

University of Groningen

Replication and maintenance of plasmids in *Bacillus subtilis*

Meijer, Wilhelmus Johannes Jozef

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1995

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Meijer, W. J. J. (1995). *Replication and maintenance of plasmids in Bacillus subtilis*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Addendum

Complete nucleotide sequences of plasmids pTA1015, pTA1040 and pTA1060

Sequence pTA1015 :

```

TTGATACTATATAGAAACATCTCAAGGCGAAAAAATAGCCCCCATCCCTTATTTGTCAAGGGTTTGACGGCTTTTTGACATGTAGAAACT 90
.
CCTTCCGCTATTATTAAGTGCCCACTAAAATAATAGAATGCTAGATTACTAGCTCAGAAGGAGTTTTTTTGTTCATGTATTTCATCTGAA 180
.
Rep -> M Y S S E
AATGATTATATCATCCTTGAGGACAAGACCGCAACAGGTAAAAAGCGGGATTGGAAGGGGAAAAAGAGACGGACGAATCTTATGGCTGAG 270
.
N D Y I I L E D K T A T G K K R D W K G K K R R T N L M A E
CACTATGAAGCTTTACAGAGTAAAACTGGTATACCTTACTATGGCAAAAAAGCTGAGAAATTGTGCAGTTGTGCGGAATGTCTTTCGTTT 360
.
H Y E A L Q S K T G I P Y Y G K K A E K L C S C A E C L S F
AAACGAGACCCGGAGACGGGCAAATTAAGTTGTATCAAGCTCAGTTTTGTAAAGTGAGGTTATGCCCGATGTGTGCGTGGCGTAGGTCT 450
.
K R D P E T G K L K L Y Q A Q F C K V R L C P M C A W R R S
TTAAAAATGCTTATCATAATAAATTAATCGTTGAGGAAGCGAATCGGCAGTACGGTTGTGGATGGATTTTTCTCACACTGACGGTTCCGG 540
.
L K I A Y H N K L I V E E A N R Q Y G C G W I F L T L T V R
AATGTCGAGGGTGACGGATTAAAACCCATGATTGCTGACATGATGAAAGGATGGAACCGCTTTTCGGATATAAACGAGTTAAGGTAGCG 630
.
N V E G D G L K P M I A D M M K G W N R L F G Y K R V K V A
ACTTTAGGTTATTTTCAGAGCTTTAGAGATTACCAAAAATCACGAAGAAGATACATATCATCCGCATTTTCATGTGTTGTTGCCTGTGAAG 720
.
T L G Y F R A L E I T K N H E E D T Y H P H F H V L L P V K
AAAAGCTATTTTACTCACAATTACATTAAGCAGTCTGAGTGGACGAGCTTATGGAAAAGGGCGATGAAACTGGACTACACGCCGATTGTT 810
.
K S Y F T H N Y I K Q S E W T S L W K R A M K L D Y T P I V
GATATCCGAAGAGTCAAGGAAGAGCTAAAATGATGCCGAACAGATTGAGAGCGATGTGCGGGAAGCCATGATGGAGCAAAAAGCTGTT 900
.
D I R R V K G R A K I D A E Q I E S D V R E A M M E Q K A V
CTTGAAATCTCTAAATATCCGGTTAAAGATACGGATGTTGTGCGCGGCAATAAGGTGACAGAAGACAATCTGAACACGGTGTTTTATTGG 990
.
L E I S K Y P V K D T D V V R G N K V T E D N L N T V F Y L
GATGATGCGCTTTCTGCCCGCCGGCTTATTGGTTACGGTGGCATCTTGAAGGAAATTCATAAAGAACTAAACCTCGGTGATGCGGAGGAC 1080
.
D D A L S A R R L I G Y G G I L K E I H K E L N L G D A E D
GGCGATCTCGTCAAGATTGAGGAAGAAGATGACGAGGTGGCGAACGAAGCATTGAAGTTATGGCTTACTGGCATCCAGGCATTA AAAAT 1170
.
G D L V K I E E E D D E V A N E A F E V M A Y W H P G I K N
TACATAATCAGATAAAAAGCAGGCGTTGTTCCCTGCTTTTTTTATACTCTAATAGTCAAATCAAGAGTTAATTTTAGATGTAATTGTGAGA 1260
.
Y I I R ***-> <-
ATTAGAGTGGCTGACCAGTATTTGAAACTTCTTGGGCTACTTTCTTAACCTTATATTA AAACTATGTATATATGTGTTGTTTTTCTATT 1350
.
ATTTTGATATTATTTACAAGTATTGAATTTTGCTAGGAGGGAAAGTTTTTATGGTTACCACGATTGGTAAAAGTAAGATGGGTAGGTA 1440
.
Orf1 -> M V T T I G K S K M W V G
TTATTGTTGTATTATCTTTATTATTGGTATCTTTTTCGCCTGCTGTAAAGGCTGATACTAAGGATAAATATTATTCTACAACCTTCTACCC 1530
.
I I V V L S L L L V S F S P A V K A D T K D K Y Y S T T S T
AATCTTCAACAAAAGTTATCGTAAAGCTAATACTAGTGGGGTCTATGTGAAAGTACTTAAAGCTGGACGTTCTCGTGATGTTGCTATTT 1620
.
Q S S T K S Y R K A N T S G V Y V K V L K A G R S R D V A I

```

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

CTGCAGGGCATAATATCATCTTACAACTATGCTGTAGAACGTTATGAAAGAATGTTCCATACAATTATTTGTTTCTAATGGTTCTG 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 ***

TTGTTTTGGTAAATGTGATAAGCGGGTTTTGAAATATAGAGGAGGAAATTTCTTTGACAAAAGAGAAAGTTTTCAAAAAAGAAAGCTCT 1980
 sip15 → M T K E K V F K K K S S

ATTTTAGAGTGGGGCAAGGCTATTGTGATAGCTGTATACTTGCTCTTCTTATTTCGGAATTTTTGTTTGAGCCATATGTGGTAGAAGGG 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

AAGTCTATGGATCCCACTTTGGTTGATTCTGAAAGATTATTTGTAATAAGACTGTGAAGTATACAGGTAATTTTAAACGAGGGGATATA 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

ATAATTTTAAACGGAAAGGAAAAAGCACACATTATGTGAAGCGATTAATTTGGTTTACCTGGAGACACTGTAGAAATGAAGAATGACCAC 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

CTTTTTATTAATGAAATGAAGTTAAGGAACCATATCTTTCTTATAATAAAGAAAATGCTAAGAAAGTGGGTATAAACCTTACAGGAGAT 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

TTTGGACCAATTAAGTTCCAAAAGATAAATATTTGTTATGGGCGATAACCGACAAGAATCAATGGATAGTCGTAATGGGCTTGGACTC 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 ***

