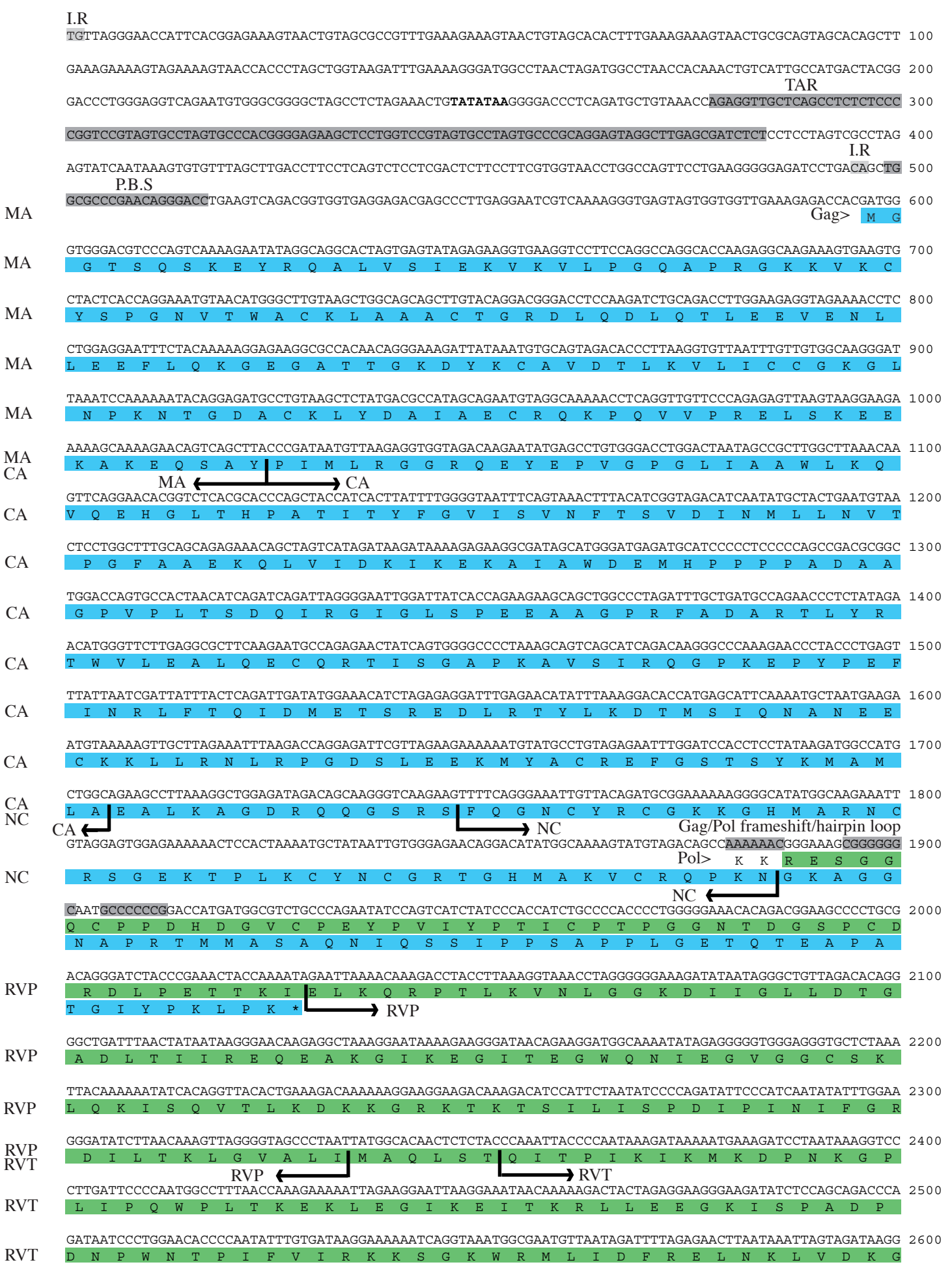
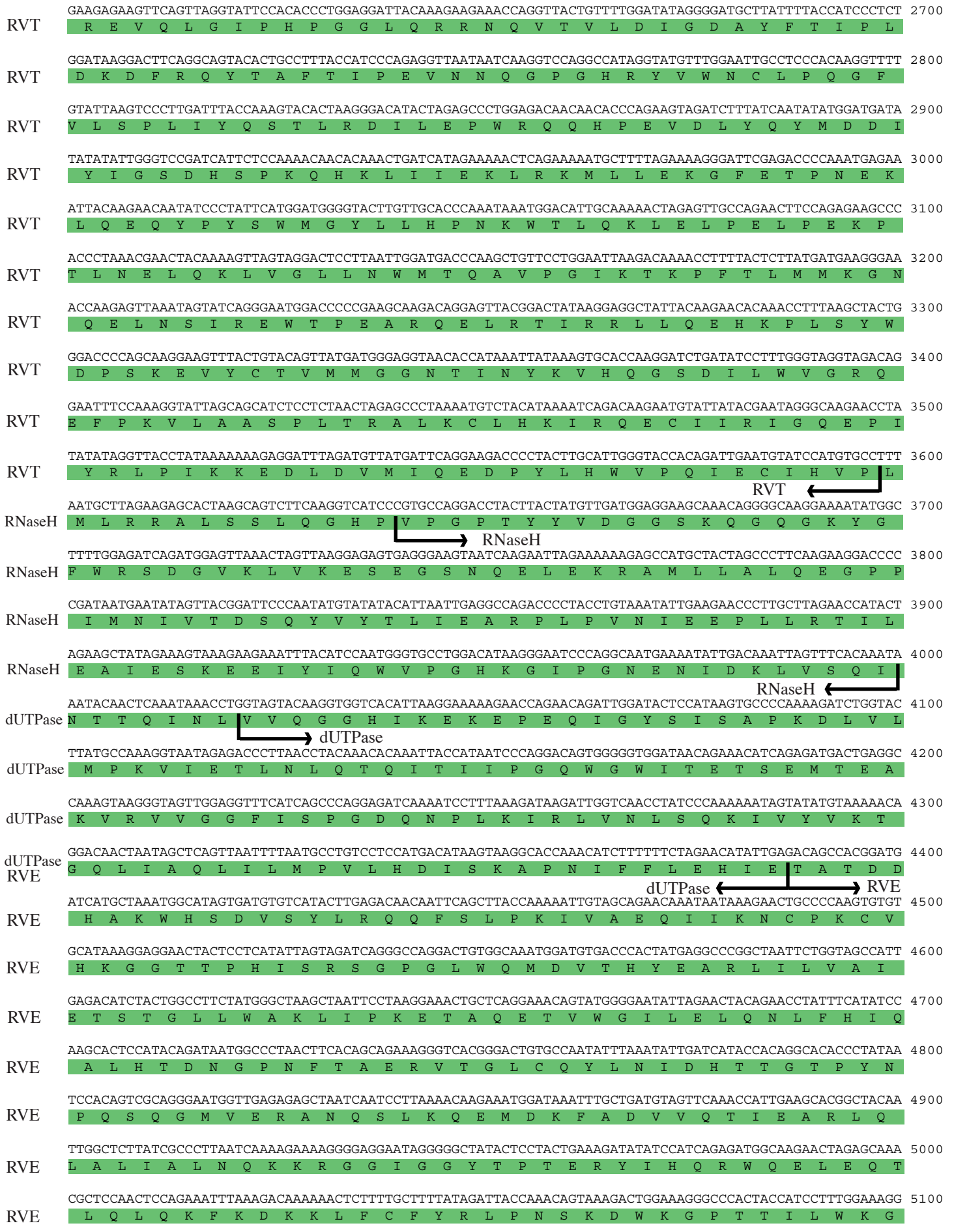
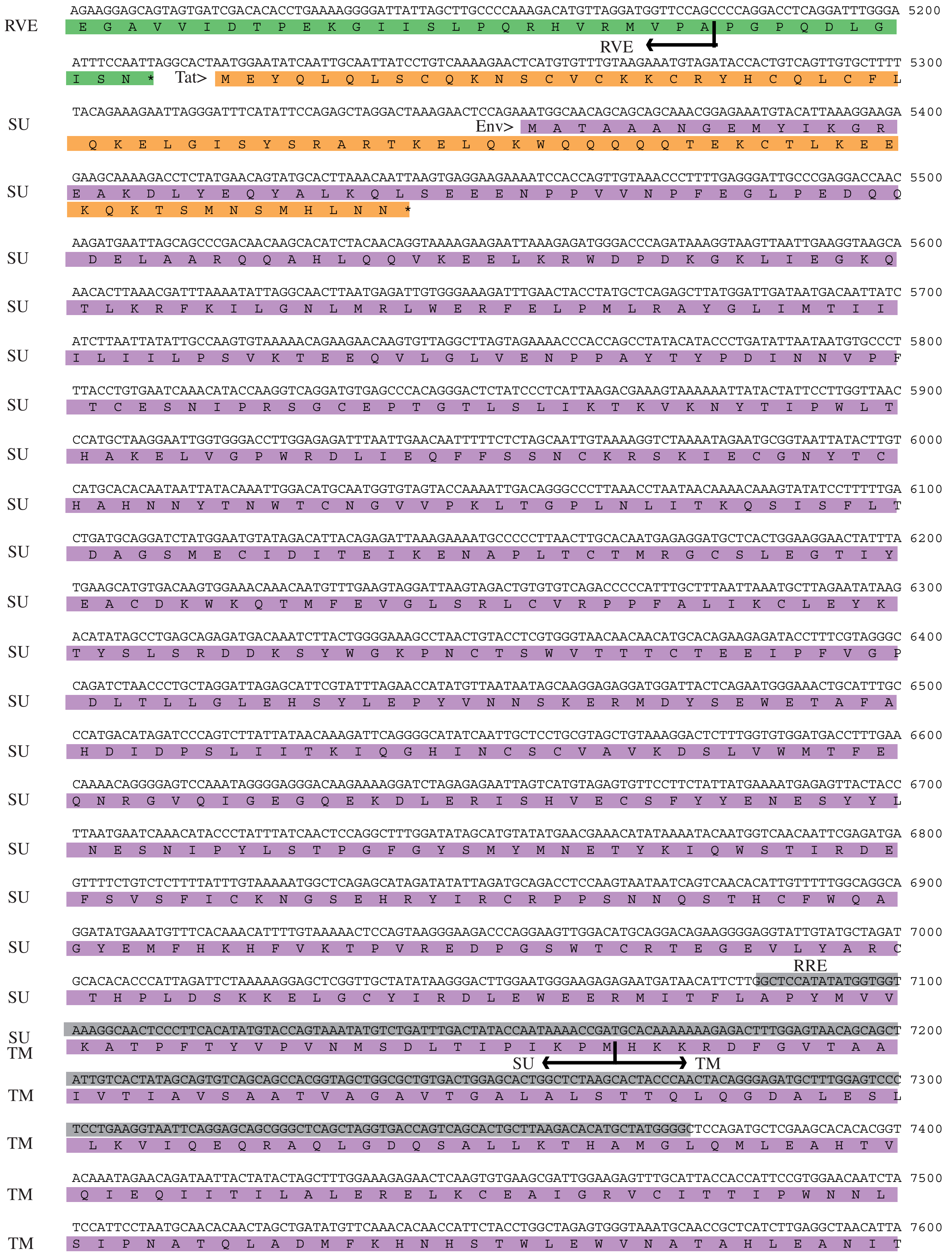
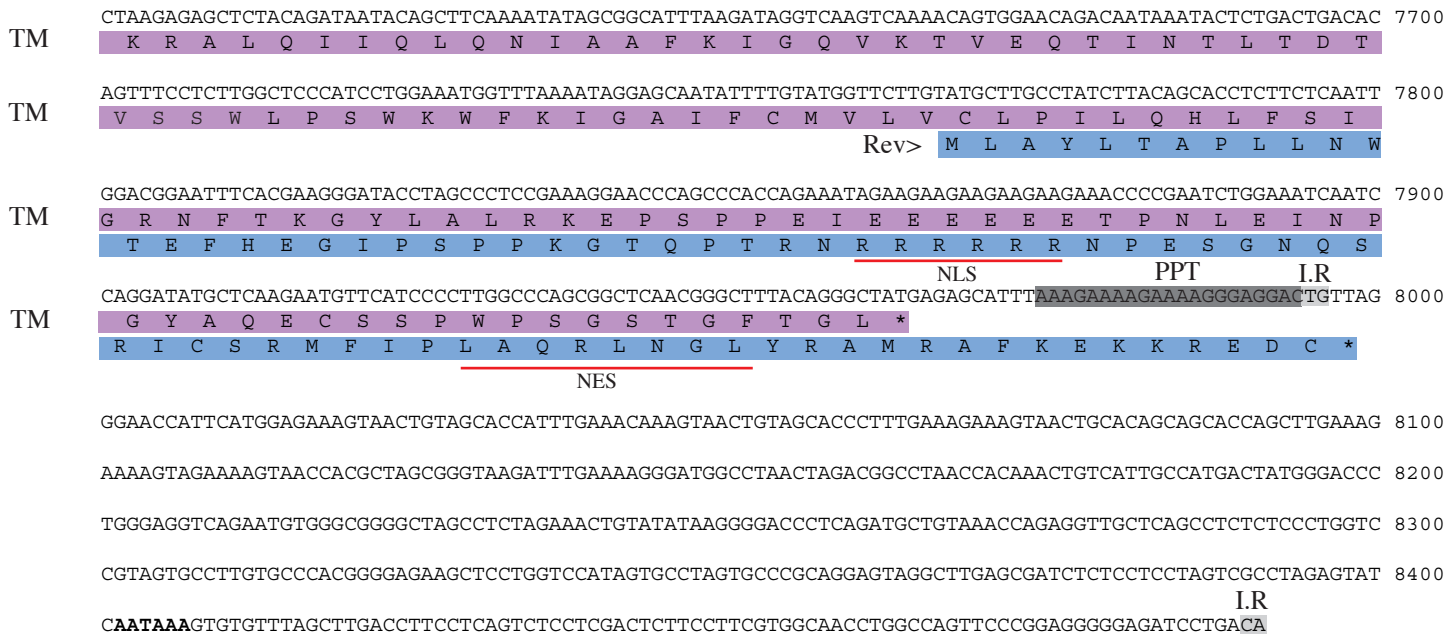

I.R
TGTTAGGGAACCATTCACGGAGAAAGTAACTGTAGCGCCGTTTGAAAGAAAGTAACTGTAGCACACTTTGAAAGAAAGTAACTGCGCAGTAGCACAGCTT 100
GAAAGAAAAGTAGAAAAGTAACCACCCTAGCTGGTAAGATTTGAAAAGGGATGGCCTAACTAGATGGCCTAACCACAAACTGTCATTGCCATGACTACGG 200
TAR
GACCCTGGGAGGTCAGAATGTGGGCGGGGCTAGCCTCTAGAAACTG**TATATAA**GGGGACCCTCAGATGCTGTAAACCAGAGGTTGCTCAGCCTCTCTCCC 300
CGGTCCGTAGTGCCTAGTGCCCACGGGGAGAAGCTCCTGGTCCGTAGTGCCTAGTGCCCGCAGGAGTAGGCTTGAGCGATCTCTCCTCCTAGTCGCCTAG 400
I.R
AGTATCAATAAAGTGTGTTTAGCTTGACCTTCCTCAGTCTCCTCGACTCTTCCTTCGTGGTAACCTGGCCAGTTCCTGAAGGGGGAGATCCTGACAGCTG 500
P.B.S
GCGCCCGAACAGGGACCTGAAGTCAGACGGTGGTGAGGAGACGAGCCCTTGAGGAATCGTCAAAAGGGTGAGTAGTGGTGGTTGAAAGAGACCACGATGG 600
MA
Gag>
M G
GTGGGACGTCCCAGTCAAAAGAATATAGGCAGGCACTAGTGAGTATAGAGAAGGTGAAGGTCCTTCCAGGCCAGGCACCAAGAGGCAAGAAAGTGAAGTG 700
MA
G T S Q S K E Y R Q A L V S I E K V K V L P G Q A P R G K K V K C
CTACTCACCAGGAAATGTAACATGGGCTTGTAAGCTGGCAGCAGCTTGTACAGGACGGGACCTCCAAGATCTGCAGACCTTGGAAGAGGTAGAAAACCTC 800
MA
Y S P G N V T W A C K L A A A C T G R D L Q D L Q T L E E V E N L
CTGGAGGAATTTCTACAAAAAGGAGAAGGCGCCACAACAGGGAAAGATTATAAATGTGCAGTAGACACCCTTAAGGTGTTAATTTGTTGTGGCAAGGGAT 900
MA
L E E F L Q K G E G A T T G K D Y K C A V D T L K V L I C C G K G L
TAAATCCAAAAAATACAGGAGATGCCTGTAAGCTCTATGACGCCATAGCAGAATGTAGGCAAAAACCTCAGGTTGTTCCCAGAGAGTTAAGTAAGGAAGA 1000
MA
N P K N T G D A C K L Y D A I A E C R Q K P Q V V P R E L S K E E
AAAAGCAAAAGAACAGTCAGCTTACCCGATAATGTTAAGAGGTGGTAGACAAGAATATGAGCCTGTGGGACCTGGACTAATAGCCGCTTGGCTTAAACAA 1100
MA
CA
K A K E Q S A Y P I M L R G G R Q E Y E P V G P G L I A A W L K Q
MA
CA
