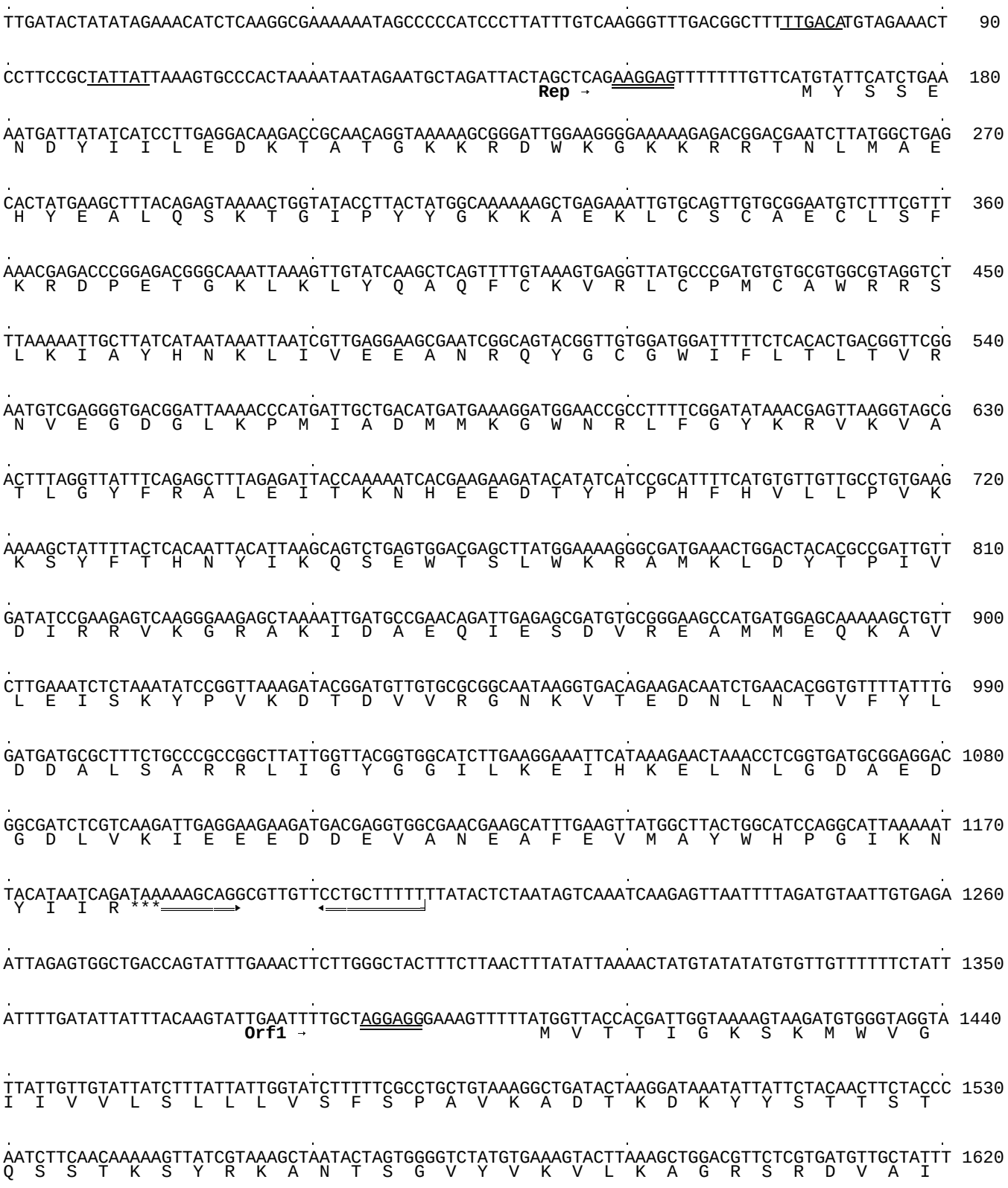
Replication and maintenance of plasmids in Bacillus subtilis

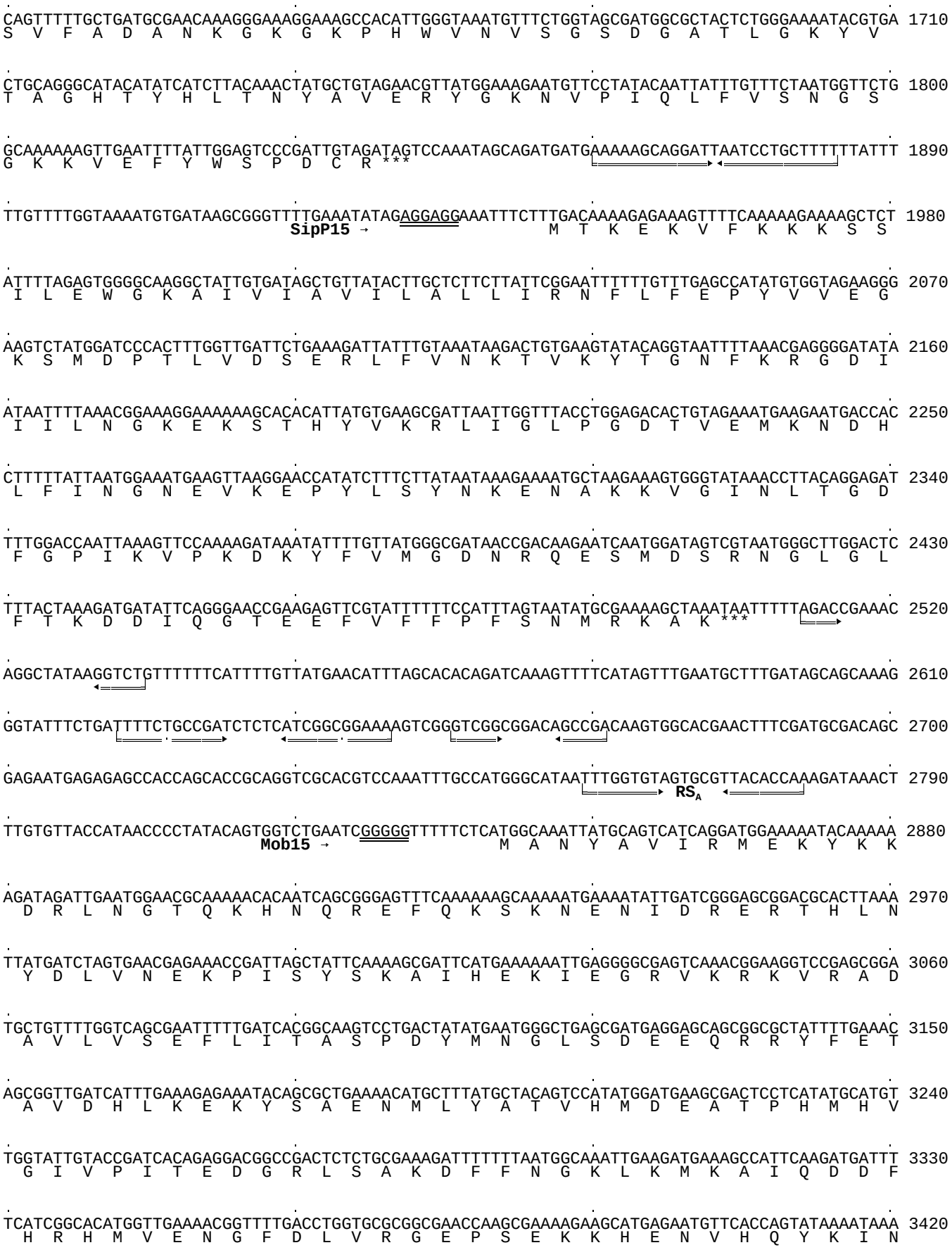
Meijer, Wilhelmus Johannes Jozef

Addendum

Complete nucleotide sequences of plasmids pTA1015, pTA1040 and pTA1060

Sequence pTA1015 :




