



150

1 GGACCTGCAGGGCCTGAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTG
 101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCAGGCTCCCAGCAGGCAG
 201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTCTCCGGTGGCCGTGAGTGGGAGAGCGCACATCGCCACAGTCCCCGA
 301 GAAAGTGGGGGAGGGGTGCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACCTGGGAAAGTGTGCTGTACTGGCTCCGCTTTTCC
 401 GAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTC
 501 GCATCTCTCTTACGCGCCCGCCGCTACCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGC
 601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCCTCACGCTTTGC

Agel (791)

701 CTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCTGTTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCAA
 801 CATGTTTGTGTTCTTGGTGTGCTTCCACTGGTTCAGTTCCTCAATGCGTTAATCTCACCACCCGAACTCACTCCACCCGCATATACAAATTCCTTACC
 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
G75V+T76I (1023)
 1001 CAATCCATGTGTCTGGGACAAACCTGATCCTCGACAACCCCTGTATTGCCATTCAATGATGGGGTGTACTTTGCCTCCACAGAGAAATCCAACAT
 67▶ A I H V S G T N V I K R F D N P V L P F N D G V Y F A S T E K S N I
 1101 CATTTCGAGGATGATTTTCGGGACTACTCTGGACTCAAAGACACAGAGCCTGTGATCGTTAACAACGCCACAAACGTTGTATCAAGTGTGCGGAGTTTC
 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
 1201 CAGTTTTGCAATGATCCCTTCTGGGAGTGTACTATACAAGAATAACAAGTCTGGATGGAGAGCGAATTTCCGGGTCTACAGCAGCGCAAACAAGTCA
 134▶ 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 V Y S S A N N C
 1301 CCTTCGAGTACGTGAGTCAACCTTTCTGATGGACCTGGAAGGAAACAGGAAACTTCAAGAACCTGAGAGAGTTTGTCTTTAAGAACATCGACGGCTA
 167▶ 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 Y
 1401 TTTAAGATCTATAGTAAGCATACGCCTATCAACCTGGTAAGGGATCTTCCCAGGGCTTTTCCAGCCTGGAACCTTTGGTTGACTTGCTATTGGTATC
 200▶ 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 I

D253N (1539)

Δ [R246,G252]

1501 AATATACCAGATTTAGACCCTTCTGGCATTGCATAACTCCTCCTCCGGGTGGACTGCCGGCGCCGCTGCCTACTATGTCGGCTATCTGCAACCAAGAA
 234▶ N I T R F Q T L L A L H N S S S G W T A G A A A Y Y V G Y L Q P R
 1601 CGTTCTGCTCAAGTACAACGAAACGGCACTATTACGGATGCTGTTGATTGTGCCCTGGACCCTCTGTCTGAGACTAAATGCACCCTCAAGAGCTTTAC
 267▶ 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 F T
 1701 CGTTGAGAAGGGGATTTACCAAACAGTAATTTCCGGTCCAACCCACCGAAAGCATTGTGCGGTTCCCAAATATACCAATCTGTGTCCTTTGGCGAA
 300▶ 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 G E

1801 GTGTTCAATGCTACAAGTTTGTCTTCTGTGTACGCATGGAATAGGAAACGCATCTCCAATTGTGTCGCTGATTACTCCGTGCTGTACAATTCGCCTCTT
 334▶ 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 S A S

1901 TCTCAACCTTCAAGTGTATGGCGTTTACCTACCAAACCTAACGACCTGTGCTTCACTAATGTGTATGCCGACTCTTTTGTGATACGAGGCGATGAAGT
 367▶ 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 F V I R G D E V

2001 GAGACAGATTGCACCAGGCGAGACCGGCAAAATGCGGACTACAACCTACAAGCTTCCAGATGACTTTACCGGATGTGTTATTGATGGAACCTCAAACAAT
 400▶ 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 V I A W N S N N

L452Q (2133)

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

F490S (2247)

2201 AGGCTGGAAGTACGCCCTGCAACGGCGTGAAGGGTCAACTGCTACTCCCATTCGAGAGTTACGGATTCCAGCCTACAAACGGGGTGGGATACCAACC
 467▶ Q A G S T P C N G V E G F N C Y S P L Q S Y G F Q P T N G V G Y Q P

2301 CTATCGTGTGCTAGTCTCTGAGTTTTGAGTCTCCTCATGCCAGCCACAGTCTGTGGCCCCAAGAAAAGCACCAATCTGGTGAAGAACAATGCGTGAAC
 500▶ 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 V N

2401 TTTAACTTTAACGGACTCACAGGAACCGCGTATTGACGGAGAGTAACAAGAAGTTCTGCCATTCCAGCAGTTCGGTTCGCGATATTGCCGACACTACCG
 534▶ 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 T T

2501 ACGCTGTCCGAGATCCCAGACATTGGAGATCTTGATATCACACCCTGTAGTTTCCGGCGAGTGAAGCTGATTACGCCCGGAACCAATACCAGCAATCA
 567▶ 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 N Q

D614G (2619)

2601 GGTTGCCGTCCTGTATCAGGGCGTGAATTGCACCGAGGTACCTGTGCGCCATCCACGCTGACCAAATTACCCACATGGCGAGTATATTCCACCGGCTCC
600 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 G S
2701 AACGTCTTTCAGACACGTGCTGGATGTCTGATCGGTGCAGAACAGTTAATAATAGCTACGAGTGTGATATCCCCATCGGTGCTGGAATATGCGCCTCTT
634 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 I G A G I C A S

Furin cleavage site (2832)