GTTCAGGAACACGGTCTCACGCACCCAGCTACCATCACTTATTTTGGGGTAATTTCAGTAAACTTTACATCGGTAGACATCAATATGCTACTGAATGTAA 1200
CA
V Q E H G L T H P A T I T Y F G V I S V N F T S V D I N M L L N V T
CTCCTGGCTTTGCAGCAGAGAAACAGCTAGTCATAGATAAGATAAAAGAGAAGGCGATAGCATGGGATGAGATGCATCCCCCTCCCCCAGCCGACGCGGC 1300
CA
P G F A A E K Q L V I D K I K E K A I A W D E M H P P P P A D A A
TGGACCAGTGCCACTAACATCAGATCAGATTAGGGGAATTGGATTATCACCAGAAGAAGCAGCTGGCCCTAGATTTGCTGATGCCAGAACCCTCTATAGA 1400
CA
G P V P L T S D Q I R G I G L S P E E A A G P R F A D A R T L Y R
ACATGGGTTCTTGAGGCGCTTCAAGAATGCCAGAGAACTATCAGTGGGGCCCCTAAAGCAGTCAGCATCAGACAAGGGCCCAAAGAACCCTACCCTGAGT 1500
CA
T W V L E A L Q E C Q R T I S G A P K A V S I R Q G P K E P Y P E F
TTATTAATCGATTATTTACTCAGATTGATATGGAAACATCTAGAGAGGATTTGAGAACATATTTAAAGGACACCATGAGCATTCAAAATGCTAATGAAGA 1600
CA
I N R L F T Q I D M E T S R E D L R T Y L K D T M S I Q N A N E E
ATGTAAAAAGTTGCTTAGAAATTTAAGACCAGGAGATTCGTTAGAAGAAAAAATGTATGCCTGTAGAGAATTTGGATCCACCTCCTATAAGATGGCCATG 1700
CA
C K K L L R N L R P G D S L E E K M Y A C R E F G S T S Y K M A M
CTGGCAGAAGCCTTAAAGGCTGGAGATAGACAGCAAGGGTCAAGAAGTTTTCAGGGAAATTGTTACAGATGCGGAAAAAAGGGGCATATGGCAAGAAATT 1800
CA
NC
L A E A L K A G D R Q Q G S R S F Q G N C Y R C G K K G H M A R N C
CA
NC
Gag/Pol frameshift/hairpin loop
GTAGGAGTGGAGAAAAAACTCCACTAAAATGCTATAATTGTGGGAGAACAGGACATATGGCAAAAGTATGTAGACAGCCAAAAAACGGGAAAGCGGGGGG 1900
Pol>
K K
NC
R S G E K T P L K C Y N C G R T G H M A K V C R Q P K N G K A G G
NC
CAATGCCCCCCGGACCATGATGGCGTCTGCCCAGAATATCCAGTCATCTATCCCACCATCTGCCCCACCCCTGGGGGAAACACAGACGGAAGCCCCTGCG 2000
Q C P P D H D G V C P E Y P V I Y P T I C P T P G G N T D G S P C D
N A P R T M M A S A Q N I Q S S I P P S A P P L G E T Q T E A P A