CAGTTTTTGCTGATGCGAACAAAGGGAAAGGAAAGCCACATTGGGTAAATGTTTCTGGTAGCGATGGCGCTACTCTGGGAAAATACGTGA 1710
S V F A D A N K G K G K P H W V N V S G S D G A T L G K Y V
CTGCAGGGCATACATATCATCTTACAAACTATGCTGTAGAACGTTATGGAAAGAATGTTCCTATACAATTATTTGTTTCTAATGGTTCTG 1800
T A G H T Y H L T N Y A V E R Y G K N V P I Q L F V S N G S
GCAAAAAAGTTGAATTTTATTGGAGTCCCGATTGTAGATAGTCCAAATAGCAGATGATGAAAAAGCAGGATTAATCCTGCTTTTTTATTT 1890
G K K V E F Y W S P D C R ***
TTGTTTTGGTAAAATGTGATAAGCGGGTTTTGAAATATAGAGGAGGAAATTTCTTTGACAAAAGAGAAAGTTTTCAAAAAGAAAAGCTCT 1980
SipP15 →
M T K E K V F K K K S S
ATTTTAGAGTGGGGCAAGGCTATTGTGATAGCTGTTATACTTGCTCTTCTTATTCGGAATTTTTTGTTTGAGCCATATGTGGTAGAAGGG 2070
I L E W G K A I V I A V I L A L L I R N F L F E P Y V V E G
AAGTCTATGGATCCCACTTTGGTTGATTCTGAAAGATTATTTGTAAATAAGACTGTGAAGTATACAGGTAATTTTAAACGAGGGGATATA 2160
K S M D P T L V D S E R L F V N K T V K Y T G N F K R G D I
ATAATTTTAAACGGAAAGGAAAAAAGCACACATTATGTGAAGCGATTAATTGGTTTACCTGGAGACACTGTAGAAATGAAGAATGACCAC 2250
I I L N G K E K S T H Y V K R L I G L P G D T V E M K N D H
CTTTTTATTAATGGAAATGAAGTTAAGGAACCATATCTTTCTTATAATAAAGAAAATGCTAAGAAAGTGGGTATAAACCTTACAGGAGAT 2340
L F I N G N E V K E P Y L S Y N K E N A K K V G I N L T G D
TTTGGACCAATTAAAGTTCCAAAAGATAAATATTTTGTTATGGGCGATAACCGACAAGAATCAATGGATAGTCGTAATGGGCTTGGACTC 2430
F G P I K V P K D K Y F V M G D N R Q E S M D S R N G L G L
TTTACTAAAGATGATATTCAGGGAACCGAAGAGTTCGTATTTTTTCCATTTAGTAATATGCGAAAAGCTAAATAATTTTTAGACCGAAAC 2520
F T K D D I Q G T E E F V F F P F S N M R K A K ***
AGGCTATAAGGTCTGTTTTTTCATTTTGTTATGAACATTTAGCACACAGATCAAAGTTTTCATAGTTTGAATGCTTTGATAGCAGCAAAG 2610
GGTATTTCTGATTTTCTGCCGATCTCTCATCGGCGGAAAAGTCGGGTCGGCGGACAGCCGACAAGTGGCACGAACTTTCGATGCGACAGC 2700
GAGAATGAGAGAGCCACCAGCACCGCAGGTCGCACGTCCAAATTTGCCATGGGCATAATTTGGTGTAGTGCGTTACACCAAAGATAAACT 2790
RS$_A$
TTGTGTTACCATAACCCCTATACAGTGGTCTGAATCGGGGGGTTTTTCTCATGGCAAATTATGCAGTCATCAGGATGGAAAAATACAAAAA 2880
Mob15 →
M A N Y A V I R M E K Y K K
AGATAGATTGAATGGAACGCAAAAACACAATCAGCGGGAGTTTCAAAAAAGCAAAAATGAAAATATTGATCGGGAGCGGACGCACTTAAA 2970
D R L N G T Q K H N Q R E F Q K S K N E N I D R E R T H L N
TTATGATCTAGTGAACGAGAAACCGATTAGCTATTCAAAAGCGATTCATGAAAAAATTGAGGGGCGAGTCAAACGGAAGGTCCGAGCGGA 3060
Y D L V N E K P I S Y S K A I H E K I E G R V K R K V R A D
TGCTGTTTTGGTCAGCGAATTTTTGATCACGGCAAGTCCTGACTATATGAATGGGCTGAGCGATGAGGAGCAGCGGCGCTATTTTGAAAC 3150
A V L V S E F L I T A S P D Y M N G L S D E E Q R R Y F E T
AGCGGTTGATCATTTGAAAGAGAAATACAGCGCTGAAAACATGCTTTATGCTACAGTCCATATGGATGAAGCGACTCCTCATATGCATGT 3240
A V D H L K E K Y S A E N M L Y A T V H M D E A T P H M H V
TGGTATTGTACCGATCACAGAGGACGGCCGACTCTCTGCGAAAGATTTTTTTAATGGCAAATTGAAGATGAAAGCCATTCAAGATGATTT 3330
G I V P I T E D G R L S A K D F F N G K L K M K A I Q D D F
TCATCGGCACATGGTTGAAAACGGTTTTGACCTGGTGCGCGGCGAACCAAGCGAAAAGAAGCATGAGAATGTTCACCAGTATAAAATAAA 3420
H R H M V E N G F D L V R G E P S E K K H E N V H Q Y K I N

```
TCAGCGGCAAGCGGAGCTTGAGCGGCTTAATGCTGAAATTGCTTTAAAGGAAAAGCAGAGAGAGGAACTGGAAAAGCAAAACAAAGCTGT 3510
  Q  R  Q  A  E  L  E  R  L  N  A  E  I  A  L  K  E  K  Q  R  E  E  L  E  K  Q  N  K  A  V

TCAAGCAGTTATAGAAGTGAAAAAAGAATCGCTGACAGCTAAGGCTGAAGAGTTGAAAATGCCGACTATTGAACATGAAAAAGCGTGGCT 3600
  Q  A  V  I  E  V  K  K  E  S  L  T  A  K  A  E  E  L  K  M  P  T  I  E  H  E  K  A  W  L

CAAAAAGGATAAAGTCATTGTGCCAGAGCGGGAACTCCATGCTTTGTATGCCTATGCGGAGCAGAAAACTAAAACGGCAGCCGAGTTGGC 3690
  K  K  D  K  V  I  V  P  E  R  E  L  H  A  L  Y  A  Y  A  E  Q  K  T  K  T  A  A  E  L  A

GGGGCAATTGAAGTCGGAAACGCAGGAAAAGGAGCGCTGGCAGTCTATCGCCCGGCAGGAAGCAGATCGGGCGGATGAAAAAGACCAACG 3780
  G  Q  L  K  S  E  T  Q  E  K  E  R  W  Q  S  I  A  R  Q  E  A  D  R  A  D  E  K  D  Q  R

GCTTCAGGAACTGCAGAGTAGGATCCATTCAGAAGTTGAAGCGTCCAAAAAGGAAATGCGGCGCAAGCTTGCAAAGGAATTTACGGAAGA 3870
  L  Q  E  L  Q  S  R  I  H  S  E  V  E  A  S  K  K  E  M  R  R  K  L  A  K  E  F  T  E  E

ACAGCGTCAGGATCTTCGGCAGGAAGTGAAAGAGGAACTGACGACTTTACGAACGGAAAACGAGGAACTGTCAGCTGAAAATAAAGTTTT 3960
  Q  R  Q  D  L  R  Q  E  V  K  E  E  L  T  T  L  R  T  E  N  E  E  L  S  A  E  N  K  V  L

GATCATTCAAAGAAATAGCGAAGCTGCGGAGAGCCTAAAACTAAAACAGGAACTTGATAAGAGAAACGGGCAGTATGCTGAGGTTTTGAG 4050
  I  I  Q  R  N  S  E  A  A  E  S  L  K  L  K  Q  E  L  D  K  R  N  G  Q  Y  A  E  V  L  S

TTTCGCCAAGAAGCAGAATCAAACGCTTGAAAAAGTGGCTGGAGAAAACAAGGCGTTAAAAAAAGAAAATAAGACACTAAAAGAGAGAGT 4140
  F  A  K  K  Q  N  Q  T  L  E  K  V  A  G  E  N  K  A  L  K  K  E  N  K  T  L  K  E  R  V

TGCCGTACTGGACCAATGGAAAGACAAAATGGTTCAGTGGGCTAAAGAAAAATTACCAAAGATGCGGAAATTAGCGGCATCGTTTTCCG 4230
  A  V  L  D  Q  W  K  D  K  M  V  Q  W  A  K  E  K  L  P  K  M  R  K  L  A  A  S  F  F  R

TACGGCTGGAATGCCTAGAGAAGCCAATAAATACAAGGACAATGAATTAGAGCGGTGAAAACGGGCAATCAATTGCCCTTCCAAAATTTC 4320
  T  A  G  M  P  R  E  A  N  K  Y  K  D  N  E  L  E  R ***           *** N  G  K  W  F  K

CACCACTTTTTTTGTTTGGTGGCTGCGATCATTTTTTGTGTCTCAAGGGACTCACGTAATGCAGCGGTGAGTGTCTCGTGCCTTTCTTGC 4410
W  W  K  K  Q  K  T  A  A  I  M  K  Q  T  E  L  S  E  R  L  A  A  T  L  T  E  H  R  E  Q


TGTCTTTTTTCGAATCGTTCCATCCGTTCTGCCATGTGGCGGTTGAATTCTTCCTGTCGTTTCATGAATTCGACTAAAGGATTGTCCTGT 4500
Q  R  K  E  F  R  E  M  R  E  A  M  H  R  N  F  E  E  Q  R  K  M  F  E  V  L  P  N  D  Q

AGCGATGTAGCGGTATCCGATATGGTCAGTTTAGAACGATATAAGCTTGCTATGTGCTTTACCGTTTCGTCGAGTGAGTGTCCATTGATC 4590
L  S  T  A  T  D  S  I  T  L  K  S  R  Y  L  S  A  I  H  K  V  T  E  D  L  S  H  G  N  I

TTAGTCATTGTACATAGATACTCTAAAGTCTTTACGTCATCCTCGGTATAGAGTCGCCATCCTTTCGAATCTTTATTGAACGAGTAGCCT 4680
K  T  M  T  C  L  Y  E  L  T  K  V  D  D  E  T  Y  L  R  W  G  K  S  D  K  N  F  S  Y  G

TGTTCTTCAAGCATACTGGCATACTTGCGGACAGTTACTGGCTCTATACCGAGGTGTTTTGCGACGTCCTTTGATGATAATTTGATTCCC 4770
Q  E  E  L  M  S  A  Y  K  R  V  T  V  P  E  I  G  L  H  K  A  V  D  K  S  S  L  K  I  G

ATATCCATAACGTATCACCTCGTAAAAAGGGTTCGCTATAGTGAGGGCAAAGCCTTTATTTAGAGCCATTTAAAAAGAAAGCTGTTCCGT 4860
M  D  M       ← Orf2C15

TCTTTTATCAGGTTGGGACTGTATTTAGTTTATGGGGGAGTGCCAATGAGTGATTTTGAAACAGGAATGAGATATGTAAGAGCAACCCTT 4950

GGCTTTGAGGGTTTAGTGCTGACGGAAGAAGAGGAAAAGCTTTTAGAAAGACGGTTTCATGGAGAAATCACAGAGGAAGAATACATACAA 5040

AAAGCGTTGAGCTTCTTGATGTGATGTACCTGCAGGGCTTATAATATTCATAGCTGTTGTTAGAACCTAAACCCTTTTACCGTTTTTGGC 5130
```