AGGCTATAAGGTCTGTTTTTTTCATTTTGTATGAACATTTAGCACACAGATCAAAGTTTTTCATAGTTTGAATGCTTTGATAGCAGCAAAG 2610

GGTATTTCTGATTTTCTGCCGATCTCTCATCGGCGAAAAGTCGGGTCGGCGGACAGCCGACAAGTGGCACGAACTTTTCGATGCGACAGC 2700

GAGAATGAGAGAGCCACCAGCACCAGGTCGCACGTCCAAATTTGCCATGGGCATAATTTGGTGTAGTGCCTTACACCAAAGATAAACT 2790
 RS_a

TTGTGTTACCATAACCCCTATACAGTGGTCTGAATCGGGGTTTTTCTCATGGCAAATTATGCAGTCATCAGGATGGAAAAATACAAAAA 2880
 Mob15 → M A N Y A V I R M E K Y K K

AGATAGATGAATGGAACGCAAAAACACAATCAGCGGGAGTTTCAAAAAGCAAAAATGAAAATATTGATCGGGAGCGGACGCACTTAAA 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

TTATGATCTAGTGAACGAGAAACCGATTAGCTATTCAAAGCGATTTCATGAAAAAATTGAGGGCGAGTCAAACGGAAGGTCGAGCGGA 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

TGCTGTTTTGGTCAGCGAATTTTTGATCACGGCAAGTCTGACTATATGAATGGGCTGAGCGATGAGGAGCAGCGGCTATTTTGA AAC 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

AGCGGTTGATCATTGAAAGAGAAAATACAGCGCTGAAAACATGCTTTATGCTACAGTCCATATGGATGAAGCGACTCCTCATATGCATGT 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

TGGTATTGTACCGATCACAGAGGACGGCCGACTCTGCGAAAGATTTTTTAAATGGCAAATTGAAGATGAAAGCCATTCAAGATGATTT 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

TCATCGGCACATGGTTGAAAACGGTTTTGACCTGGTGC GCGGCAACCAAGCGAAAAGAAGCATGAGAATGTTCCACAGTATAAAAATAAA 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

TCAAGCAGTTATAGAAGTGAAAAAGAATCGCTGACAGCTAAGGCTGAAGAGTTGAAAATGCCGACTATTGACATGAAAAAGCGTGGCT 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

CAAAAAGGATAAAGTCATTGTGCCAGAGCGGAACTCCATGCTTTGTATGCCTATGCGGAGCAGAAAACATAAACGGCAGCCGAGTTGGC 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

GGGCAATTGAAGTCGAAAACGCAGGAAAAGGAGCGCTGGCAGTCTATCGCCCGCAGGAAGCAGATCGGGCGGATGAAAAAGACCAACG 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

GCTTCAGGAACTGCAGAGTAGGATCCATTGAGAAGTTGAAGCGTCCAAAAAGGAAATGCGGCGCAAGCTTGCAAAGGAATTTACGGAAGA 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

ACAGCGTCAGGATCTTCGGCAGGAAGTGAAAGAGGAACTGACGACTTTACGAACGAAAACGAGGAACTGTCAGCTGAAAATAAAGTTTT 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

GATCATTCAAAGAAATAGCGAAGCTGCGGAGAGCCTAAAACATAACAGGAACTTGATAAGAGAAACGGGCAGTATGCTGAGGTTTTGAG 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

TTTCCGCAAGAAGCAGAATCAAACGCTTGAAAAAGTGGCTGGAGAAAACAAGGCGTTAAAAAAGAAAATAAGACTAAAAGAGAGAGT 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

TGCCGTAAGTGGACCAATGGAAGACAAAATGGTTTCAGTGGGCTAAAGAAAATTACCAAAGATGCGGAAATTAGCGGCATCGTTTTTCCG 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

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

CACCCTTTTTTTGTTGGTGGCTGCGATCATTTTTTTGTGTCTCAAGGACTCACGTAATGCAGCGGTGAGTGTCTCGTGCCTTTCTTGC 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

TGTCTTTTTTCGAATCGTTCCATCCGTTCTGCCATGTGGCGTTGAATTCTTCTGTCTTTCATGAATTCGACTAAAGGATTGTCCTGT 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

TTAGTCATGTACATAGATACTCTAAAGTCTTTACGTCATCCTCGGTATAGAGTCGCCATCCTTTTGAATCTTTATTGAACGAGTAGCCT 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

ATATCCATAACGTAATCACCCTCGTAAAAAGGTTTCGCTATAGTGAGGGCAAAGCCTTTATTTAGAGCCATTTAAAAAGAAAGCTGTTCCGT 4860
 M D M ATCACCCTCGT + **Orf2C15**

TCTTTTATCAGGTTGGGACTGTATTTAGTTTATGGGGAGTGCCAAATGAGTGATTTTGAACAGGAATGAGATATGTAAGAGCAACCCTT 4950

GGCTTTGAGGTTTATGTCTGACGGAAGAAGAGGAAAAGCTTTTAGAAAGACGTTTCATGGAGAAATCACAGAGGAAGAATACATACAA 5040

AAAGCGTTGAGCTTCTTGATGTGATGTACCTGCAGGGCTTATAATATTCATAGCTGTTGTTAGAACCTAAACCCTTTTACCGTTTTTGGC 5130

. GGTGAGGGTTCTTTTTTTTAGGCAGTGCATTTGGCGTGAGTCAACGGTAACCGGACCGTAGGGAGGATTAAGGAGTTGACTCGC 5220
. TCAGCGCCACCCGAACCCCTTCAGCACTCAAACAAACCCGTTTGTGACGCCAACCGGCAGGGAGCCCCCGAAGAAGCGGGGTTGGG 5310
. GGGATTGAATGCTGGCATCCAACGGCCGTCCGTTGGTGGGTTTGGGCAAAGCCAAGAAGTGTGCAAGGCTCGTTGAGAATAAAGAATGC 5400
. TTTTCAGGATGCTTAGAATCGTTTCTGAGAGCTTCAAATAAAAAAGATGACCTTTTATAGGGGAAGCTCTTAAAATTGAATGTAGGGGC 5490
. ATTTAAACACGTTTAAAAATAAAAAAGCAGACTCTTTAGAGTCCGCCTTGTTATTTTTAACCCAGTGCTCCATTTTCGGCTGTTTGGAA 5580
. ATCTTTTGAGATGCCGAACCATCCATTTCTTTTGTTCATGAAAAAGTGCTTTTGGATGCTTAAAAGGCTTTTTCGTATAAAAAAAG 5670
. CCCGATTTTGAAAAAAAATCTCCCCCTGCGGAAGAATGGTTTGTATCTTTGGGTTTTAGGTTTTAAAAAAGCCGGCTGTTTTCAG 5760
. CCGTTTTTTTCGATTTTGGCGGACGGAAATCGGGTCTTTTCTTATC 5807

