA
434▶ N K L D S K V G G N Y N Y R Y R L F R K S N L K P F E R D I S T E

S477N, T478K (2214)

E484A (2235)

F486V (2241)

Q498R (2277)

N501Y (2286)

Y505H (2298)

2201 TCTACCAGGCTGGAACAAACCTGCAACGGCGTGGCTGGGTTGAACTGCTACTTCCATTGCAGAGTTACGGATTACAGCTTACATACGAGGTTGGGTCA
467▶ I Y Q A G N K P C N G V A G V N C Y F P L Q S Y G F R P T Y G V G H

2301 CCAACCTATCGTGTGCTAGTCTGAGTCTTCTGAGTCTCCATGCCAGCCACAGTCTGTGGCCCAAGAAAAGCACCAATCTGGTGAAGAACAATGC
500▶ 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 C

2401 GTGAACCTTTAACCTTTAACGGACTCACAGGAACCGCGTATTGACGGAGAGTAAACAAGAAGTTCCTGCCATTCCAGCAGTTCGGTCGCGATATTGCCGACA
534▶ 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 D

2501 CTACCGACGCTGTCCGAGATCCCAGACATTGGAGATTCTTGATATCACACCCTGTAGTTTCGGCGGAGTGAGCGTGATTACGCCCGGAACCAATACCAG
567▶ 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 S

D614G (2625)

2601 CAATCAGGTTGCCGTCTGTATCAGGGCGTGAATTGCACCGAGGTACCTGTGCCATCCACGCTGACCAACTACACCACATGGCGAGTATATTCCACC
600▶ 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 T

H655Y (2748)

2701 GGCTCCAACGCTTTTCAGACACGTGCTGGATGTCTGATCGGTGCAGAAATATGTTAATAATAGCTACGAGTGTGATATCCCATCGGTGCTGGAATATGCG
634▶ 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 C

P681H (2826)

N679K (2820) Furin cleavage site

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

2901 CGCTACAGCAACAACCTCATTGCTATCCCTACTAACTTACAATCAGTGTGACAACCTGAAATCTGCCGTATCTATGACCAAAACAAGCGTTGACTGC
700▶ 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 C

N764K (3075)

3001 ACCATGTACATCTGTGGCGATTCTACCGAATGTAGCAATCTCCTCCTGCAATACGGATCATTCTGCACTCAGCTGAAGCGTGCCTCACAGGTATTGCAG
734▶ 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 A

D796Y (3171)

3101 TTGAGCAGGACAAGAATACGAGGAAGTGTTCGCCAGGTGAAGCAAATCTACAAAACCTCACCCATAAAAATCTTTGGCGATTCAATTTCTCACAGAT
767▶ 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 Q I

3201 CCTGCCGATCCCTCAAACCTCCAAGCGTAGCTTTATCGAGGATCTGCTCTTCAACAAGGTAACCTCGCAGATGCCGGTTTCATCAAGCAGTATGGC
800▶ 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 G

3301 GATTGTCTGGGAGACATCGCCGCTCGGGACCTGATCTGTGCACAGAAGTTCATGGACTGACCGTGCCTCCCTTCTGACCGACGAGATGATAGCCC
834▶ 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 A

3401 AATACACTAGCGCCCTGCTGGCCGGCACCATCACTTCTGGGTGGACATTCCGGAGTGGCGTGCCTTTCAGATTCTTTTGTATGCAGATGGCCTACCG
867▶ 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 R

3501 CTTTAAACGGCATCGGTGTGACACAAAACGTTCTGTATGAAAACAGAACTCATCGCAACAGTTCACAGTGTATCGGTAAGATACAGGATAGCCTG
900▶ 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 L

Q954H (3645) N969K (3690)

3601 TCATCCACTGCCAGCGCATTGGGAAAGTTCAGGATGTAGTGAACCAATGCCAGGCACCTAACACCCTGGTGAACAGCTCTCTTCAAGTTTGGTG
934▶ 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 G

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

3801 GACCTATGTGACCAGCAGCTCATCCGCGTGTGAAATTCGCGCATCCGCTAACCTGGCAGCAACCAAAATGTCCGAGTGTGTGCTGGGTGAGTCTAAG
1000▶ 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 S K

3901 AGAGTGGACTTTTGGGGAAGGGTATCACCTGATGTCTTTCTCAGTCTGCACCCCATGGTGTGGTCTTTCTGCACGTGACTTATGTCCCAGCTCAGG
1034▶ 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 Q

4001 AAAAGAATCTCACTACAGCCCCAGCCATCTGCCACGATGGGAAAGCCACTTTCCAGGGAAGGCGTATTCTGTCCAATGGTACTCATTGGTTCTGCAC
1067▶ 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 T

4101 TCAGAGAAATTTCTACGAGCCCCAGATTATAACCACTGACAATACATTTGTATCCGGCAATTTGTATGGTTATCGGGATTGTGAATAACTGTTTAC
1100▶ 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 Y

4201 GATCCTTTGCAGCCAGAGCTGGACTCCTTCAAGGAGGAGCTTGACAAAATTTTAAAGAAATCACACATCACCTGACGTGCACCTCGGAGATATTTAGGAA
1134▶ 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 G

4301 TCAATGCTCCGTGGTCAATATTGAGAAGGAGATAGACAGGCTGAATGAGGTTGCCAAGAACCTCAACGAGTCTCTGATCGATCTGCAGGAGTTGGGCAA
1167▶ 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 K

4401 GTACGAACAGTATATCAAATGGCCTTGGTACATTTGGCTTGGGTTTATTGCTGGGCTGATAGCTATCGTCATGGTGACAATATGTTGTGTTGCATGACA
1200▶ 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 T

NheI (4552)

4501 TCCTGCTGTAGTTGTCTGAAGGGCTGCTGCTCATGCGGAGCTGTTGCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAAC
1234▶ S C C S C L K G C C S C G S C C •

4601 ACAACTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATATAAGCTGCAATAAACAAGTTAAACA
ATA

4701 ACAAATTGCATTATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATAC
ATA

4801 AGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCAATGTGCATT
ATA

4901 AGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTCTTCAATTTCTTTATGTTTTAAATGCACTG
ATA

5001 ACTCCCACATTCCTTTTATGTAATAATTCAGAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTATTAGGCAGAAATCCAGATGCTCA
ATA

5101 AGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGT
ATA

5201 TCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAAGGATGATGAGAT
139▶ 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 I

5301 GAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTCTC
106▶ 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 K

5401 TGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAG
72 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 T
5501 TCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGA
39 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 S
5601 GATGTTGAAGTCTTCATGATGGCCCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCG
6 I N F T K M
5701 TCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTCGCTCAATGGGGCGGAGTTGTTA
5801 CGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATC
5901 CACGCCATTGATGTA CTG CCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTA CTG CCAAAGTAGGAAAGTCCCAT AAGGTCATGTA
6001 CTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGGAGTTTAC
6101 CGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGAAACATACGTCATTATTGACGTCAATGGGCGGGGGTCTTGGGGC
6201 GTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAG
BspLU11I (6256)
6301 GCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
6401 GATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGT
6501 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCGACCGC
6601 TCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
6701 ATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
6801 CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGA
6901 TCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTA
7001 AATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACG
7101 AAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAAGTGCCAGAACATTTCTCTATCGAA