ACAGGGATCTACCCGAAACTACCAAAATAGAATTAAAACAAAGACCTACCTTAAAGGTAAACCTAGGGGGGAAAGATATAATAGGGCTGTTAGACACAGG 2100
RVP
R D L P E T T K I E L K Q R P T L K V N L G G K D I I G L L D T G
T G I Y P K L P K *
RVP
GGCTGATTTAACTATAATAAGGGAACAAGAGGCTAAAGGAATAAAAGAAGGGATAACAGAAGGATGGCAAAATATAGAGGGGGTGGGAGGGTGCTCTAAA 2200
RVP
A D L T I I R E Q E A K G I K E G I T E G W Q N I E G V G G C S K
TTACAAAAAATATCACAGGTTACACTGAAAGACAAAAAAGGAAGGAAGACAAAGACATCCATTCTAATATCCCCAGATATTCCCATCAATATATTTGGAA 2300
RVP
L Q K I S Q V T L K D K K G R K T K T S I L I S P D I P I N I F G R
GGGATATCTTAACAAAGTTAGGGGTAGCCCTAATTATGGCACAACTCTCTACCCAAATTACCCCAATAAAGATAAAAATGAAAGATCCTAATAAAGGTCC 2400
RVP
RVT
D I L T K L G V A L I M A Q L S T Q I T P I K I K M K D P N K G P
RVP
RVT
CTTGATTCCCCAATGGCCTTTAACCAAAGAAAAATTAGAAGGAATTAAGGAAATAACAAAAAGACTACTAGAGGAAGGGAAGATATCTCCAGCAGACCCA 2500
RVT
L I P Q W P L T K E K L E G I K E I T K R L L E E G K I S P A D P
GATAATCCCTGGAACACCCCAATATTTGTGATAAGGAAAAAATCAGGTAAATGGCGAATGTTAATAGATTTTAGAGAACTTAATAAATTAGTAGATAAGG 2600
RVT
D N P W N T P I F V I R K K S G K W R M L I D F R E L N K L V D K G


GAAGAGAAGTTCAGTTAGGTATTCCACACCCTGGAGGATTACAAAGAAGAAACCAGGTTACTGTTTTGGATATAGGGGATGCTTATTTTACCATCCCTCT 2700
RVT R E V Q L G I P H P G G L Q R R N Q V T V L D I G D A Y F T I P L
GGATAAGGACTTCAGGCAGTACACTGCCTTTACCATCCCAGAGGTTAATAATCAAGGTCCAGGCCATAGGTATGTTTGGAATTGCCTCCCACAAGGTTTT 2800
RVT D K D F R Q Y T A F T I P E V N N Q G P G H R Y V W N C L P Q G F
GTATTAAGTCCCTTGATTTACCAAAGTACACTAAGGGACATACTAGAGCCCTGGAGACAACAACACCCAGAAGTAGATCTTTATCAATATATGGATGATA 2900
RVT V L S P L I Y Q S T L R D I L E P W R Q Q H P E V D L Y Q Y M D D I
TATATATTGGGTCCGATCATTCTCCAAAACAACACAAACTGATCATAGAAAAACTCAGAAAAATGCTTTTAGAAAAGGGATTCGAGACCCCAAATGAGAA 3000
RVT Y I G S D H S P K Q H K L I I E K L R K M L L E K G F E T P N E K