Sequence pTA1040 :

```

TTGATACATATAGAAACAATGAGATTTTCAAAAAGTGC GTTCAATCCAGTATTGTCAAGGGTTTGATGGCAATTTGACAGTGAAAAACCC 90
.
CATCTGCTATTATTAAAGTGTCTGAGCAAAAATAATAGAGTGCTAGAGAAGCTAGCGTCAGAAAGGGTTTTTCACTTTGTATTTCATCTGAAT 180
.                               Rep →                               M Y S S E
CGAATTATAGCATCTTTTCAGGACAAAACCGCAACAGGTAAAAAGCGGGATTGGAAGGGGAAAAAGGTGAGATCGTTACTTATGGCGGCTC 270
S N Y S I F Q D K T A T G K K R D W K G K K V R S L L M A A
.
ATTATGAAGGACTTGAAAAGCGTACCGGTGCGCCTTACTACGGAAAAAGGCAGAGAAAAGTATGTGATTGTGCGGAATGTCTTGCGTTTA 360
H Y E G L E K R T G A P Y Y G K K A E K V C D C A E C L A F
.
AGAGAGATTTAGAAACGGGCAGACTACGATTATATCAAGCCTATTTTTGCAAAGTGAAGTTATGCCCTATGTGTGCGTGGCGAAGGTCAAT 450
K R D L E T G R L R L Y Q A Y F C K V R L C P M C A W R R S
.
TGAAAATTGCTTATCACAACAAATTGATCGTTGAAGAAGCCAATCGGCAATATAAGCCC GCATGGATATTTCTCACGTTGACAGTAAAAA 540
L K I A Y H N K L I V E E A N R Q Y K P A W I F L T L T V K
.
ACGTTGAGGGCGATGATCTCAAACAGACCATTTTCGGACATGATGAAGGGTTTTAATAAATTGATGAAATATAAGAAAAGTAGATTCTGCGG 630
N V E G D D L K Q T I S D M M K G F N K L M K Y K K V D S A
.
TTCGTTGATATTTTTCAGAGCTTTAGAGATCACGAAGAATCATGAAGAAAATACATATCACCCGCATTTTCATGTGTTGATTCTGTTTCGTA 720
V R G Y F R A L E I T K N H E E N T Y H P H F H V L I P V R
.
CTGTTTACTTTGGTAAAACTATATTAATCAAAAAGAATGGACAAGCTTTTGGAAAGCGTGCAATGAAATGGGATTACACACCGATTGTTTC 810
T G Y F G K N Y I N Q K E W T S F W K R A M K W D Y T P I V
.
ATGTGCAGCGTGTCAAAGGAAAAAGGGAATCGATGCAGAAGCGATTGAACAAGAAGTGCCTAAAGAGATGGAAGAGSSCAAGGCTATTC 900
H V Q R V K G K K G I D A E A I E Q E V R K E M E E L K A I
.
TTGAAATCTCGAAGTATCCTGTTAAGGATACGGATGTCATACGTGGTAATGAAGTGACCGAGGAAAATTTGGATACGGTCTATTACTTGG 990
L E I S K Y P V K D T D V I R G N E V T E E N L D T V Y Y L
.
ATGATGCATTAAGTGCCAGGCGTTTTGATCGGGTACGGCGGGATCTTGAAAGAGATTGATAAGGAACTGAATTTGACGGATGCTGAAGACG 1080
D D A L S A R R L I G Y G G I L K E I H K E L N L T D A E D
.
GAGATCTGATCCGTATTGAAGAAGATGATGATGATGTAGCCAATGAGACGTTTGAAGTTATGGCGCATTGGCATGTTGGTATTAAGAATT 1170
G D L I R I E E D D D D V A N E T F E V M A H W H V G I K N
.
ACATAATCACATAGAAAAGCAGGCTAGAAAATAGCCTGCTTTTTCTTGTATGGATTAGTTGATTTTTGCTTTTTAAACTCTATAATCTCCA 1260
Y I I T ***→←*** N I K A K L V R Y D G
.
CTGCCGTAAAGTCCTACGTCAACACGAGAAGAGTATTCAACAGGTGTGATTTTCAGTGCTTTGAAAGTTCCTCTGCTACGAATTTCCCAA 1350
S G Y L G V D V R S S Y E V P T Y K L A K F T G R S R I E W
.
ACGCTAGGGTCTAGTTTGTGTTGTTTTGTTGAGTGTGATTTTGTAGTAAAAGACGCCAGTTAAATAAGCATCATAACTGCTAGATACTAAT 1440
V S P D L K H Q K T S H S K S Y F V G T L Y A D Y S S S V L
.
TTGCCTTTTTTATAAGTAAATTTACCACTTACTTCATGAGCAAAACCAAAAACAGGAGTTGTATTTGAAACCTTTGCTTTGAATTTCTTT 1530
K G K E Y T F K G S V E H A F G F V P T T N S V K A K F K K
.
TGTTGACTGCTTTTTGTTGAAATGTCTTCAACGTTTCATAGACAATTCGAGTTTGTCTCCTTGTACGTTTAGAGGTTCTTGTACTACAACG 1620
Q Q V A K T S I D E V N M S L E L K D G Q V N L P E Q V V V

```

TTATCTAATCCTTTTTTAAGTTTGGCGTTTGCCTTATCAGTTAGTTGATCTAGTTTGTGTTGATCGTTTCAACTTCTTTTCAGATGTGCC 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

.
 ACTTGTTCACCTTTGTGTGCTTTTGCATCTGGGGCGTATATTCCCTAAAGATAAAGTGATCGCAGCAAGTCCATTAAAACAGTGGAGCTG 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

.
 GTCAGAACTAATTACTAAAGATCGCCCTATTAATCATTTCGATCATCATAGTGATCTCACTTATTGAAGCTATAATTGGTATCGTTTTGG 2070

.
 CTTACTATAAACACCCTTTTTGTAGAATTGATCTATAGAAATCACTCCTTTTTTTGGGTATATTTTCATCCTAACATTGAATGATTTAG 2160

.
 CGTTGGGAATTCATTGGAAAATTTCTTTTTTGGTTGCGGTGTCCGCAATCATTTTTTTTATTTTTTCAGTGCTTCGTTTTAGTGGATTTTT 2250

.
 TTGTGATAAAATCCGTTCACTGTGGAGGTGAATGTGATGATGATGTAATTAATGTGATGGATTTAGGAGCATTCATTTTTTGTACTTTG 2340

.
 ATGTTTCATTGCATGTGCAGTTTTCAGAAACAGAAGAAGCCGGCTATTGAGTCGGCTTTTTCTTTACACGCTTGACGAAAATCGAGAGATT 2430

.
 TAGCAAGCACACATCTGTGTGTATAAGTGCCTATTTTCATAGGTGTTCTATTTCCAAAATGCCACCACCTTTTTTTGTTTCATTGGATT 2520

.
 TTTTCAAAGCGAGCATGAGATTTTCATCTCGCTTTTTTTTGCTGTTTTTCAGATTGTTTCGATTCGTTTTTAGGATGGCCCGATTGAATTCC 2610

.
 TCTTGTCTTTAAATGATCGGCGAGCGGTATCCGATATGGACAAGTTCGACCAATAACAAGGTGCTACAAGTTGGCTACTTCTTCTAGTC 2700

.
 TTATTAGGAGCCTATTTTTGTTTTTTCTTAACTTTACCTTAAAAAACATTAATAGATAATTTTATATAATTTAATGGTATTATTTTCT 2790

.
 TTGTAATATATTTCTGAAAGGATGGAAGAAGAAGTATGAAGTTGAAATTTGTTACAATGATTCTTTAGCTTTTGTATTGATTTTATCTC 2880
Orf1 → M K L K F V T M I S L A F V L I L S