GGTGAGGGTTCTTTTTTTTTAGGCAGTGATGCGATTTTGGCGTGAGTCAACGGTAACCGGACCGTAGGGAGGATTAAGGAGTTGACTCGC 5220

TCAGCGCCACCCGAACCCTTTCAGCACTCAAACAAACCCGTTTGTTTGACGCCAACCGGCAGGGAGCCCCCCGAAGAAGCGGGGGTTGGG 5310

GGGATTGAATGCTGGCATCCAACGGCCGTCCGTTGGTGGGTTTGGGCAAAGCCAAGAACTGTTGCAAGGCTCGTTGAGAATAAAGAATGC 5400

TTTTCAGGATGCTTAGAATCGTTTCTGAGAGCTTCAAATAAAAAAGATGACCTTTTATAGGGGGAAGCTCTTAAAATTGAATGTAGGGGC 5490

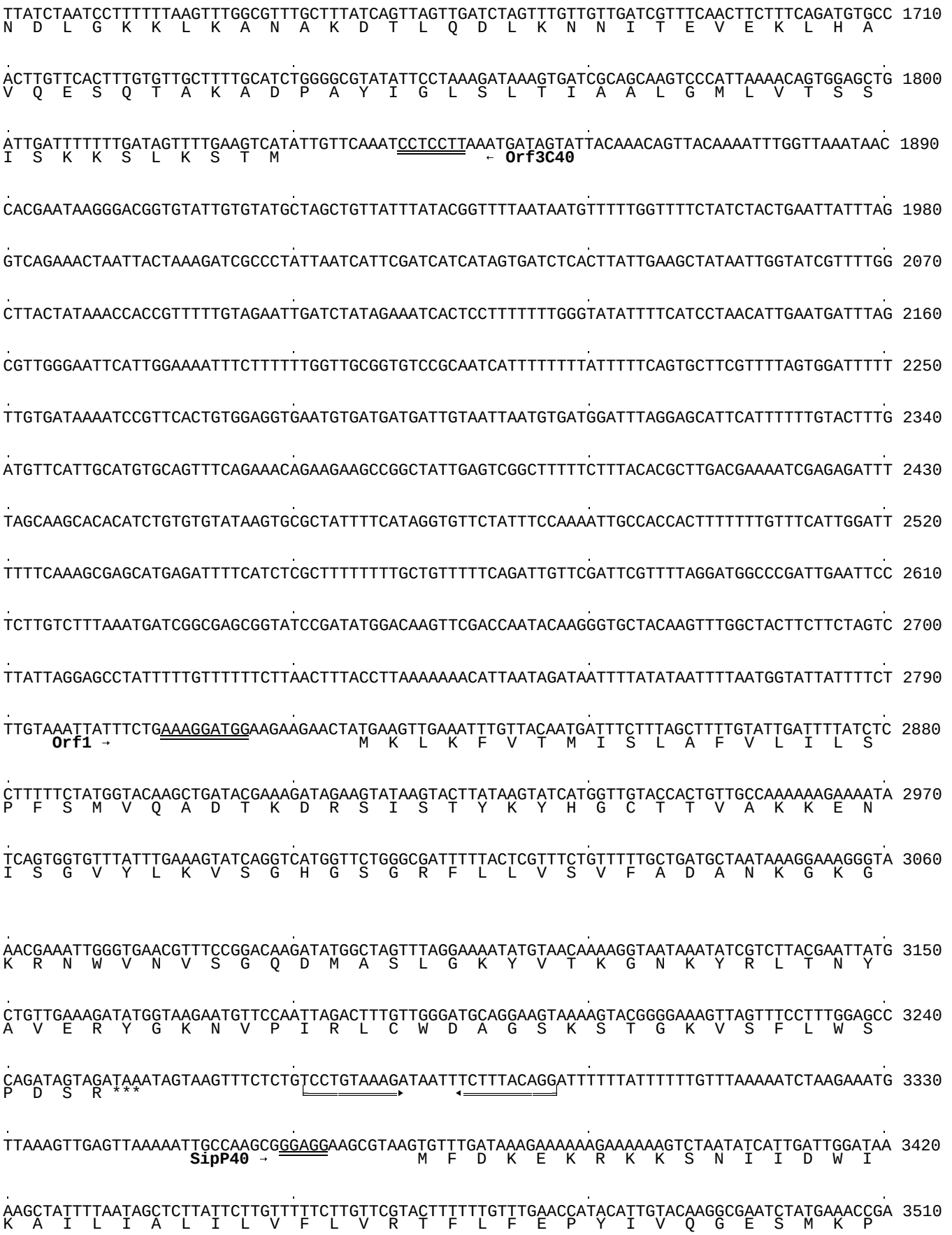
ATTTAAACACGTTTAAAAATAAAAAAAGCAGACTCTTTAGAGTCCGCCTTGTTATTTTTAACCCAGTGCTCCATTTTCGGCTGTTTGGAA 5580

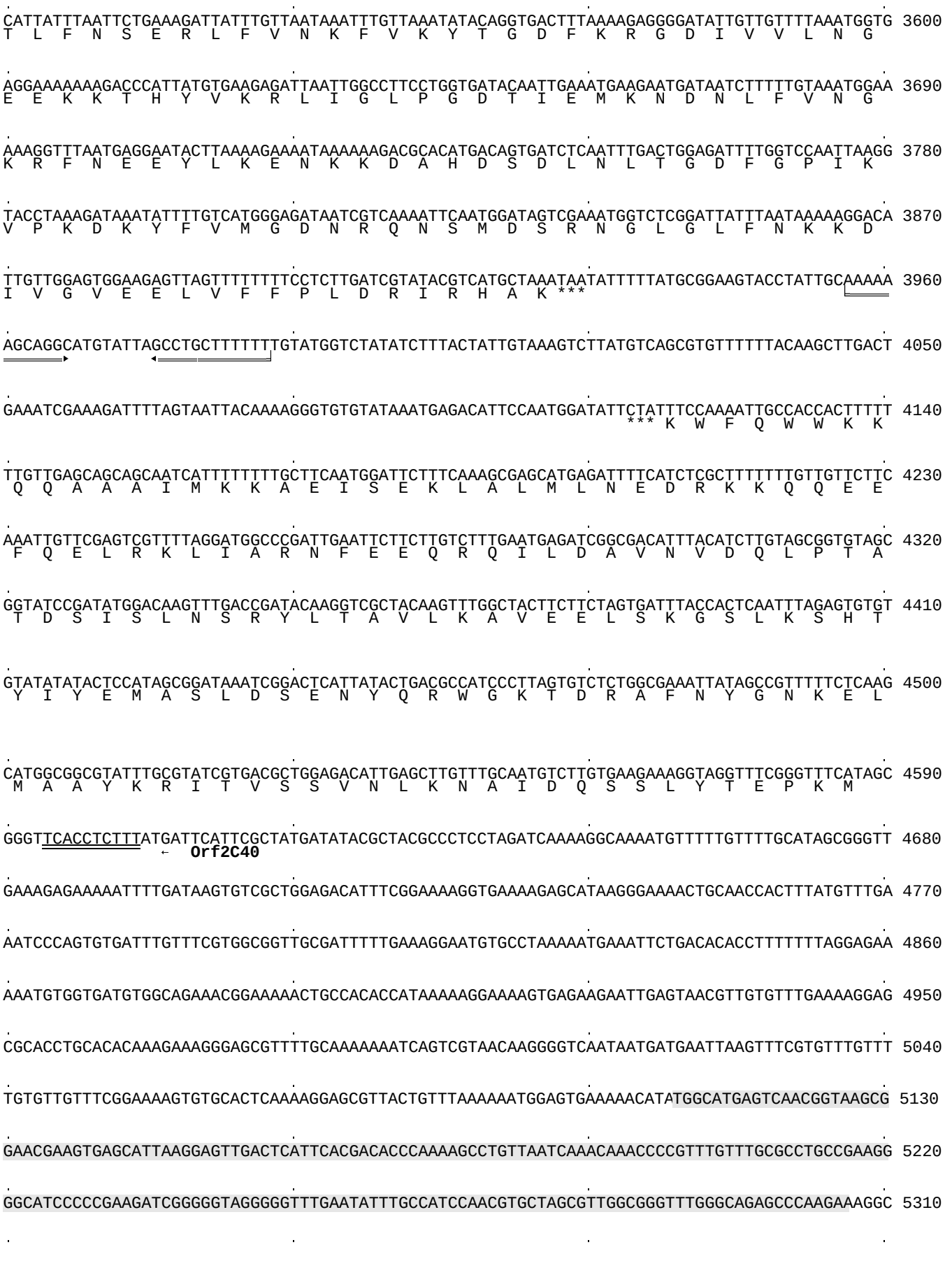
ATCTTTTGAGATGCCGAACCATCCATTTTCTTTTGTTCCATGAAAAAAGTGCTTTTGGATGCTTAAAAAGGCTTTTTCGTATAAAAAAAG 5670