ATTACAAGAACAATATCCCTATTCATGGATGGGGTACTTGTTGCACCCAAATAAATGGACATTGCAAAAACTAGAGTTGCCAGAACTTCCAGAGAAGCCC 3100
RVT L Q E Q Y P Y S W M G Y L L H P N K W T L Q K L E L P E L P E K P
ACCCTAAACGAACTACAAAAGTTAGTAGGACTCCTTAATTGGATGACCCAAGCTGTTCCTGGAATTAAGACAAAACCTTTTACTCTTATGATGAAGGGAA 3200
RVT T L N E L Q K L V G L L N W M T Q A V P G I K T K P F T L M M K G N
ACCAAGAGTTAAATAGTATCAGGGAATGGACCCCCGAAGCAAGACAGGAGTTACGGACTATAAGGAGGCTATTACAAGAACACAAACCTTTAAGCTACTG 3300
RVT Q E L N S I R E W T P E A R Q E L R T I R R L L Q E H K P L S Y W
GGACCCCAGCAAGGAAGTTTACTGTACAGTTATGATGGGAGGTAACACCATAAATTATAAAGTGCACCAAGGATCTGATATCCTTTGGGTAGGTAGACAG 3400
RVT D P S K E V Y C T V M M G G N T I N Y K V H Q G S D I L W V G R Q
GAATTTCCAAAGGTATTAGCAGCATCTCCTCTAACTAGAGCCCTAAAATGTCTACATAAAATCAGACAAGAATGTATTATACGAATAGGGCAAGAACCTA 3500
RVT E F P K V L A A S P L T R A L K C L H K I R Q E C I I R I G Q E P I
TATATAGGTTACCTATAAAAAAAGAGGATTTAGATGTTATGATTCAGGAAGACCCCTACTTGCATTGGGTACCACAGATTGAATGTATCCATGTGCCTTT 3600
RVT Y R L P I K K E D L D V M I Q E D P Y L H W V P Q I E C I H V P L
RVT
AATGCTTAGAAGAGCACTAAGCAGTCTTCAAGGTCATCCCGTGCCAGGACCTACTTACTATGTTGATGGAGGAAGCAAACAGGGGCAAGGAAAATATGGC 3700
RNaseH M L R R A L S S S L Q G H P V P G P T Y Y V D G G S K Q G Q G K Y G
RNaseH
TTTTGGAGATCAGATGGAGTTAAACTAGTTAAGGAGAGTGAGGGAAGTAATCAAGAATTAGAAAAAAGAGCCATGCTACTAGCCCTTCAAGAAGGACCCC 3800
RNaseH F W R S D G V K L V K E S E G S N Q E L E K R A M L L A L Q E G P P
CGATAATGAATATAGTTACGGATTCCCAATATGTATATACATTAATTGAGGCCAGACCCCTACCTGTAAATATTGAAGAACCCTTGCTTAGAACCATACT 3900
RNaseH I M N I V T D S Q Y V Y T L I E A R P L P V N I E E P L L R T I L
AGAAGCTATAGAAAGTAAAGAAGAAATTTACATCCAATGGGTGCCTGGACATAAGGGAATCCCAGGCAATGAAAATATTGACAAATTAGTTTCACAAATA 4000
RNaseH E A I E S K E E I Y I Q W V P G H K G I P G N E N I D K L V S Q I
RNaseH
AATACAACTCAAATAAACCTGGTAGTACAAGGTGGTCACATTAAGGAAAAAGAACCAGAACAGATTGGATACTCCATAAGTGCCCCAAAAGATCTGGTAC 4100
dUTPase N T T Q I N L V V Q G G H I K E K E P E Q I G Y S I S A P K D L V L
dUTPase
TTATGCCAAAGGTAATAGAGACCCTTAACCTACAAACACAAATTACCATAATCCCAGGACAGTGGGGGTGGATAACAGAAACATCAGAGATGACTGAGGC 4200
dUTPase M P K V I E T L N L Q T Q I T I I P G Q W G W I T E T S E M T E A
CAAAGTAAGGGTAGTTGGAGGTTTCATCAGCCCAGGAGATCAAAATCCTTTAAAGATAAGATTGGTCAACCTATCCCAAAAAATAGTATATGTAAAAACA 4300
dUTPase K V R V V G G F I S P G D Q N P L K I R L V N L S Q K I V Y V K T
GGACAACTAATAGCTCAGTTAATTTTAATGCCTGTCCTCCATGACATAAGTAAGGCACCAAACATCTTTTTTCTAGAACATATTGAGACAGCCACGGATG 4400
dUTPase G Q L I A Q L I L M P V L H D I S K A P N I F F L E H I E T A T D D
RVE
dUTPase
RVE
ATCATGCTAAATGGCATAGTGATGTGTCATACTTGAGACAACAATTCAGCTTACCAAAAATTGTAGCAGAACAAATAATAAAGAACTGCCCCAAGTGTGT 4500
RVE H A K W H S D V S Y L R Q Q F S L P K I V A E Q I I K N C P K C V
GCATAAAGGAGGAACTACTCCTCATATTAGTAGATCAGGGCCAGGACTGTGGCAAATGGATGTGACCCACTATGAGGCCCGGCTAATTCTGGTAGCCATT 4600
RVE H K G G T T P H I S R S G P G L W Q M D V T H Y E A R L I L V A I
GAGACATCTACTGGCCTTCTATGGGCTAAGCTAATTCCTAAGGAAACTGCTCAGGAAACAGTATGGGGAATATTAGAACTACAGAACCTATTTCATATCC 4700
RVE E T S T G L L W A K L I P K E T A Q E T V W G I L E L Q N L F H I Q
AAGCACTCCATACAGATAATGGCCCTAACTTCACAGCAGAAAGGGTCACGGGACTGTGCCAATATTTAAATATTGATCATACCACAGGCACACCCTATAA 4800
RVE A L H T D N G P N F T A E R V T G L C Q Y L N I D H T T G T P Y N
TCCACAGTCGCAGGGAATGGTTGAGAGAGCTAATCAATCCTTAAAACAAGAAATGGATAAATTTGCTGATGTAGTTCAAACCATTGAAGCACGGCTACAA 4900
RVE P Q S Q G M V E R A N Q S L K Q E M D K F A D V V Q T I E A R L Q
TTGGCTCTTATCGCCCTTAATCAAAAGAAAAGGGGAGGAATAGGGGGCTATACTCCTACTGAAAGATATATCCATCAGAGATGGCAAGAACTAGAGCAAA 5000
RVE L A L I A L N Q K K R G G I G G Y T P T E R Y I H Q R W Q E L E Q T
CGCTCCAACTCCAGAAATTTAAAGACAAAAAACTCTTTTGCTTTTATAGATTACCAAACAGTAAAGACTGGAAAGGGCCCACTACCATCCTTTGGAAAGG 5100
RVE L Q L Q K F K D K K L F C F Y R L P N S K D W K G P T T I L W K G



AGAAGGAGCAGTAGTGATCGACACACCTGAAAAGGGGATTATTAGCTTGCCCCAAAGACATGTTAGGATGGTTCCAGCCCCAGGACCTCAGGATTTGGGA 5200
RVE E G A V V I D T P E K G I I S L P Q R H V R M V P A P G P Q D L G
RVE
ATTTCCAATTAGGCACTAATGGAATATCAATTGCAATTATCCTGTCAAAAGAACTCATGTGTTTGTAAGAAATGTAGATACCACTGTCAGTTGTGCTTTT 5300
I S N * Tat> M E Y Q L Q L S C Q K N S C V C K K C R Y H C Q L C F L
TACAGAAAGAATTAGGGATTTCATATTCCAGAGCTAGGACTAAAGAACTCCAGAAATGGCAACAGCAGCAGCAAACGGAGAAATGTACATTAAAGGAAGA 5400
SU Env> M A T A A A N G E M Y I K G R
Q K E L G I S Y S R A R T K E L Q K W Q Q Q Q Q T E K C T L K E E
GAAGCAAAAGACCTCTATGAACAGTATGCACTTAAACAATTAAGTGAGGAAGAAAATCCACCAGTTGTAAACCCTTTTGAGGGATTGCCCGAGGACCAAC 5500
SU E A K D L Y E Q Y A L K Q L S E E E N P P V V N P F E G L P E D Q Q
K Q K T S M N S M H L N N *
AAGATGAATTAGCAGCCCGACAACAAGCACATCTACAACAGGTAAAAGAAGAATTAAAGAGATGGGACCCAGATAAAGGTAAGTTAATTGAAGGTAAGCA 5600
SU D E L A A R Q Q A H L Q Q V K E E L K R W D P D K G K L I E G K Q
AACACTTAAACGATTTAAAATATTAGGCAACTTAATGAGATTGTGGGAAAGATTTGAACTACCTATGCTCAGAGCTTATGGATTGATAATGACAATTATC 5700
SU T L K R F K I L G N L M R L W E R F E L P M L R A Y G L I M T I I
ATCTTAATTATATTGCCAAGTGTAAAAACAGAAGAACAAGTGTTAGGCTTAGTAGAAAACCCACCAGCCTATACATACCCTGATATTAATAATGTGCCCT 5800
SU I L I I L P S V K T E E Q V L G L V E N P P A Y T Y P D I N N V P F
TTACCTGTGAATCAAACATACCAAGGTCAGGATGTGAGCCCACAGGGACTCTATCCCTCATTAAGACGAAAGTAAAAAATTATACTATTCCTTGGTTAAC 5900
SU T C E S N I P R S G C E P T G T L S L I K T K V K N Y T I P W L T
CCATGCTAAGGAATTGGTGGGACCTTGGAGAGATTTAATTGAACAATTTTTCTCTAGCAATTGTAAAAGGTCTAAAATAGAATGCGGTAATTATACTTGT 6000
SU H A K E L V G P W R D L I E Q F F S S N C K R S K I E C G N Y T C
CATGCACACAATAATTATACAAATTGGACATGCAATGGTGTAGTACCAAAATTGACAGGGCCCTTAAACCTAATAACAAAACAAAGTATATCCTTTTTGA 6100
SU H A H N N Y T N W T C N G V V P K L T G P L N L I T K Q S I S F L T
CTGATGCAGGATCTATGGAATGTATAGACATTACAGAGATTAAAGAAAATGCCCCCTTAACTTGCACAATGAGAGGATGCTCACTGGAAGGAACTATTTA 6200
SU D A G S M E C I D I T E I K E N A P L T C T M R G C S L E G T I Y
TGAAGCATGTGACAAGTGGAAACAAACAATGTTTGAAGTAGGATTAAGTAGACTGTGTGTCAGACCCCCATTTGCTTTAATTAAATGCTTAGAATATAAG 6300
SU E A C D K W K Q T M F E V G L S R L C V R P P F A L I K C L E Y K
ACATATAGCCTGAGCAGAGATGACAAATCTTACTGGGGAAAGCCTAACTGTACCTCGTGGGTAACAACAACATGCACAGAAGAGATACCTTTCGTAGGGC 6400
SU T Y S L S R D D K S Y W G K P N C T S W V T T T C T E E I P F V G P
CAGATCTAACCCTGCTAGGATTAGAGCATTCGTATTTAGAACCATATGTTAATAATAGCAAGGAGAGGATGGATTACTCAGAATGGGAAACTGCATTTGC 6500
SU D L T L L G L E H S Y L E P Y V N N S K E R M D Y S E W E T A F A
CCATGACATAGATCCCAGTCTTATTATAACAAAGATTCAGGGGCATATCAATTGCTCCTGCGTAGCTGTAAAGGACTCTTTGGTGTGGATGACCTTTGAA 6600
SU H D I D P S L I I T K I Q G H I N C S C V A V K D S L V W M T F E
CAAAACAGGGGAGTCCAAATAGGGGAGGGACAAGAAAAGGATCTAGAGAGAATTAGTCATGTAGAGTGTTCCTTCTATTATGAAAATGAGAGTTACTACC 6700
SU Q N R G V Q I G E G Q E K D L E R I S H V E C S F Y Y E N E S Y Y L
TTAATGAATCAAACATACCCTATTTATCAACTCCAGGCTTTGGATATAGCATGTATATGAACGAAACATATAAAATACAATGGTCAACAATTCGAGATGA 6800
SU N E S N I P Y L S T P G F G Y S M Y M N E T Y K I Q W S T I R D E
GTTTTCTGTCTCTTTTTATTTGTAAAAATGGCTCAGAGCATAGATATATTAGATGCAGACCTCCAAGTAATAATCAGTCAACACATTGTTTTTGGCAGGCA 6900
SU F S V S F I C K N G S E H R Y I R C R P P S N N Q S T H C F W Q A
GGATATGAAATGTTTCACAAACATTTTGTAAAAACTCCAGTAAGGGAAGACCCAGGAAGTTGGACATGCAGGACAGAAGGGGAGGTATTGTATGCTAGAT 7000
SU G Y E M F H K H F V K T P V R E D P G S W T C R T E G E V L Y A R C
RRE
GCACACACCCATTAGATTCTAAAAAGGAGCTCGGTTGCTATATAAGGGACTTGGAATGGGAAGAGAGAGAATGATAACATTCTTGGCTCCATATATGGTGGT 7100
SU T H P L D S K K E L G C Y I R D L E W E E R M I T F L A P Y M V V
AAAGGCAACTCCCTTCACATATGTACCAGTAAATATGTCTGATTTGACTATACCAATAAAACCGATGCACAAAAAAAGAGACTTTGGAGTAACAGCAGCT 7200
SU TM K A T P F T Y V P V N M S D L T I P I K P M H K K R D F G V T A A
SU TM
ATTGTCACTATAGCAGTGTCAGCAGCCACGGTAGCTGGCGCTGTGACTGGAGCACTGGCTCTAAGCACTACCCAACTACAGGGAGATGCTTTGGAGTCCC 7300
TM I V T I A V S A A T V A G A V T G A L A L S T T Q L Q G D A L E S L
TCCTGAAGGTAATTCAGGAGCAGCGGGCTCAGCTAGGTGACCAGTCAGCACTGCTTAAGACACATGCTATGGGGCTCCAGATGCTCGAAGCACACACGGT 7400
TM L K V I Q E Q R A Q L G D Q S A L L K T H A M G L Q M L E A H T V
ACAAATAGAACAGATAATTACTATACTAGCTTTGGAAAGAGAACTCAAGTGTGAAGCGATTGGAAGAGTTTGCATTACCACCATTCCGTGGAACAATCTA 7500
TM Q I E Q I I T I L A L E R E L K C E A I G R V C I T T I P W N N L
TCCATTCCTAATGCAACACAACTAGCTGATATGTTCAAACACAACCATTCTACCTGGCTAGAGTGGGTAAATGCAACCGCTCATCTTGAGGCTAACATTA 7600
TM S I P N A T Q L A D M F K H N H S T W L E W V N A T A H L E A N I T



Key

MA	Matrix
CA	Capsid
NC	Nucleocapsid
RVP	Protease
RVT	Reverse transcriptase
RVE	Endonuclease/integrase
SU	Envelope Surface Unit protein
TM	Envelope Transmebrane protein
P.B.S	Primer Binding Site
PPT	Polypurine Tract
I.R.	Inverted Repeats
TAR	Transactivation responsive region
RRE	Rev Responsive Element
NLS	Nuclear Localisation Signal
NES	Nuclear Export Signal