.
 CTTTTCTATGGTACAAGCTGATACGAAAGATAGAAGTATAAGTACTTATAAGTATCATGGTTGTACCACTGTTGCCAAAAAGAAAATA 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

.
 TCAGTGGTGTATTATTGAAAGTATCAGGTCATGGTTCTGGGCGATTTTACTCGTTTCTGTTTTTGGCTGATGCTAATAAAGGAAAGGGTA 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

.
 AACGAAATGGGTGAACGTTTCCGGACAAGATATGGCTAGTTTAGGAAAATATGTAACAAAAGGTAATAAATATCGTCTTACGAATTATG 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

.
 CAGATAGTAGATAAATAGTAAGTTTCTCTGTCCTGTAAGATAATTCTTTACAGGATTTTTTATTTTTTGTTTAAAATCTAAGAAATG 3330
 P D S R ***

.
 TTAAAGTTGAGTTAAAAATGCAAGCGGGAGGAGCGTAAGTGTGTTGATAAAGAAAAAGAAAAAGTCTAATATCATTGATTGGATAA 3420
Sip40 → M F D K E K R K K S N I I D W I

.
 AAGCTATTTAATAGCTCTTATTCTGTGTTTTCTTGTTCGTAATTTTTGTTTGAACCATACATTGTACAAGCGGAATCTATGAAACCGA 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

CATTATTTAATTCTGAAAGATTATTTGTTAATAAATTTGTTAAATATACAGGTGACTTTAAAGAGGGGATATGTTGTTTTAAATGGTG 3600
 T L F N S E R L F V N K F V K Y T G D F K R G D I V V L N G

AGGAAAAAAGACCCATTATGTGAAGAGATTAATTGGCCTTCCTGGTGATACAATTGAAATGAAGAATGATAATCTTTTTGTAAATGGAA 3690
 E E K K T H Y V K R L I G L P G D T I E M K N D N L F V N G

AAAGGTTTAAATGAGGAATACTTAAAAGAAAATAAAAAAGACGCACATGACAGTGATCTCAATTGACTGGAGATTTGGTCCAATTAAGG 3780
 K R F N E E Y L K E N K K D A H D S D L N L T G D F G P I K

TACCTAAAGATAAATATTTTGTTCATGGGAGATAATCGTCAAAATTCATGGATAGTCGAAATGGTCTCGGATTATTTAATAAAAAGGACA 3870
 V P K D K Y F V M G D N R Q N S M D S R N G L G L F N K K D

TTGTTGGAGTGAAGAGTTAGTTTTTTTTCTCTGTATACGTCATGCTAAATAATATTTTTATGCGGAAGTACCTATTGCAAAAA 3960
 I V G V E E L V F F P L D R I R H A K ***

AGCAGGCATGTATTAGCCTGCTTTTTTGTATGGTCTATATCTTTACTATTGTAAAGTCTTATGTCAGCGTGTTTTTTACAAGCTTGACT 4050
AGCAGGCATGTATTAGCCTGCTTTTTTGTATGGTCTATATCTTTACTATTGTAAAGTCTTATGTCAGCGTGTTTTTTACAAGCTTGACT

GAAATCGAAAGATTTTAGTAATTACAAAAGGGTGTGTATAAATGAGACATTCCAATGGATATCTATTTCCAAAATGGCCACCACCTTTTT 4140
 *** K W F Q W W K K

TTGTTGAGCAGCAGCAATCATTTTTTTGTCTTCAATGGATTCTTTCAAAGCGAGCATGAGATTTTCATCTCGCTTTTTTTGTTGTTCTTC 4230
 Q Q A A A I M K K A E I S E K L A L M L N E D R K K Q Q E E

AAATGTTTCGAGTCGTTTTAGGATGGCCCGATTGAATCTTCTTGTCTTTGAATGAGATCGGCGACATTTACATCTTGTAGCGGTGTAGC 4320
 F Q E L R K L I A R N F E E Q R Q I L D A V N V D Q L P T A

GGTATCCGATATGGACAAGTTTGACCGATAACAAGTTCGCTACAAGTTTGGCTACTTCTTCTAGTGATTTACCACTCAATTTAGAGTGTGT 4410
 T D S I S L N S R Y L T A V L K A V E E L S K G S L K S H T

GTATATACTCCATAGCGGATAAATCGGACTCATTATACTGACCCATCCCTTAGTGTCTCTGGCGAAATATAGCCGTTTTTCTCAAG 4500
 Y I Y E M A S L D S E N Y Q R W G K T D R A F N Y G N K E L

CATGGCGCGTATTTGCGTATCGTGACGCTGGAGACATTGAGCTTGTGTTGCAATGTCTTGTGAAGAAAGGTAGGTTTCGGGTTTCATAGC 4590
 M A A Y K R I T V S S V N L K N A I D Q S S L Y T E P K M

GGGTTCACCTCTTTATGATTCATTGCTATGATATACGCTACGCCCTCCTAGATCAAAGGCAAATGTTTTGTTTTGCATAGCGGGTT 4680
GGGTTCACCTCTTTATGATTCATTGCTATGATATACGCTACGCCCTCCTAGATCAAAGGCAAATGTTTTGTTTTGCATAGCGGGTT
 Orf2C40

GAAAGAGAAAAATTTTGATAAGTGTGCTGGAGACATTTGCGAAAAGGTGAAAAGAGCATAAGGAAAACCTGCAACCCTTTATGTTTGA 4770

AATCCCAGTGTGATTTGTTTTGTTGGCGGTTGCGATTTTTGAAAGGAATGTGCCTAAAAATGAAATCTGACACACCTTTTTTTAGGAGAA 4860

AAATGTGGTGATGTGGCAGAAACGGAAAACTGCCACACCATAAAAAGGAAAAGTGAGAAGAATTGAGTAACGTTGTGTTTGAAGAGGAG 4950

CGCACCTGCACAAAGAAAGGGAGCGTTTTTGCAAAAAAATCAGTCGTAACAAGGGGTCAATAATGATGAATTAAGTTTCGTGTTTGT 5040

TGTGTTGTTTCGGAAAAGTGTGCACTCAAAGGAGCGTTACTGTTTAAAAAATGGAGTGAAAAACATATGGCATGAGTCAACGGTAAGCG 5130

GAACGAAGTGAGCATTAAAGGAGTTGACTCATTACGACACCCAAAAGCCTGTTAATCAAACAAACCCCGTTTGTGCGCCTGCCGAAGG 5220

GGCATCCCCGAAGATCGGGGGTAGGGGTTTTGAATATTTGCCATCCAACGTGCTAGCGTTGGCGGGTTTTGGGCAGAGCCCAAGAAAGGC 5310

GGTCGGATGATGAAAATGACAAGAGACGCTTTTGGAAAAGGCATGAGATATGCGAAAGCGACACATGCTATTGAAGGAATTTATCTTACA 5400

GCTGACGAGGAAGAATTATTATGGCAGCATGCATCCGGTCAGATCACAGATGAGGAATTTGAGAGAAAGGCTCTTGAATTGGCTTATAAA 5490

GTGATAACAATAGAGTTCTAGGCGAAGGTGCGTAGCGGAAGACGACGACAGTCGTCGAGTGGCAACGAAAAGCGAAGCACCCTAGCCAGCT 5580

TGAACCCCTCACTCTAACAGGGGGTAGTAATGTAAGATGAAAATAAAACGTATAAAAAAATATGTCGTGCATTTCAAAAATGCCGTTG 5670

AGAGCGTTCAGATTGGATTCTGAGCGCTTTTATTTTGTAGGTAATTGAAATAGGTGAAAGCGTCAAAAACGCTTATAAGGGCTTTAA 5760

ATGCGTTTTTTCGTTTTGTCTGCTGTCATGATTCCTAAAATTTGATGGTGAGATACCGAGCAAAATTCAAAAACGGTTGCGAGATACCGA 5850
#####

AAATTTTCAAAAATGGAACCTATAATGGATTTTATAACGTACAAATTATGCTAGCATTTAACATAAAAAGGGATGTCGAAGGACGAGATG 5940
ComA-box ##### **Rap40** →

TGATGGAAAGTATTCTGCTGCTGTGGTCCGACAGAAAGATTAATGAATGGTATCGGCATATCAAAAATTTGAATGTAACAGATGCGGAAA 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

AGCAGATGCTTGAGTATGTAATCCTTCTCAAAATCAAGTGAACAAAGCTGATTATTTGCGAGCTGTTGAAGGTGAAGGTAAAAAGTTGA 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

CGGGTATCATGGAATACTACGTTCACTTTTTTAAAGGGATGTATGAATTTCTCTAAGGGTGAGTACATAAAAGCCATTGCTTTTTTATAGAC 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

AATGTCGTTTTGTCTATTGCAGGAAATTTATGATGATTTATGTACACATGAAAAAGCGATACTCATTTAAAAGAAGCGGAAGAGTCGGCCA 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

GGAATTTAGGCAACAAGTATATGCTTGCTAAAGCGTTGCTTAACTCGGTGATTCTTATAAAAAATATGGGTGATCAGGACAGAGCGCTTT 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

AGAAAAAGGAGCATGAATCGGGAAAACACTATTTTGAGTTGGGTGTTGAAAGTCTGAAAGTTCAATGATTGGTTTTTTCACGAATTTGT 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

TAAATGTACTGAAGCTTCTTTATTTAGAGGGGGCTCAACCGACTGCTGTTGCGGATGCTTTAGAAAGAACTTGACGATATTCGAGGCTATC 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

CATATTTAGAAGAAGCTGGCGCTTGTGCTGCTGAATTTTATACTGAAATTTGGGCATATGGATGAATCTGTATATTTTTATAACAAAATGA 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

TCTATGCTCAAAAACAAATTCGAGGGGCGAGTCTTGTATGAAATTTAATGCCTTACTGCTTTTAAATGTTTGTGCTAGTTTATTAATC 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

.
 GTTAGTGGTTCTTCTTTGTAATCCAACAAGATTCAAATGTGTCAGTTGCTTCAAGAAAGGCTACATGATGATGAGTTTTATTCGATTTG 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
 .
 TGTTCATGCTCTTCTGTTTGTCTCTTTTCAATCCCACCATTTTGTGGGCGTGTGATGGCAGCTGTATGTTGAAAGGCGGGTTTAGTC 7470
 .
 AGGTTTTGAGTGAAAAGCTACAGAAAAGGTGTTGCTGATCGTAAAGGCTACAGAAACGGTGTGGTGTTTTTTCTTGCCAAAATAAACCCCT 7560
 .
 ATGATTGCAAGGTTTTGATATGTATTTGAGTGATTAAAAGCTATAAAAACGGTGTAAAAGCTAATGAAAAGCTACAAAAAAGTGTGAG 7650
 .
 GCTATAAAAACGGGGGACTCAATTTTGTAGTAGGGGCAAAAAATCAAAAAAATCTCCCCTACGGGGAAATTTGGTTTAGTCTTTG 7740
 .
 GTCTTTTAGGTTTTGAGTTTTGGTTTTTAAAAGCTGGCGGTTCCGCCAGCTTTTGGGCGATTTTGACGGAGTCAGAATCGTGTCTTT 7830
 .
 TCTTATC 7837

Sequence pTA1060 :

```

TTGATACTATATAGAAACAACATCATTTTTTCAAAAATTAGGTCAAAGCCTTGTGTATCAAGGGTTTGATGGTTCTTTGACAGGTAAAAACT 90
CCTTCTGCTATTATTAAGGTGTGCGAATCAAATAATAGAATGCTAGAGAAGCTCAGAAAGGAGTTTTTTTTGTTGATTATTTCATCTGA 180
Rep - M I Y S S E
AAATGATTATAGCATCCTCGAAGATAAAAACCGCAACAGGTAAAAAGCGGGATTGGAAGGGGAAAAAGAGACGGACGAACCTCATGGCGGA 270
N D Y S I L E D K T A T G K K R D W K G K K R R T N L M A E
GCATTACGAAGCGTTAGAGAGTAAGATTGGGGCACCTTACTATGGCAAAAAGGCTGAAAACTAATTAGTTGTGCAGAGTATCTTTTCGTT 360
H Y E A L E S K I G A P Y Y G K K A E K L I S C A E Y L S F
TAAGAGAGACCCGGAGACGGGCAAGTTAAAACGTATCAAGCCCATTTTTTGTAAAGTGAGGTTATGTCCGATGTGTGCGTGGCGCAGGTC 450
K R D P E T G K L K L Y Q A H F C K V R L C P M C A W R R S
GTTAAAAATTGCTTATCACAATAAGTTGATCGTAGAGGAAGCCAATAGACAGTACGGCTGCGGATGGATTTTTCTCACGCTGACGATTCTG 540
L K I A Y H N K L I V E E A N R Q Y G C G W I F L T L T I R
AAATGTAAAGGGAGAACGGCTGAAGCCACAAATTTCTGCGATGATGGAAGGCTTTAGGAAACTGTTCCAGTACAAAAAGTAAAAACTTC 630
N V K G E R L K P Q I S A M M E G F R K L F Q Y K K V K T S
GGTTCTTGATTTTTTCAGAGCTTTAGAGATTACCAAAAAATCATGAAGAAGATACATATCATCCTCATTTCATGTGTTGATACCAAGTAAG 720
V L G F F R A L E I T K N H E E D T Y H P H F H V L I P V R
GAAAAATTATTTTGGGAAAAACTATATTAAGCAGGCGGAGTGGACGAGCCTTTGGAAAAAGGCGATGAAATTGGATTACACTCCAATTGT 810
K N Y F G K N Y I K Q A E W T S L W K K A M K L D Y T P I V
CGATATTCGTCGAGTGAAGGTAAGCTAAGATTGACGCTGAACAGATTGAAAACGATGTGCGGAACGCAATGATGGAGCAAAAAGCTGT 900
D I R R V K G K A K I D A E Q I E N D V R N A M M E Q K A V
TCTCGAAATCTCTAAATATCCGGTTAAGGATACGGATGTTGTGCGCGGTAATAAGGTGACTGAAGACAATCTGAACACGGTGCTTTACTT 990
L E I S K Y P V K D T D V V R G N K V T E D N L N T V L Y L
GGATGATCGCTTGGCAGCTCGAAGGTTAATTGGATACGGTGGCATTTTGAAGGAGATACATAAAGAGCTGAATCTTGGTGATGCGGAGGA 1080
D D A L A A R R L I G Y G G I L K E I H K E L N L G D A E D
CGGCGATCTGGTCAAGATTGAGGAAGAAGATGACGAGGTTGCAAATGGTGCATTGAGGTTATGGCTTATTGGCATCCTGGCATTAAAAA 1170
G D L V K I E E E D D E V A N G A F E V M A Y W H P G I K N
TTACATAATCAAATAAAAAAGCAGACCTTTAGAAGGCCTGCTTTTTTAACTAACCCATTTGTATTGTGTTGAAATATGTTTTGTTTCGTT 1260
Y I I K ***=====→ ←=====*** S V W K Y Q T S I H K T R K
TTTTATCTTTGTAGCTTTTTATGATTAATTGTCTATGATAACCTAGGGAATCTTTACATTGTCTGTGTTGAATAGTTACATAAACTGCTT 1350
K D K Y S K I I L Q R H Y G L S D K C Q R H Q I T V Y V A K
TGCTCTTAGGATAGTTTAAATATACTGTTGTCATTGCTCCGAAAACAAAGCTTTAATGGTCTTTAGCGATATAGTTTCTATAACTC 1440
S K P Y N L Y V T T M A G F V F A K L P G K A I Y N G I V G
CAACTACGCCACCTGCAATTCCTGGAATAATTTTTATTTCCCAATCTGTATATTCACATTATCATACTTTTTGGTAGCGACAATTTTAC 1530
V V G G A I G P I I K N E W D T Y K V N D Y K K T A V I K C
ATGAATTTAAATAGTCGCCAGGAAGTTGTTGGATAGTTATATTTGATTAGTTTGGTTTGAAGAACAATTTGTAATATCATTTTGAGAAA 1620
S N L Y D G P L Q Q I T I N Q N T Q N S S C N T I D N Q S F

```

AAGTGGTTGGAAC TATTGGTGTGTTATATCTTTACTCTGTTTCATGTTATTTTGTGAGCAGCAAGTATACTATTTCGTAGGAATCAATA 1710
 T T P V I P T T I D K S Q E N N N Q Q A A L I S N T P I L L

AAATAAGACATTACTAAAGCTAATAACAATTTTTATGCTACCTTCAATTAATAAACATCCCTTCTTTGGTTATATGGGCCGTATATA 1800
 V L S M V L A L V I K I S G K M CCCTTCTTTGGTTATATGGGCCGTATATA ← **Orf7C60**

ATATACCTCTTATTGGGATTTTTTTGTTAAATTTTGAATTGAATTAAGAGAAAAATTAAGAAAGGTATGAATTCATAGTGAATTACTAA 1890

TGAGTTTTGTGTTTGGAGCTTTAGGAGCATTATTAATGGTAAGTGTGAGAGAAAACTTTTCTTTACTGTGACATTAATAGTTTTAGTTG 1980

TTGTTATCATCGTTGCTTTTATTCTTGATGAGTTATGGAAAAGGCATAAAAAAAGAGAGAAGAAGATTAATTTTCTTCTCTTTTTTTT 2070

TATGCCTTGTCTCTCTATGGGGAAGATGTTTACAAAATGAGAAAGTAGATGCACGGTCCGCAGCATACTTTTCGGAATTGAAAATTTG 2160

ACACACAGAAACAAATGTTTCATGCGGTCAAATGCTGCGAGAGCAGCATGGGTTTTTCCGATTTTCTGCCGATCTCTCATCGGTGAAAAAT 2250
CGGTTCTGCGGACAGCCGAC → CAAGTGGCAGAACTTTCGATGCGATAGCGAGAATGAGAGAGCCACCAGCACCCGAGGTCGCACGTCCAAA ←

TTTGCCATGGGCATAATTTGGTGTAGTGC GTTACACCAAAGATAAACTTTGTGTTACCATAACCCCTAGATAGTGGTCTGAATCGGGGGT 2430
 → CAAGTGGCAGAACTTTCGATGCGATAGCGAGAATGAGAGAGCCACCAGCACCCGAGGTCGCACGTCCAAA ← **RS_A**

TTTTGTTCATGGCAAATTTATGCGGTTCATAAGGATGGAAAAATACAAAAAGATAGATTGAATGGAAACGCAAAAACAATCAGCGGGAGTT 2520
 M A N Y A V I R M E K Y K K D R L N G T Q K H N Q R E F

TCAAAAAGCAAAAATGAAAATATAGATCGGGAGCGGACGTACTTAAATTATGATCTGGTGAACGAGAGACCGATTAGTTATTCAAAAGC 2610
 Q K S K N E N I D R E R T Y L N Y D L V N E R P I S Y S K A

GGTTTCATGAAAAATGAGGGGCGAGTCAAACGGAAGGTCCGAGCGGATGCTGTTTGGTTCAGCGAATTTTGGATCACGGCAAGTCCTGA 2700
 V H E K I E G R V K R K V R A D A V L V S E F L I T A S P D

CTATATGAATGGGCTGAGCGATGAGGAGCAGCGGCTATTTTGAACAGCGGTTGATCATTGAAAGAGAAATACAGCGCTGAAAATAT 2790
 Y M N G L S D E E Q R R Y F E T A V D H L K E K Y S A E N M

GCTTTATGCCACAGTCCATATGGATGAAGCGACTCCCATATGCATGTTGGTATTGTACCGATCACGGAGGACGGCAGGCTCTCTGCGAA 2880
 L Y A T V H M D E A T P H M H V G I V P I T E D G R L S A K

AGATTTTTTTTAAATGGCAAATGAAGATGAAAGCCATTCAAGATGATTTTACAGGCACATGGTTGAAAACGGTTTTGACCTGGTGC GCGG 2970
 D F F N G K L K M K A I Q D D F H R H M V E N G F D L V R G

CGAACCGAGCGAAAAGAAGCATGAGAATGTTCCACAGTATAAAATCAATCAGCGACAAGCGGAGTTAGAGCGGCTTAATGCTGAAATTGC 3060
 E P S E K K H E N V H Q Y K I N Q R Q A E L E R L N A E I A

TTTAAAGGAAAAGCAGAGAGAGAACTGGAAAAGCAAACAAGCTGTTCAAGCAGTTATAGAAGTGA AAAAAGAATCGCTGACAGCTAA 3150
 L K E K Q R E E L E K Q N K A V Q A V I E V K K E S L T A K

GGTTGAAGAGTTGAAAATGCCGACTATTGAACATGAAAAGCGTGGCTCAAAAAGGATAAAGTCATTGTGCCGAGCGGGAAGTCCATGA 3240
 V E E L K M P T I E H E K A W L K K D K V I V P E R E L H D

TTTGTACGCTTATGCGGAGCAGAAAATAAACCGGCATCCGAGTTGGCGGGCAATTGAAGTTGGAAAACGAGGAAAAGGAGTGCTGGCA 3330
 L Y A Y A E Q K T K T A S E L A G Q L K L E T Q E K E C W Q

GTCTATCGCCCGCAGGAAGCAGATCGGGCGGATGAAAAGACCGACGGCTTCAGGAAGTGCAGAGTAGGATCCATTCAGAAGTTGAAGC 3420
 S I A R Q E A D R A D E K D R R L Q E L Q S R I H S E V E A

GTCCAAAAGGAAATGCGGCGCAAGCTTGCAAAGGAATTTACGGAAGAGCAGCGTGAGGATCTGCGGCAGGAAGTAAAAGAGGAGCTGAC 3510

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

GGCTTTACGAACGGAAAATAAGGAACTGTCAGCTGAAAATAAAGTTTTGATCATTCAAAAGAAATAGCGAAGCTGCGGAGAGCCTAAACCT 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

AAAACAGGAACTTGATAAGAAAAATGAAAAGTATGCTGAGCTTTTTAGTTTCGCCAAGAAACAGAATCAAACACTTGAAAAAGTACTGG 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

AGAAAATCAGGTGCTAAAAACAGAAAATAGGACATTAAAAGAGAGAGTTGCCGTACTGGAACAATGGAAGACAAAATGGTTCAGTGGGC 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

TGAATTAGAGCGGTGAAAACGGGCAATTATTGCCACTTAATTGCCTTTCCAAAATTCACCACACTTTTTTTGTTTGGTGGCTGCGATCA 3960
E L E R *** $\xrightarrow{\hspace{1cm}}$ $\xleftarrow{\hspace{1cm}}$ *** 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

CCATGCGGCGGTTGAATTTCTTCTGTGCTTTCATGAATTCACCAAAGGATTGTCTTGTAGCGATGTAGCGGTATCCGATATGGTCAGTT 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

TTGAACGATATAAGCTTGCTATATGCTTTACCGTTTCGTGAGTGAGTGTCCATTGATTTTAGTCATTGTACATAGATACTCTAAAGTCT 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

TTACGTATCCTCGGTATAGAGTCGCCATCCTTTTGAATCTTTATGAATGAATAGCCTTGTTCTTCAAGCATACTGGCATATTTGCGGA 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

CAGTGACTGGCTCTATAACCAAGGTGTTTTCGAACGTCCTTTGACGATAATTTGACTCCCATATCCATCACGTATCACCTGGTAAAAAAGG 4410
T V P E I G L H K A V D K S S L K V G M D M TATCACCTGGTAAAAAAGG + **Orf2C60**

TTTCGCTATAGCGATGCTAAAGCCCTTTAAGTGCTGCACCCTTTATTTGAGTTGTTTTGTAAAAGGATGCACTGTTACTCAAAAAATAAAA 4500

AAAGCATTGTATCTTTGAATTTGAGTTTGTTTTAGAAACATCAGGTGCGCAACAGGTGAATTTGTCGCTTTTTTTATGTTTTCTAAAAA 4590

AAATAAAAACTACGTTTCGACCGAAAAACATTTTCGGGATTGGTTTTTCGTGTTTACTCGTCTATGTTATTCCGTATTGTTATCCGTATAGT 4680

ATTCGTATAGTTTGAACGAAATAAAAAAGCAAGTAAACCACGAAAAAAGCACACTGTAATGATAAAAATATCTGCGTTTTAAGCGCGTTTTGA 4770

GTATGTGTTTTTTTTTCGTTTATGATGAATCTATCATAcataAAGGAGATGTTTTAATGAGTGGTTTTAGACTTTCTTTACCTTCATCGGAT 4860
Orf4 - M S G F R L S L P S S D

TCTGATATTAATGATATGCTTTGAAGCATTACTATAAGAGTTGGAGAACGGTTAATCAAACAATAAGTCTTCTTTTATAGCTTTAGAT 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

AATGCTTTTCGAGATACATATTTGAAAGATTGGAACCTGGACCTTTAAAATTGTATCTCTATTTTTGTTTCGCTGCTAATAATGATCAT 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

GGTCATAGTTGGCATGGTATTCAGAAAATCGCTGATTATTTTCGATTGTCAGACTCGGACTATTGATAAATTGGATAAAAAGTCTTGTTGAT 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

AAGGATTTGGTTTTATCGAACAAAGAAAGGTCATAGGTCACATACTACCTATCTTTTGCCTTTTCAGTGACACTATTCTTATACAAAAAGCT 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

TCGAAGAAGTACATTAAAGATGATCAGGACTTGCTTAATGGTCTATTA AAAAGAATAAGAGAACTTGACTTTATTTATGGGGAAATTATT 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

AAGGTCCATCATCTTTTCAATGGTCTGTTTCAAAGAAAAAACTGTTGATGGAGATAAAAGCGATCAATATCTTTTGGTTATTACTAAA 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

CGAAAAACGGTGTCTTAATTGGACATGTTTATGTGTTGAGAAAATCTGATCATTTGAGTGTGAATGAATTACAGATTGAAGAACCATCT 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

GTATTTGATAGTCTTTTTATCTTAATGAATCAAGAATAATTGGTCTTGCTTTGCCTCATTATCCTAAAATTTTTTCTAAGTCTGGATTA 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

AATGATTTATTGGGTTTAAATAAGGGACTTAGTGATTATTGAAGATTGGCAGTTGGAAGATCGATCAAACTTATTATTGGTGAAAAGCGG 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

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

TTATGAATTTTGAACAAATAGGGGTATTAAATGATACTGTGATTAATCAGTTAAATTTGACTATTAGCCCACTCCAATTTTTATTGGAA 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

AATCTAATATTGATCATATGATTAAGAGCATCCCGAGGATTTTAAACAATATGGTTGTTACTTAGAAGAAATAATTTCTTCCGGATT 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

ATGTTTCTCTTCATCCGAACAATGGATCTATTCAATATATTAAGTTTTTTTTTGACGAAGATAATGAAAAACATGTTTTAGTTGCAGTCC 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

GAGCAACAAAGAAAGGAGTATTTTTTGCGCGTCTTTGTTTGTGATGAGCGATTTAAAAGTAAAGAGCTATACAAATAAGGGGATGCTGA 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

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

CCCGCATCTCCAACATAGGGTGATGGCTGCCTGTTCATCCTCAGAAATCCTTCTTATTAGTTTATTCAACTTTCTATATTTATTATACA 6300

TGTGAAAATTTTATGTTTCAAGGGTGGTTTTATTAAACGAGTTTTTATGGCGTGAGTCAACGGTAACCGGACCGTAGGGAGGATTAAGG 6390

AGTTGACTCGCTCAGTGCTACCCGAACCCTTTTCAGCACTCAAACAACTCGTTTGTGTTGACGCCAACCGGCGAGGGAGCCCCCGGAGAT 6480

CGGGGGGTTGGGGGATGAATGTTGGAATCCAACGGCCATCCGTTGGTGGGTTTGGGCAAAGCCCAAGTTCTACTTTAACTCTTGAT 6570

GCAATTTGTGTTTTTGTGCGAGCAAGACAATTCCTTCAGGGCTTCGTTCTGGGAAGTGTGTGTGAGATTTAGCTTTTGATCAGAAAGCA 6660

ACTCCTGCATTTTCTTTTGTGCTTCAGCGAGCGTTTCGATTTTCAATTTTTATTCCTCCGGCGTTTTTGTGATAATACTAGCATGAGGTG 6750

CGCTACCAGTAGCCACGTAAAAATAGATAGCTGGATAAGAGGGACGGATGCTTTTTAGCAATCCGCTTTTTTGTGTTGAGAGGTTAA 6840

TAGATAATGATAAATAGCTCTGAGGATTTCTCAGATTCGTTTCTGAGAGTTTTAATTTTTAAATAGAGTAAAAGGTCACAGGGGAATTTA 6930

GGGACGCTGTATTGAGCTTTAAATGGATTTTTTGATCTTTGTTTTGAGAGAAAATTTTTGTTTTGAAAAATGAAACGAAAGGGCTTTT 7020

TTTCATGAGTAAGGATGCTTTTGTGTA AAAATAATCCTAATCGATTTTGTTAATTTAAATATTTTCGCTTTTCGGCTTACCGAAAATTTAA 7110
ComA-box

GGTTCGGCGTCCGCAACCCCGATTTTACAGCTTTTCTGTGCTACCATTTACAAAAAAGTATTAAATTAGGGGAGCAGTTACTCGGAAT 7200
Rap60 → M

GGAAGTTATTCATCTGCGAATGTAGGGCTAAAGATAAATGAGTGGCACATGTACATAAATCGATTTCGATGTTAGAAAACGCTGAAAAGGT 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

AAAAGCGGAAGTTGAACAGATGATTGATGTTATGGAAGAAAATCAAGATTTGCTCATATATTATCAGCTGATCGAGTTTCAGACACCGTAT 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

GATGCTTGATTATTTATTTCCGGCTGATCCCTCTAAAAAATCAATAAAGGTGATTATCTAAAAAATGAGGGGAAAGGCCAAAAAAT 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

GTCTGGCATGATGGAATACTACACCTATTTTAAAGGGATGTACGATTTTCTATTGGTGAGTACATAAGACCATTTCAGTCATACAA 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

AAAGGCTGAGAAAAGCTTTTCGAGGGTTTGGATGAAATAGAGAAAGCTGAGTTTATTTTAAATGGCTGAAGTCTTTTACCATATGAA 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

ACAAACGCATGTATCTATGTATTATGTGCAACAGGCATATGACATTTATAAAAAGCACGATACATATAAAATTCATCGTATCAAGTGTC 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

TTTTGTTATTCGCGCAAAATATGATGAATTAAGGACTCATGACAGAGCGCTTTTCTCATATTGAAAAGCCTTAATTTTGCTGGATAGTAT 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

GGAGGACCATTCCCCTTTAAAGGCAAAAGCTTTGTTGAATATGGGAATAGTTATAACCAGATGGGAAGCCTTTTTTCGGCTGTACCTTA 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

TTATCACAAAGCTATAAAAGGCAGCTAAAATAAGCGGAGCCAAGGAAATTACCCAGGCTTATTATGATTTAGCTCTGATTCATTTTAGGAA 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

CAATGAAAAGATTGAAGGACGCGGATTTCTTTGGAAAAGCAATGGAGCAAGCTGAAGAATTTAACGATAGTCTGTTTCAGGATTTATTTAA 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

TGTTCTAAAGGCATTATTTATTGAGACTGGAAGCAGGCAAAAGGTCATGAATGCCTTGAGGCGTTGAGAACAGGACAAGGCTATCCATA 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

TTTTGAAGAGCTAGCTTTAATCGCTGCAGAATTTTATACTATGGATAAGCGCATGGAAGATTCTATATATTTTACAATGAGATGGTGTG 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

CGCTCAAAGACAAATTCAAAGGGGCGATTTCTGTATGAAGTTTAAAGGTCTTTTTTCAGCGTCTTATTGTAAGCCTGTTGGTTGGAG 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

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

AAGATATGAAAAACAGATCGTTATTGATCTGTTTTTTCTTTGTGTGCGACCATGCGGGCATGCTGCGGAGATAACACGGGCTGCTGCGC 8550
↔ ↔

AGATGATGCGGGCATGGGTGCGTTGATCGTGCCGAGACTGAAAAGGCAAAATTTCAAAAAAATCTCCCCCCTACGGGGGGGAAGAAT 8640

GGTTTTGATCTTTTGTTTTGGGTTTTAAAAAAGCCGGCTGTTTTTCAGCCGGTCTTTTTTCGATTTTGGCGAAGCCGAAATCGGGTCTTT 8730

TCTTATC 8837