CCCGATTTTTGAAAAAAAAATCTCCCCCCCTGCGGAAGAATGGTTTTGATCTTTGGGTTTTAGGTTTTAAAAAAAGCCGGCTGTTTTCAG 5760

CCGGTTTTTTTCGATTTTGGCGGACGGAAATCGGGTCTTTTCTTATC 5807

Sequence pTA1040 :

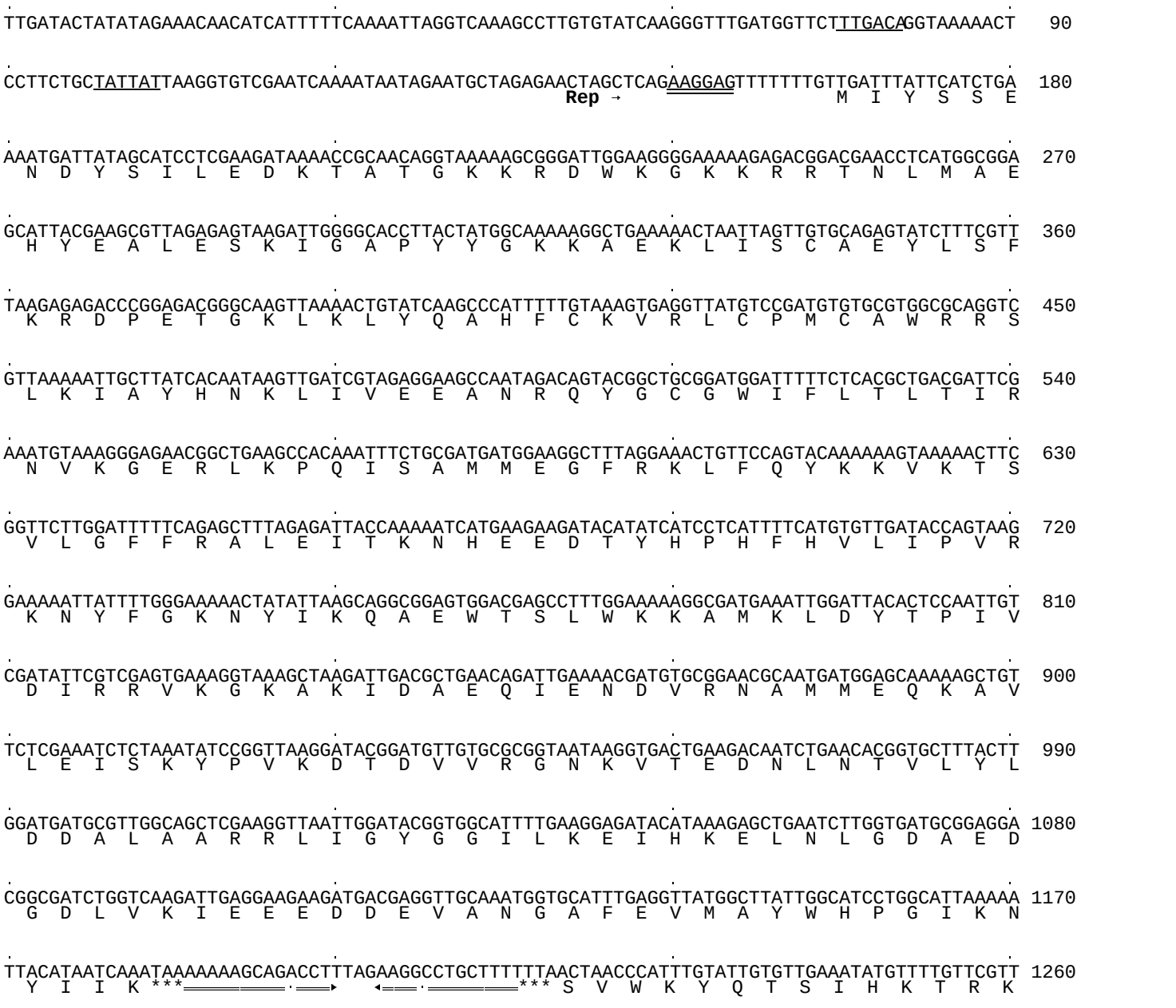

TTATCTAATCCTTTTTTAAGTTTGGCGTTTGCTTTATCAGTTAGTTGATCTAGTTTGTTGTTGATCGTTTCAACTTCTTTCAGATGTGCC 1710
N D L G K K L K A N A K D T L Q D L K N N I T E V E K L H A
ACTTGTTCACTTTGTGTTGCTTTTGCATCTGGGGCGTATATTCCTAAAGATAAAGTGATCGCAGCAAGTCCCATTAAAACAGTGGAGCTG 1800
V Q E S Q T A K A D P A Y I G L S L T I A A L G M L V T S S
ATTGATTTTTTTGATAGTTTTGAAGTCATATTGTTCAAATCCTCCTTAAATGATAGTATTACAAACAGTTACAAAATTTGGTTAAATAAC 1890
I S K K S L K S T M
Orf3C40
CACGAATAAGGGACGGTGTATTGTGTATGCTAGCTGTTATTTATACGGTTTTAATAATGTTTTTGGTTTTCTATCTACTGAATTATTTAG 1980
GTCAGAAACTAATTACTAAAGATCGCCCTATTAATCATTCGATCATCATAGTGATCTCACTTATTGAAGCTATAATTGGTATCGTTTTGG 2070
CTTACTATAAACCACCGTTTTTGTAGAATTGATCTATAGAAATCACTCCTTTTTTTTGGGTATATTTTCATCCTAACATTGAATGATTTAG 2160
CGTTGGGAATTCATTGGAAAATTTCTTTTTTGGTTGCGGTGTCCGCAATCATTTTTTTTATTTTTCAGTGCTTCGTTTTAGTGGATTTTT 2250
TTGTGATAAAATCCGTTCACTGTGGAGGTGAATGTGATGATGATTGTAATTAATGTGATGGATTTAGGAGCATTCATTTTTTGTACTTTG 2340
ATGTTCATTGCATGTGCAGTTTCAGAAACAGAAGAAGCCGGCTATTGAGTCGGCTTTTTCTTTACACGCTTGACGAAAATCGAGAGATTT 2430
TAGCAAGCACACATCTGTGTGTATAAGTGCGCTATTTTCATAGGTGTTCTATTTCCAAAATTGCCACCACTTTTTTTGTTTCATTGGATT 2520
TTTTCAAAGCGAGCATGAGATTTTCATCTCGCTTTTTTTTGCTGTTTTTCAGATTGTTCGATTCGTTTTAGGATGGCCCGATTGAATTCC 2610
TCTTGTCTTTAAATGATCGGCGAGCGGTATCCGATATGGACAAGTTCGACCAATACAAGGGTGCTACAAGTTTGGCTACTTCTTCTAGTC 2700
TTATTAGGAGCCTATTTTTGTTTTTTCTTAACTTTACCTTAAAAAAACATTAATAGATAATTTTATATAATTTTAATGGTATTATTTTCT 2790
TTGTAAATTATTTCTGAAAGGATGGAAGAAGAACTATGAAGTTGAAATTTGTTACAATGATTTCTTTAGCTTTTGTATTGATTTTATCTC 2880
Orf1
M K L K F V T M I S L A F V L I L S
CTTTTTCTATGGTACAAGCTGATACGAAAGATAGAAGTATAAGTACTTATAAGTATCATGGTTGTACCACTGTTGCCAAAAAAGAAAATA 2970
P F S M V Q A D T K D R S I S T Y K Y H G C T T V A K K E N
TCAGTGGTGTTTATTTGAAAGTATCAGGTCATGGTTCTGGGCGATTTTTACTCGTTTCTGTTTTTGCTGATGCTAATAAAGGAAAGGGTA 3060
I S G V Y L K V S G H G S G R F L L V S V F A D A N K G K G
AACGAAATTGGGTGAACGTTTCCGGACAAGATATGGCTAGTTTAGGAAAATATGTAACAAAAGGTAATAAATATCGTCTTACGAATTATG 3150
K R N W V N V S G Q D M A S L G K Y V T K G N K Y R L T N Y
CTGTTGAAAGATATGGTAAGAATGTTCCAATTAGACTTTGTTGGGATGCAGGAAGTAAAAGTACGGGGAAAGTTAGTTTCCTTTGGAGCC 3240
A V E R Y G K N V P I R L C W D A G S K S T G K V S F L W S
CAGATAGTAGATAAATAGTAAGTTTCTCTGTCCTGTAAAGATAATTTCTTTACAGGATTTTTTATTTTTTGTTTAAAAATCTAAGAAATG 3330
P D S R ***
TTAAAGTTGAGTTAAAAATTGCCAAGCGGGAGGAAGCGTAAGTGTTTGATAAAGAAAAAAGAAAAAAGTCTAATATCATTGATTGGATAA 3420
SipP40
M F D K E K R K K S N I I D W I
AAGCTATTTTAATAGCTCTTATTCTTGTTTTTCTTGTTCGTACTTTTTTGTTTGAACCATACATTGTACAAGGCGAATCTATGAAACCGA 3510
K A I L I A L I L V F L V R T F L F E P Y I V Q G E S M K P

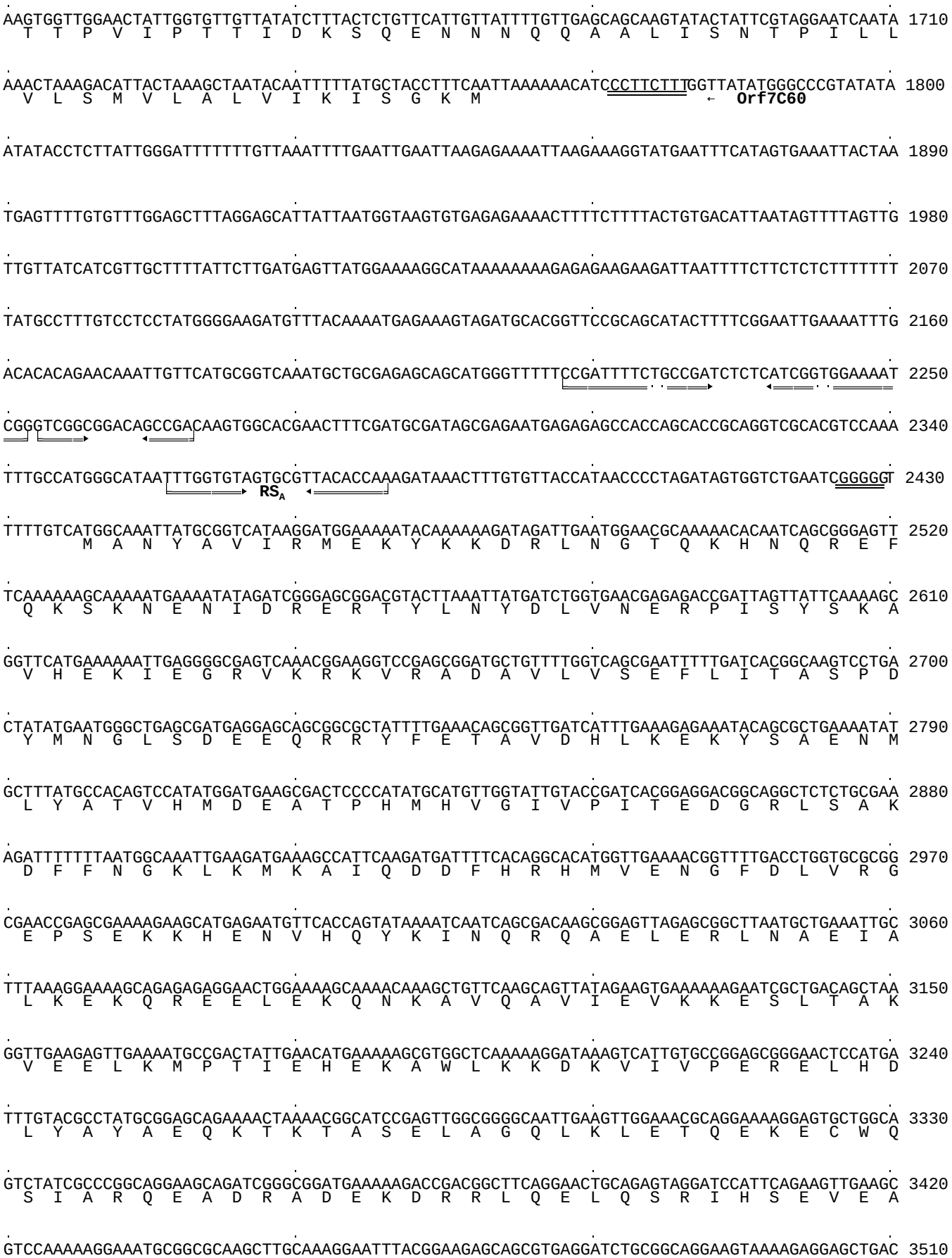

Orf2C40

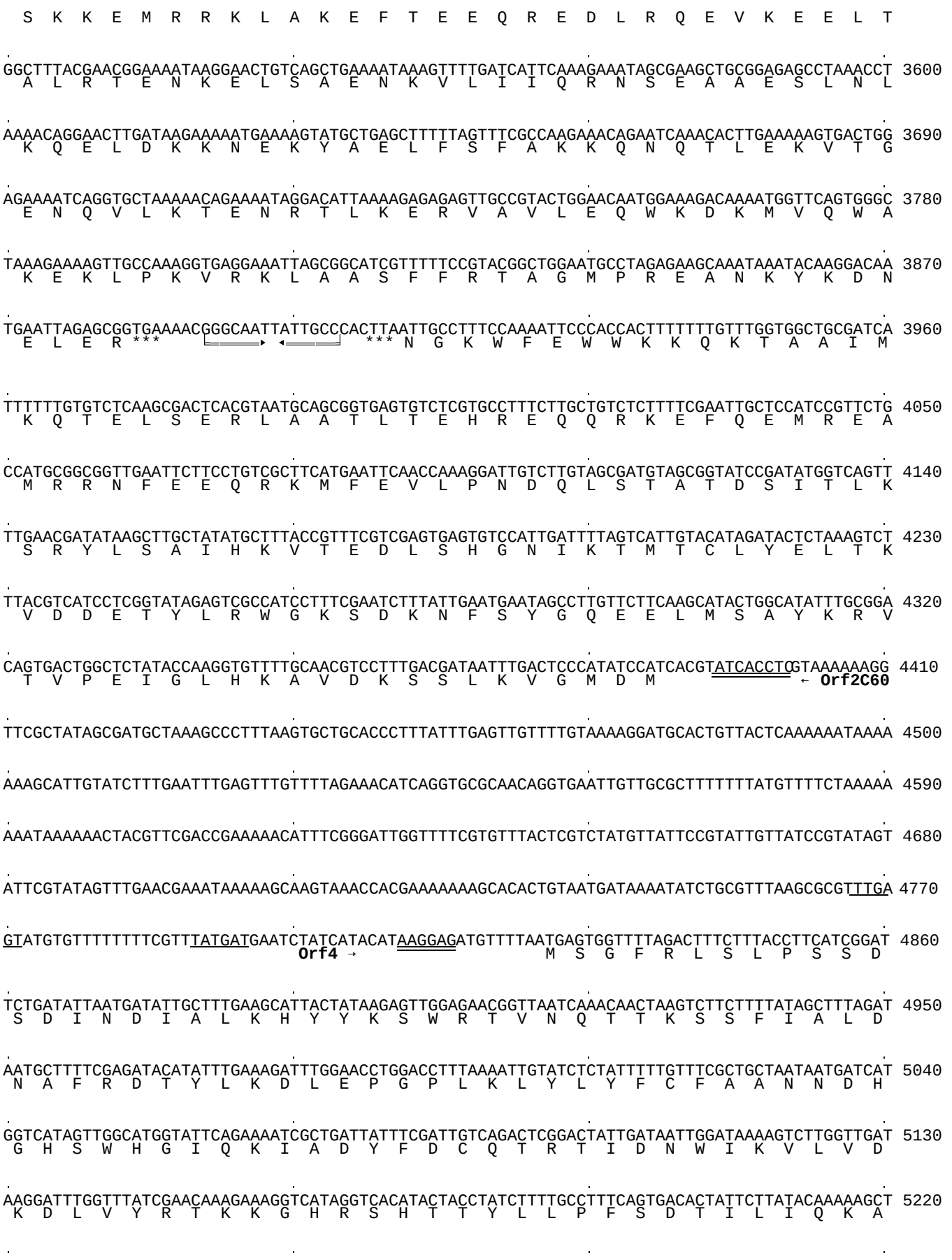
```
GGTCGGATGATGAAAATGACAAGAGAGCGTTTTGAAAAAGGCATGAGATATGCGAAAGCGACACATGCTATTGAAGGAATTTATCTTACA 5400

GCTGACGAGGAAGAATTATTATGGCAGCATGCATCCGGTCAGATCACAGATGAGGAATTTGAGAGAAAGGCTCTTGAATTGGCTTATAAA 5490

GTGATACAATAGAGTTCTAGGCGAAGGTGCGTAGCGGAAGACGACGACAGTCGTCGAGTGGCAACGAAAAGCGAAGCACCGTAGCCAGCT 5580

TGAACCCCCTCACTCTAACAGGGGGGTAGTAATGTAAGATGAAAATAAAACGTATAAAAAAATATGTCGTGCATTTCAAAAATGCCGTTG 5670

AGAGCGTTCAGATTGGATTCTGAGCGCTTTTATTTTTGTTTAGGTAATTGAAATAGGTGAAAGCGTCAAAAACGCTTATAAGGGCTTTAA 5760

ATGCGTTTTTTCGTTTTGTCTGCTGTCATGATTCCTAAAATTTGATGGTGAGATACCGAGCAAAATTCAAAAACGGTTGCGAGATACCGA 5850
                                                                                ###############

AAATTTTCAAAATTGGAACTATAATGGATTTTATAACGTACAAATTATGCTAGCATTTAACATAAAAAGGGATGTCGAAGGAGAGAGATG 5940
## ComA-box #################                                    Rap40  →

TGATGGAAAGTATTCCTGCTGCTGTGGTCGGACAGAAGATTAATGAATGGTATCGGCATATCAAAAAATTGAATGTAACAGATGCGGAAA 6030
   M  E  S  I  P  A  A  V  V  G  Q  K  I  N  E  W  Y  R  H  I  K  K  L  N  V  T  D  A  E

TGCTTCGTGAAGAGATTAAACGCGAATTAGATGTGATGGAAGAAGATGAACAAGCTGTTCTCTATTTTCAGTTGATGGACTTTCGTCATG 6120
M  L  R  E  E  I  K  R  E  L  D  V  M  E  E  D  E  Q  A  V  L  Y  F  Q  L  M  D  F  R  H

AGCAGATGCTTGAGTATGTAAATCCTTCTCAAAATCAAGTGAACAAAGCTGATTATTTGCGAGCTGTTGAAGGTGAAGGTAAAAAGTTGA 6210
E  Q  M  L  E  Y  V  N  P  S  Q  N  Q  V  N  K  A  D  Y  L  R  A  V  E  G  E  G  K  K  L


CGGGTATCATGGAATACTACGTTCACTTTTTTAAAGGGATGTATGAATTCTCTAAGGGTGAGTACATAAAAGCCATTGCTTTTTATAGAC 6300
T  G  I  M  E  Y  Y  V  H  F  F  K  G  M  Y  E  F  S  K  G  E  Y  I  K  A  I  A  F  Y  R

ATGCAGAAAAGAAGCTGGATAAAGTAGCTGATGAGCTTGAACGGGCTGAATTTTACTATAAGATGTCAGAGATTTTCTATCACATGAAGC 6390
H  A  E  K  K  L  D  K  V  A  D  E  L  E  R  A  E  F  Y  Y  K  M  S  E  I  F  Y  H  M  K

AAACGCATTTGTCTATGTATTATATCGAGCAGGCGTATCATGCTTATAAAGCTCATAAAGCCAACGGGATGTATGAAATCAAGGTGATTC 6480
Q  T  H  L  S  M  Y  Y  I  E  Q  A  Y  H  A  Y  K  A  H  K  A  N  G  M  Y  E  I  K  V  I

AATGTCGTTTTGTCATTGCAGGAAATTATGATGATTTATGTACACATGAAAAAGCGATACCTCATTTAAAAGAAGCGGAAGAGTCGGCCA 6570
Q  C  R  F  V  I  A  G  N  Y  D  D  L  C  T  H  E  K  A  I  P  H  L  K  E  A  E  E  S  A

GGAATTTAGGCAACAAGTATATGCTTGCTAAAGCGTTGCTTAACCTCGGTGATTCTTATAAAAATATGGGTGATCAGGACAGAGCGCTTT 6660
R  N  L  G  N  K  Y  M  L  A  K  A  L  L  N  L  G  D  S  Y  K  N  M  G  D  Q  D  R  A  L

ATTTTTTGATCAAAGCTGTTGATCAGGCTAAAGAGTCTGGAGCTAAAGAGCTTACACAGATTTATTATGAGTTGGCTTTACTGCATTTTA 6750
Y  F  L  I  K  A  V  D  Q  A  K  E  S  G  A  K  E  L  T  Q  I  Y  Y  E  L  A  L  L  H  F

AGAAAAAGGAGCATGAATCGGGAAAACACTATTTTGAGTTGGGTGTTGAAAGTGCTGAAAGGTTCAATGATTGGTTTTTCACGAATTTGT 6840
K  K  K  E  H  E  S  G  K  H  Y  F  E  L  G  V  E  S  A  E  R  F  N  D  W  F  F  T  N  L

TAAATGTACTGAAGCTTCTTTATTTAGAGGGGGCTCAACCGACTGCTGTTGCGGATGCTTTAGAAGAACTTGACGATATTCGAGGCTATC 6930
L  N  V  L  K  L  L  Y  L  E  G  A  Q  P  T  A  V  A  D  A  L  E  E  L  D  D  I  R  G  Y

CATATTTAGAAGAACTGGCGCTTGTGTCTGCTGAATTTTATACTGAAATTGGGCATATGGATGAATCTGTATATTTTTATAACAAAATGA 7020
P  Y  L  E  E  L  A  L  V  S  A  E  F  Y  T  E  I  G  H  M  D  E  S  V  Y  F  Y  N  K  M

TCTATGCTCAAAAACAAATTCGGAGGGGCGAGTTCTTGTATGAAATTTAATGCCTTACTGCTTTTAATTGTTTGTGCTAGTTTATTAATC 7110
I  Y  A  Q  K  Q  I  R  R  G  E  F  L  Y  E  I
       Rap40B  →                    M  K  F  N  A  L  L  L  L  I  V  C  A  S  L  L  I
```

```
GTTAGTGGTTCTTCTTTTGTAATCCAACAAGATTCAAATGTGTCAGTTGCTTCAAGAAAGGCTACATGATGATGAGTTTTATTCGATTTG 7200
 V  S  G  S  S  F  V  I  Q  Q  D  S  N  V  S  V  A  S  R  K  A  T ***

GTACATGAAAAGTAACAGACCTTTAGGGTCTGTTTTTTTGTGGAAATATCATGTGTAGTCACTTGGAAATAACACGTAAACTATTAGGTT 7290

TTTGATTAAACGTGATATTTTCACTTGGTGCAGCCGACAGGAACCTATCGGCCATAAGATCGGACCGAATCGCGATGAAATGACCTTACT 7380

TGTTCTATGCTCTTCTGTTTGTCTCTTTTTCAATCCCACCATTTTTGTGGGCGTGTGATGGCAGCTGTATGTTGAAAGGCGGGTTTAGTC 7470

AGGTTTTGAGTGAAAAGCTACAGAAAAGGTGTTGCTGATCGTAAAGGCTACAGAAACGGTGTGGTGTTTTTTCTTGGCAAAATAAACCCT 7560

ATGATTGCAAGGTTTTGATATGTATTTGAGTGATTAAAAGCTATAAAAACGGTGTTAAAAGCTAATGAAAAGCTACAAAAAAGTGTTGAG 7650

GCTATAAAAACGGGGGGACTCAATTTTTGAGTAGGGGGCAAAAAATCAAAAAAAAATCTCCCCCTACGGGGAAATTTGGTTTAGTCTTTG 7740

GTCTTTTAGGTTTTGAGGTTTTGGTTTTTAAAAAGCTGGCGGTTCCGCCAGCTTTTTGGGCGATTTTGACGGAGTCAGAATCGTGTCTTT 7830

TCTTATC 7837
```

Sequence pTA1060 :




Orf7C60
RS$_A$


S K K E M R R K L A K E F T E E Q R E D L R Q E V K E E L T
GGCTTTACGAACGGAAAATAAGGAACTGTCAGCTGAAAATAAAGTTTTGATCATTCAAAGAAATAGCGAAGCTGCGGAGAGCCTAAACCT 3600
A L R T E N K E L S A E N K V L I I Q R N S E A A E S L N L
AAAACAGGAACTTGATAAGAAAAATGAAAAGTATGCTGAGCTTTTTAGTTTCGCCAAGAAACAGAATCAAACACTTGAAAAAGTGACTGG 3690
K Q E L D K K N E K Y A E L F S F A K K Q N Q T L E K V T G
AGAAAATCAGGTGCTAAAAACAGAAAATAGGACATTAAAAGAGAGAGTTGCCGTACTGGAACAATGGAAAGACAAAATGGTTCAGTGGGC 3780
E N Q V L K T E N R T L K E R V A V L E Q W K D K M V Q W A
TAAAGAAAAGTTGCCAAAGGTGAGGAAATTAGCGGCATCGTTTTTCCGTACGGCTGGAATGCCTAGAGAAGCAAATAAATACAAGGACAA 3870
K E K L P K V R K L A A S F F R T A G M P R E A N K Y K D N
TGAATTAGAGCGGTGAAAACGGGCAATTATTGCCCACTTAATTGCCTTTCCAAAATTCCCACCACTTTTTTTGTTTGGTGGCTGCGATCA 3960
E L E R *** *** N G K W F E W W K K Q K T A A I M
TTTTTTGTGTCTCAAGCGACTCACGTAATGCAGCGGTGAGTGTCTCGTGCCTTTCTTGCTGTCTCTTTTCGAATTGCTCCATCCGTTCTG 4050
K Q T E L S E R L A A T L T E H R E Q Q R K E F Q E M R E A
CCATGCGGCGGTTGAATTCTTCCTGTCGCTTCATGAATTCAACCAAAGGATTGTCTTGTAGCGATGTAGCGGTATCCGATATGGTCAGTT 4140
M R R N F E E Q R K M F E V L P N D Q L S T A T D S I T L K
TTGAACGATATAAGCTTGCTATATGCTTTACCGTTTCGTCGAGTGAGTGTCCATTGATTTTAGTCATTGTACATAGATACTCTAAAGTCT 4230
S R Y L S A I H K V T E D L S H G N I K T M T C L Y E L T K
TTACGTCATCCTCGGTATAGAGTCGCCATCCTTTCGAATCTTTATTGAATGAATAGCCTTGTTCTTCAAGCATACTGGCATATTTGCGGA 4320
V D D E T Y L R W G K S D K N F S Y G Q E E L M S A Y K R V
CAGTGACTGGCTCTATACCAAGGTGTTTTGCAACGTCCTTTGACGATAATTTGACTCCCATATCCATCACGTATCACCTCGTAAAAAAGG 4410
T V P E I G L H K A V D K S S L K V G M D M ← Orf2C60
TTCGCTATAGCGATGCTAAAGCCCTTTAAGTGCTGCACCCTTTATTTGAGTTGTTTTGTAAAAGGATGCACTGTTACTCAAAAAATAAAA 4500
AAAGCATTGTATCTTTGAATTTGAGTTTGTTTTAGAAACATCAGGTGCGCAACAGGTGAATTGTTGCGCTTTTTTTATGTTTTCTAAAAA 4590
AAATAAAAAACTACGTTCGACCGAAAAACATTTCGGGATTGGTTTTCGTGTTTACTCGTCTATGTTATTCCGTATTGTTATCCGTATAGT 4680
ATTCGTATAGTTTGAACGAAATAAAAAGCAAGTAAACCACGAAAAAAAGCACACTGTAATGATAAAATATCTGCGTTTAAGCGCGTTTGA 4770
GTATGTGTTTTTTTTCGTTTATGATGAATCTATCATACATAAGGAGATGTTTTAATGAGTGGTTTTAGACTTTCTTTACCTTCATCGGAT 4860
Orf4 → M S G F R L S L P S S D
TCTGATATTAATGATATTGCTTTGAAGCATTACTATAAGAGTTGGAGAACGGTTAATCAAACAACTAAGTCTTCTTTTATAGCTTTAGAT 4950
S D I N D I A L K H Y Y K S W R T V N Q T T K S S F I A L D
AATGCTTTTCGAGATACATATTTGAAAGATTTGGAACCTGGACCTTTAAAATTGTATCTCTATTTTTGTTTCGCTGCTAATAATGATCAT 5040
N A F R D T Y L K D L E P G P L K L Y L Y F C F A A N N D H
GGTCATAGTTGGCATGGTATTCAGAAAATCGCTGATTATTTCGATTGTCAGACTCGGACTATTGATAATTGGATAAAAGTCTTGGTTGAT 5130
G H S W H G I Q K I A D Y F D C Q T R T I D N W I K V L V D
AAGGATTTGGTTTATCGAACAAAGAAAGGTCATAGGTCACATACTACCTATCTTTTGCCTTTCAGTGACACTATTCTTATACAAAAAGCT 5220
K D L V Y R T K K G H R S H T T Y L L P F S D T I L I Q K A

```
TCGAAGAAGTACATTAAAGATGATCAGGACTTGCTTAATGGTCTATTAAAAAGAATAAGAGAACTTGACTTTATTTATGGGGAAATTATT 5310
 S  K  K  Y  I  K  D  D  Q  D  L  L  N  G  L  L  K  R  I  R  E  L  D  F  I  Y  G  E  I  I

AAGGTCCATCATCTTTTTCAATGGTCTGTTTCAAAAGAAAAAACTGTTGATGGAGATAAAAGCGATCAATATCTTTTGGTTATTACTAAA 5400
 K  V  H  H  L  F  Q  W  S  V  S  K  E  K  T  V  D  G  D  K  S  D  Q  Y  L  L  V  I  T  K

CGAAAAAACGGTGTCTTAATTGGACATGTTTATGTGTTGAGAAAATCTGATCATTTGAGTGTGAATGAATTACAGATTGAAGAACCATCT 5490
 R  K  N  G  V  L  I  G  H  V  Y  V  L  R  K  S  D  H  L  S  V  N  E  L  Q  I  E  E  P  S

GTATTTGATAGTCCTTTTTATCTTAATGAATCAAGAATAATTGGTCTTGCTTTGCCTCATTATCCTAAAATTTTTTCTAAGTCTGGATTA 5580
 V  F  D  S  P  F  Y  L  N  E  S  R  I  I  G  L  A  L  P  H  Y  P  K  I  F  S  K  S  G  L

AATGATTTATTGGGTTTAATAAGGGACTTAGTGATTATTGAAGATTGGCAGTTGGAAGATCGATCAAAACTTATTTATGGTGAAAAGCGG 5670
 N  D  L  L  G  L  I  R  D  L  V  I  I  E  D  W  Q  L  E  D  R  S  K  L  I  Y  G  E  K  R

GATTTGATTCAAATAGAAGAAATGGATAATCAAAGTCAGGTAGATGAAGATAGTGAAGATAAGAATGAGTAAAAGGGGATGGGGGAGTTA 5760
 D  L  I  Q  I  E  E  M  D  N  Q  S  Q  V  D  E  D  S  E  D  K  N  E  ***

TTATGAATTTTGAACAAATAGGGGTTATTAATGATACTGTGATTAATCAGTTAAATTTGACTATTAGCCCAACTCCAATTTTTATTGGAA 5850
   M  N  F  E  Q  I  G  V  I  N  D  T  V  I  N  Q  L  N  L  T  I  S  P  T  P  I  F  I  G

AATCTAATATTGATCATATGATTAAAGAGCATCCCGAGGATTTTAACAAATATGGTTGTTACTTAGAAGAAATAATTTCTTCTCCGGATT 5940
K  S  N  I  D  H  M  I  K  E  H  P  E  D  F  N  K  Y  G  C  Y  L  E  E  I  I  S  S  P  D

ATGTTTCTCTTCATCCGAACAATGGATCTATTCAATATATTAAAGTTTTTTTTGACGAAGATAATGAAAAACATGTTTTAGTTGCAGTCC 6030
Y  V  S  L  H  P  N  N  G  S  I  Q  Y  I  K  V  F  F  D  E  D  N  E  K  H  V  L  V  A  V

GAGCAACAAAGAAAGGAGTATTTTTTGCGCGTTCTTTGTTTGTGATGAGCGATTTAAAAGTAAAGAGCTATACAAATAAGGGGATGCTGA 6120
R  A  T  K  K  G  V  F  F  A  R  S  L  F  V  M  S  D  L  K  V  K  S  Y  T  N  K  G  M  L

AAAAATGTGTAGTTCCAAAACAAAAGACCACAACAGAGTGAGTGTTATGGTCTCATTATCTTTTCGCAAGGATTTAGGCGGGGTGGCACT 6210
K  K  C  V  V  P  K  Q  K  T  T  T  E  ***

CCCGCATCTCCAACATAGGGTGATGGCTGCCTGTTCCATCCTCAGAAATCCTTCTTATTAGTTTATTCAACTTTCTATATTTATTATACA 6300

TGTGAAAATTTTATGTTTCAAGGGTGGTTTTATTTAAACGAGTTTTTATGGCGTGAGTCAACGGTAACCGGACCGTAGGGAGGATTAAGG 6390

AGTTGACTCGCTCAGTGCTACCCGAACCCTTTCAGCACTCAAACAAACTCGTTTGTTTGACGCCAACCGGCGAGGGAGCCCCCCGGAGAT 6480

GCGGGGGGTTGGGGGGATTGAATGTTGGAATCCAACGGCCATCCGTTGGTGGGTTTGGGCAAAGCCCAAGGTTCTACTTTAACTCTTGAT 6570

GCAATTTGTTGTTTTTGTCGAGCAAGACAATTCCTTCAGGGCTTCGTTCTGGGAAGTGTTGTGTGAGATTTAGCTTTTGATCAGAAAGCA 6660

ACTCCTGCATTTTCTTTTGTGCTTCAGCGAGCGTTTCGATTTTCAATTTTTATTCCTCCGGCGTTTTTGTCATAATACTAGCATGAGGTG 6750

GCGCTACCAGTAGCCCACGTAAAAAATAGATAGCTGGATAAGAGGGACGGATGCTTTTTAGCAATCCGTCTTTTTTGTTTGAGAGGTTAA 6840

TAGATAATGATAAATAGCTCTGAGGATTCTCAGATTCGTTCCTGAGAGGTTTTAATTTTAAAATAGAGTAAAAGGTCACAGGGGAATTTA 6930

GGGACGCTGTATTGAGCTTTAAATGGATTTTTTGCATCTTTGTTTTCGAGAGAAAATTTTTGTTTTGAAAAATGAAACGAAAGGGCTTTT 7020

TTTCATGAGTAAGGATGCTTTTGTGTAAAAATAATCCTAATCGATTTTGTTAATTTAAAATATTTCGCTTTCGGCTTACCGAAAATTTAA 7110
                                                ################ ComA-box
```

```
GGTTCGGCGTCCGCAACCCCGATTTTACAGCTTTTTCTGTGCTACCATTTACAAAAAAAGTATTAAATTAGGGGAGGAGTTACTCGGAAT 7200
  ################                                                 Rap60 →      M

.                   .                   .                   .
GGAAGTTATTCCATCTGCGAATGTAGGGCTAAAGATAAATGAGTGGCACATGTACATAAATCGATTCGATGTTAGAAACGCTGAAAAGGT 7290
  E   V   I   P   S   A   N   V   G   L   K   I   N   E   W   H   M   Y   I   N   R   F   D   V   R   N   A   E   K   V

.                   .                   .                   .
AAAAGCGGAAGTTGAACAGATGATTGATGTTATGGAAGAAAATCAAGATTTGCTCATATATTATCAGCTGATCGAGTTCAGACACCGTAT 7380
  K   A   E   V   E   Q   M   I   D   V   M   E   E   N   Q   D   L   L   I   Y   Y   Q   L   I   E   F   R   H   R   M

.                   .                   .                   .
GATGCTTGATTATTTATTTCCGGCTGATCCCTCTAAAAAAATCAATAAAGGTGATTATCTAAAAAAAAATTGAGGGGAAAGGCAAAAAACT 7470
  M   L   D   Y   L   F   P   A   D   P   S   K   K   I   N   K   G   D   Y   L   K   K   I   E   G   K   G   K   K   L

.                   .                   .                   .
GTCTGGCATGATGGAATACTACACCTATTTTTTTAAAGGGATGTACGATTTTTCTATTGGTGAGTACATAAGAGCCATTCAGTCATACAA 7560
  S   G   M   M   E   Y   Y   T   Y   F   F   K   G   M   Y   D   F   S   I   G   E   Y   I   R   A   I   Q   S   Y   K

.                   .                   .                   .
AAAGGCTGAGAAAAAGCTTTCGAGGGTTTTGGATGAAATAGAGAAAGCTGAGTTTTATTTTAAAATGGCTGAAGTCTTTTACCATATGAA 7650
  K   A   E   K   K   L   S   R   V   L   D   E   I   E   K   A   E   F   Y   F   K   M   A   E   V   F   Y   H   M   K

.                   .                   .                   .
ACAAACGCATGTATCTATGTATTATGTGCAACAGGCATATGACATTTATAAAAAGCACGATACATATAAAATTCATCGTATCAAGTGTCA 7740
  Q   T   H   V   S   M   Y   Y   V   Q   Q   A   Y   D   I   Y   K   K   H   D   T   Y   K   I   H   R   I   K   C   H


.                   .                   .                   .
TTTTGTTATTGCCGGCAAATATGATGAATTAAGGACTCATGACAGAGCGCTTTCTCATATTGAAAAAGCCTTAATTTTGCTGGATAGTAT 7830
  F   V   I   A   G   K   Y   D   E   L   R   T   H   D   R   A   L   S   H   I   E   K   A   L   I   L   L   D   S   M

.                   .                   .                   .
GGAGGACCATTCCCCTTTAAAGGCAAAAGCTTTGTTGAATATGGGGAATAGTTATAACCAGATGGGAAGCCTTTTTTCGGCTGTACCTTA 7920
  E   D   H   S   P   L   K   A   K   A   L   L   N   M   G   N   S   Y   N   Q   M   G   S   L   F   S   A   V   P   Y

.                   .                   .                   .
TTATCACAAGGCTATAAAGGCAGCTAAAATAAGCGGAGCCAAGGAAATTACCCAGGCTTATTATGATTTAGCTCTGATTCATTTTAGGAA 8010
  Y   H   K   A   I   K   A   A   K   I   S   G   A   K   E   I   T   Q   A   Y   Y   D   L   A   L   I   H   F   R   N

.                   .                   .                   .
CAATGAAAAGATTGAAGGACGCGGATTCTTTGGAAAAGCAATGGAGCAAGCTGAAGAATTTAACGATAGTCTGTTTCAGGATTTATTAAA 8100
  N   E   K   I   E   G   R   G   F   F   G   K   A   M   E   Q   A   E   E   F   N   D   S   L   F   Q   D   L   L   N

.                   .                   .                   .
TGTTCTAAAGGCATTATTTATTGAGACTGGAAGCAGGCAAAAGGTCATGAATGCCCTTGAGGCGTTGAGAACAGGACAAGGCTATCCATA 8190
  V   L   K   A   L   F   I   E   T   G   S   R   Q   K   V   M   N   A   L   E   A   L   R   T   G   Q   G   Y   P   Y

.                   .                   .                   .
TTTTGAAGAGCTAGCTTTAATCGCTGCAGAATTTTATACTATGGATAAGCGCATGGAAGATTCTATATATTTTTACAATGAGATGGTGTG 8280
  F   E   E   L   A   L   I   A   A   E   F   Y   T   M   D   K   R   M   E   D   S   I   Y   F   Y   N   E   M   V   C

.                   .                   .                   .
CGCTCAAAGACAAATTCAAAGGGGCGATTTCCTGTATGAAGTTTAAAGGTCTTTTTTCAGCGGTCCTTATTGTAAGCCTGTTGGTTGGAG 8370
  A   Q   R   Q   I   Q   R   G   D   F   L   Y   E   V
      Rap60B →                       M   K   F   K   G   L   F   S   A   V   L   I   V   S   L   L   V   G

.                   .                   .                   .
CGGGGTATTCTTTCGTGCACCATGATGAGGTTTCAGTTGCTTCAAGGAATGCAACATGATGAAGAATAAAGAATCGAGATTTACAACTTA 8460
A   G   Y   S   F   V   H   H   D   E   V   S   V   A   S   R   N   A   T  ***

.                   .                   .                   .
AAGATATGGAAAACAGATCGTTATTGATCTGTTTTTTTCTTTGTGTGCGACCATGCGGGCATGCTGCGGAGATAACACGGGCTGCTGCGC 8550

.                   .                   .                   .
AGATGATGCGGGCATGGGTGCGTTGATCGTGCCGAGACTGAAAAGGCAAAATTTCAAAAAAAAAATCTCCCCCCCTACGGGGGGGAAGAAT 8640

.                   .                   .                   .
GGTTTTGATCTTTTGTTTTGGGTTTTAAAAAAAGCCGGCTGTTTTTCAGCCGGTCTTTTTCGATTTTGGCGAAGCCGAAATCGGGTCTTT 8730

.
TCTTATC 8837
```