2801 ATCAAACCTCAAACCAACTCTCCTAGCGGGCAACGTAGTGTAGCATCCCAAAGTATCATTGCCTACACAATGAGCCTCGGTGCTGAGAATTCTGTCGCCTA
667 Y Q T Q T N S P R R A R S V A S Q S I I A Y T M S L G A E N S V A Y
2901 CAGCAACAACCTCCATTGTATCCCTACTAACTTCACAATCAGTGTGACAACTGAAATCTGCCGATCTATGACCAAAAACAGCGTTGACTGCACCATG
700 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 T M
3001 TACATCTGTGGCATTCTACCGAATGTAGCAATCTCCTCTGCAATACGGATCATTCTGCACCTCAGTGAATCGCCCTCACAGGTATTGCGATTGAGC
734 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 A L T G I A V E
3101 AGGACAAGAATACGCAGGAAGTGTGGCCAGGTGAAGCAAATCTACAAAACCTCACCCATAAAAAGACTTTGGCGGATTCAATTTCTCACAGATCCTGCC
767 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 F N F S Q I L P
3201 CGATCCCTCAAAACCTCAAGCGTAGCTTTATCGAGGATCTGCTCTTCAACAAGGTAACCCTCGCAGATGCCGGTTTCATCAAGCAGTATGGCGATTGT
800 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 D C

T859N (3354)

3301 CTGGGAGACATCGCCGCTCGGGACCTGATCTGTGCACAGAAGTTCAATGGACTGAACGTGCTGCCCTTCTGCTGACCGAGAGATGATAGCCCAATACA
834 L G D I A A R D L I C A Q K F N G L N V L P P L L T D E M I A Q Y
3401 CTAGCGCCCTGCTGGCCGGCACCATCACTTCTGGTGGACATTCGGAGCTGGCGCTGCCCTTCCAGATTCTTTTGTATGCAGATGGCTACCCTTTAA
867 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 F N
3501 CGGCATCGGTGTGACACAAAACGTTCTGTATGAAAACAGAACTCATCGCAACAGTTCAACAGTCTATCGGTAAGATACAGGATAGCCTGTCATCC
900 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 S S
3601 ACTGCCAGCGCATTGGGAAAGTTGCAGGATGTAGTGAACCAGAAAGTCCAGGCACTTAACACCTGGTGAACAGCTCTTCAAATTTTGGTGCCATTT
934 T A S A L G K L Q D V V N Q N A Q A L N T L V K Q L S S N F G A I
3701 CTAGSCTGAATGACATACTGAGCCGTTGGACAAGGTGGAGGCTGAAAGTCAGATTGATAGGCTGATAACTGGGCGCCTTCAAGTCTTTCAGACCTA
967 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 Q T Y
3801 TGTGACCAGCAGCTCATCCGCGCTGCTGAAATTCGCGCATCCGCTAACCTGGCAGCAACCAAAATGTCGAGTGTGTGCTGGGTGAGTCTAAGAGAGTG
1000 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 R V
3901 GACTTTTTCGGGAAGGGGTATCACTGATGCTTTTCTCAGTCTGCACCCATGGTGTGGTCTTTCTGCACGTGACTTATGTCCAGCTCAGGAAAAGA
1034 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 E K
4001 ACTTCACTACAGCCCACTGCCACGATGGGAAAGCCACTTTCCAGGGAAGGCGTATTCGTGCAATGGTACTCATTGGTTCGTCACTCAGAG
1067 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 Q R
4101 AAATTTCTACGAGCCCCAGATTATAACCACTGACAATACATTTGTATCCGGCAATTGTGATGTGGTTATCGGGATTGTGAATAATACTGTTTACGATCCT
1100 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 D P
4201 TTGAGCCAGAGCTGGACTCCTCAAGGAGGAGCTTGACAAATATTTAAGAATCACACATCACTGACGTGACCTCGGAGATATTCAGGAATCAATG
1134 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 I N
4301 CTTCCGTGGTCAATATTCAGAAGGAGATAGACAGGCTGAATGAGGTTGCCAAGAACCTCAACGAGTCTCTGATCGATCTGCAGGAGTTGGGCAAGTACGA
1167 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 Y E
4401 ACAGTATATCAAATGGCTGGTACATTTGGCTGGGTTTCACTGGGCTGATAGCTATCGTCATGGTGACAATTATGTTGTGTTGCATGACATCCTGC
1200 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 S C

NheI (4546)

4501 TGTAGTTGCTGAAGGGCTGCTGCTCATGCGGCAGCTGTTGCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAACCACAAT
1234 C S C L K G C C S C G S C C
4601 AGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAT
4701 GCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCAAAATACAGCATA
4801 GCAAAAACCTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGT
4901 TTGAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTAAATGACTGACCTCC
5001 CACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCC
5101 TTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAATTTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGG
5201 TGTACTTGAGGGGATGAGTTCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTC
137 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 L E
5301 TCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCG
104 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 Q G
5401 TTGCTCACAGCAGACCAATGGCAATGGCTTACGACAGACAGTACCTGCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGG
70 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 K T
5501 TCCTGATGGCCGCCCCGACATGGTGTGTTGTCTCATAGAGCAGTGGTGTCTCAGTGGCGACTCCACCAGCTCCAGATCCTGCTGAGAGATGTT
37 R I A A G V H G K N D E Y L M T I K E T A V E V L E L D Q Q S I N
5601 GAAGTCTTCATGATGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCA
4 F T K M
5701 GCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACAT

5801 TTTGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCC
5901 CATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATTAAGGTCATGTACTGGGC
6001 ATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAA
6101 TACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGC
6201 CAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCG
6301 TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC
6401 AGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCT
6501 TTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCCGACCGCTGCGCC
6601 TTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAG
6701 GCGGTGCTACAGAGTTCTGAAGTGGTGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA
6801 AAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAA
6901 GAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAG
7001 CGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAA
7101 AACAACTAGCAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA