Supplemental Figure 2. Alignement of regions of chromosomes 5, 1 and 4 comprising At5g58260 (A), At1g74880 (B) and At4g37920 (C), respectively and corresponding cDNA sequences. Upper line shows the genomic sequences retrieved from *Arabidopsis* information resource (TAIR, http://www.arabidopsis.org, September 2004, Rhee *et al*, 2003); in red, untranslated regions; in blue and capital, exons; in blue and small letters, introns. Second line (capital, black, underlined) shows the full length cDNA sequences (RAFL, Seki *et al*, 2002) retrieved from Riken BioResource Center database (http://www.brc.riken.jp/lab/epd/catalog/cdnaclone.html, September 2004), modified nucleotides are indicated in italics. Third line (black and capital letters) shows amino acid sequences deduced from cDNA sequences. Dotted lines indicated that sequences were contiguous. RAFL04-10-E22 and RAFL06-16-D17 perfectly matches At5g58260 and At1g74880 annotations. At4g37920 predicted sequence corresponds to two RAFL cDNA clones (RAFL09-82-J05 and RAFL15-26-D09) indicating that At4g37920 must be split into two genes At4g37920 and At4g37925 as proposed by the Riken BioResource Center and MIPS Arabidopsis thaliana database (MATDB, http://mips.gsf.de/proj/thal/db/index.html, September 2004.) RAFL09-82-J05 is truncated at the 5’ end. A PCR reaction performed with a primer encompassing the ATG codon located 56 base pairs upstream the ATG codon at the 5’ end of RAFL09-82-J05 and another primer located in the midle of RAFL09-82-J05 using cDNA synthesized from *Arabidopsis* RNA as template lead to the amplification of the expected fragment, demonstrating that this sequence is included in the RNA (data not shown). Alignement between the genomic sequence of At4g37925 and RAFL09-82-J05 also revealed the misannotation of an exon (amino acid sequence between brackets).

Supplemental Figure 2

A

```
    At5g58260 (TAIR):             gttagtaaagaaATGGGAAGCCGTGCAATATGTATACAACGTGTGGCACCACCGTGTTTTGAAGCCT
 RAFL04-10-E22 (RIKEN BRC):          AGTAAAGAAATGGGAAGCCGTGCAATATGTATACAACGTGTGGCACCACCGTGTTTTGAAGCCT
                                               M  G  S  R  A  I  C  I  Q  R  V  A  P  P  C  F  E  A

CGCAAGTGAAGAAGATCAAAACAGTTGGTTCCTTTTTGGTGAATACTAGGAGTAAGAGAAGAAGAAGCACTGGTGTGAAATGTAGCAGCATAGCTGACTA
CGCAAGTGAAGAAGATCAAAACAGTTGGTTCCTTTTTGGTGAATACTAGGAGTAAGAGAAGAAGAAGCACTGGTGTGAAATGTAGCAGCATAGCTGACTA
S  Q  V  K  K  I  K  T  V  G  S  F  L  V  N  T  R  S  K  R  R  R  S  T  G  V  K  C  S  S  I  A  D  Y

CATCGGAGGTGACTTGGTGAAGCCGGACATCGGACAATGGCTTCAAGACGTGGAAGAACACAAGGCAATCGCCATTTACGCTCCTCATGAGGGTGGTTAC
CATCGGAGGTGACTTGGTGAAGCCGGACATCGGACAATGGCTTCAAGACGTGGAAGAACACAAGGCAATCGCCATTTACGCTCCTCATGAGGGTGGTTAC
  I  G  G  D  L  V  K  P  D  I  G  Q  W  L  Q  D  V  E  E  H  K  A  I  A  I  Y  A  P  H  E  G  G  Y

GAAGGTCGTTATCTTAACCGTCTCAAAATGCAGGGCTACTACTTTCTTGACATCTCTGCTCGTGGCCTTGGTGACCCTGAGACTACCCTTCTCAAGAACT
GAAGGTCGTTATCTTAACCGTCTCAAAATGCAGGGCTACTACTTTCTTGACATCTCTGCTCGTGGCCTTGGTGACCCTGAGACTACCCTTCTCAAGAACT
 E  G  R  Y  L  N  R  L  K  M  Q  G  Y  Y  F  L  D  I  S  A  R  G  L  G  D  P  E  T  T  L  L  K  N

ACCCTGTTTGCCCTgtatgtctactttcatccatcttctactttaatcttgtttcaatataattttgtgtgaaccatgtatcttctcttcttttagGCTC
ACCCTGTTTGCCCT................................................................................. GCTC
Y  P  V  C  P................................................................................... A

ATCTTGGGAAACAACCTATAGCAAGATGGTATTATCCACCAGAAGTTGATTACAGGCTCGCTGCTCTTCCCCCTAGTGCCAAGGGTCTTGTTGTATGGGT
ATCTTGGGAAACAACCTATAGCAAGATGGTATTATCCACCAGAAGTTGATTACAGGCTCGCTGCTCTTCCCCCTAGTGCCAAGGGTCTTGTTGTATGGGT
H  L  G  K  Q  P  I  A  R  W  Y  Y  P  P  E  V  D  Y  R  L  A  A  L  P  P  S  A  K  G  L  V  V  W  V

CCTCGAAGCAAAGgtttgagtttctctcatttgatctgttctcccacatgctctttctcatatttccatttcactcagGTCCTCTCAAAATCGGAGCTTC
CCTCGAAGCAAAG................................................................ GTCCTCTCAAAATCGGAGCTTC
  L  E  A  K.................................................................. V  L  S  K  S  E  L

AATTTCTCGCTTTACTTCCCTCTCTTCGTCCTAACGTCAGGGTTATTGCTGAATGTGGCAACTGgtgagtatatcattccttaccatacacactaatcat
AATTTCTCGCTTTACTTCCCTCTCTTCGTCCTAACGTCAGGGTTATTGCTGAATGTGGCAACTG....................................
Q  F  L  A  L  L  P  S  L  R  P  N  V  R  V  I  A  E  C  G  N  W....................................

ctatgtttgcaatttgcatgaccattaccaattcccctgcagGAGAAAGTTCGTGTGGAAGCCTCTAGCAGAAATTGCAAATCTAGCTGCACAAGAGTAA
......................................... GAGAAAGTTCGTGTGGAAGCCTCTAGCAGAAATTGCAAATCTAGCTGCACAAGAGTAA
........................................... R  K  F  V  W  K  P  L  A  E  I  A  N  L  A  A  Q  E  *

cttaatggcacaagtggagctaccagtgtatagtttttcatgaatgcatctcccgcaagtgtagcaatactcatcccatagtaaaagttaaaatctgtat
CTTAATGGCACAAGTGGAGCTACCAGTGTATAGTTTTTCATGAATGCATCTCCCGCAAGTGTAGCAATACTCATCCCATAGTAAAAGTTAAAATCTGTAT

ttgtaacttggtagtatattcccatggagtagttaaagataatttctgactctgctttgtcacaaataaaagtatgtataatcagtaagattgataggac
TTGTAACTTGGTAGTATATTCCCATGGAGTAGTTAAAGATAATTTCTGACTCTGCTTTGTCACAAATAAAAGTATGTATAATCAGTAAGATTGGAAAAAA

AAAAAAAAAA
```

B

```
         At1g74880 (TAIR) :                                                gcggtcttatcccgaaactccaagaacacataattt
    RAFL06-16-D17 (RIKEN BRC) :                                            GCGGTCTTATCCCGAAACTCCAAGAACACATAATTT

tttcttggtggtctctccaATGGCTTTCTCTGCGACTGTGTCTCAGCTTTCTTCTCTTTCAACCATCTCCTCTTCATTACCAATTTCTTCTAGAAGACTC
TTTCTTGGTGGTCTCTCCAATGGCTTTCTCTGCGACTGTGTCTCAGCTTTCTTCTCTTTCAACCATCTCCTCTTCATTACCAATTTCTTCTAGAAGACTC
                    M  A  F  S  A  T  V  S  Q  L  S  S  L  S  T  I  S  S  S  L  P  I  S  S  R  R  L

CCTCATCGATCTCTTCCTCAATTCACAGTCAAAGCAGAAGCAGAGAAAGAGAAACAGAGCACACAGGGCAAGTCTGATGGGGAAGCATCACCAGCTGCAA
CCTCATCGATCTCTTCCTCAATTCACAGTCAAAGCAGAAGCAGAGAAAGAGAAACAGAGCACACAGGGCAAGTCTGATGGGGAAGCATCACCAGCTGCAA
 P  H  R  S  L  P  Q  F  T  V  K  A  E  A  E  K  E  K  Q  S  T  Q  G  K  S  D  G  E  A  S  P  A  A

CCAAAACCCCTAAAACCCTCCCCAAGAAACCGGTTTACTCGAgtctgttccctcacttttgtttctaagtgttgaatccaataacagattttgataatt
CCAAAACCCCTAAAACCCTCCCCAAGAAACCGGTTTACTCGA.........................................................
T  K  T  P  K  T  L  P  K  K  P  V  Y  S.............................................................

ctttgaatttctgtttctggtggatgagaaaacagTGAAGAAGGGCCAAATCGTTCGTGTGGAAAAAGAGAAGTACCTCAACAGCATCAATgtatagtc
................................... TGAAGAAGGGCCAAATCGTTCGTGTGGAAAAAGAGAAGTACCTCAACAGCATCAAT........
................................... M  K  K  G  Q  I  V  R  V  E  K  E  K  Y  L  N  S  I  N.......

gaaatctttcccagttttgaagctgcttcttcaaaaatctgtttctgaaacttttgttttgctacatttgttgaaattacagTACTTATCAGTTGGACAT
.................................................................................. TACTTATCAGTTGGACAT
................................................................................... Y  L  S  V  G  H

CCTCCTTTCTACAAAGGACTTGATTACATATACGAAGATCGCGGCGAGgtataacacaatccctttgaaaataaatcacaagagatttcaagcttattg
CCTCCTTTCTACAAAGGACTTGATTACATATACGAAGATCGCGGCGAG....................................................
 P  P  F  Y  K  G  L  D  Y  I  Y  E  D  R  G  E..................................................

aacaagtaatctaaatcttgcagGTCTTGGACCTTCGTGTCTTTGAGACAGGAGAGTATGCACTTgtaagtcagtctccacacaaattgtttattctgtt
.......................GTCTTGGACCTTCGTGTCTTTGAGACAGGAGAGTATGCACTT...................................
....................... V  L  D  L  R  V  F  E  T  G  E  Y  A  L..................................

tcataagatgagctaatggattgtgaattttggagactttagGTTGGATGGGTTGGTATCCCCACCGCACCAGCTTGGCTCCCAACAGATATGCTCATCA
.......................................... GTTGGATGGGTTGGTATCCCCACCGCACCAGCTTGGCTCCCAACAGATATGCTCATCA
...........................................V  G  W  V  G  I  P  T  A  P  A  W  L  P  T  D  M  L  I

AggtactcataatttctcagtacatcaacaatacgtagcattgtcttgatgtcaagtcatgattgtttctgtgttcttggttttgttccaccgcagTG
AG................................................................................................TG
K..................................................................................................C

TGAGAAACTTGTTTACGAGCGAATGTAGagtcaagaaggttccaagaagaatcctttggaatgcaactgatataccaaggattgtgttaagctgagatga
TGAGAAACTTGTTTACGAGCGAATGTAGAGTCAAGAAGGTTCCAAGAAGAATCCTTTGGAATGCAACTGATATACCAAGGATTGTGTTAAGCTGAGATGA
  E  K  L  V  Y  E  R  M  *

Acgataacaatggcgcaccatcatgatcatgtgttcaagagtttgacaaagtatcaacgttacaattagttacaaaaactgttgtggagtgtggaatatg
ACGATAACAATGGCGCACCATCATGATCATGTGTTCAAGAGTTTGACAAAGTATCAACGTTACAATTAGTTACAAAAACTGTTGTGGAGTGTGGAATATG

Tcatcagactcctataataagaatgcagcttcagtaaacggatgatggtaattttatctcatcgatctcccaacataaactcatttctcacagcttcatag
TCATCGGACTCCTATAATAAGAATGCAGCTTCAGG
```

C

```
At4g37925    acagaaaaATGGTTGCAGCATTCTCTTACACTGCCTGCACCAAGCTTTCTCTATTACATCCTTCTATGGTTGCTCAAATCCGACCAA
RAFL09-82-J05                                                        CATCCTTCTATGGTTGCTCAAATCCGACCAA
                      M  V  A  A  F  S  Y  T  A  C  T  K  L  S  L  L  H  P  S  M  V  A  Q  I  R  P

GAACAACACAAAAGGCGTTCGTGGTGACAAATCCGGAACAAGACAGTACTCTTGAAGTACAAGAAACAGAGACACTCAAAGAAGAGCAATCAACAGAGAA
GAACAACACAAAAGGCGTTCGTGGTGACAAATCCGGAACAAGACAGTACTCTTGAAGTACAAGAAACAGAGACACTCAAAGAAGAGCAATCAACAGAGAA
R  T  T  Q  K  A  F  V  V  T  N  P  E  Q  D  S  T  L  E  V  Q  E  T  E  T  L  K  E  E  Q  S  T  E  K

GATGAAGAAGCAGCCAACACCATTGAGACCAGTGGAGAAACAGCTGAACGTGAAGAGCAAGGGTATGGGCGATTTTGGTGGTCAGTGGCTAAGCAGTGTC
GATGAAGAAGCAGCCAACACCATTGAGACCAGTGGAGAAACAGCTGAACGTGAAGAGCAAGGGTATGGGCGATTTTGGTGGTCAGTGGCTAAGCAGTGTC
  M  K  K  Q  P  T  P  L  R  P  V  E  K  Q  L  N  V  K  S  K  G  M  G  D  F  G  G  Q  W  L  S  S  V

ACACGTCACGTTAGGATCTACGCTGCTTATATTGATCCAGAGACTTGTGAGTTTGACCAGTCACAAATGGATAAACTCACTTTGATTCTTGATCCTACTG
ACACGTCACGTTAGGATCTACCCTGCTTATATTGATCCAGAGACTTGTGAGTTTGACCAGTCACAAATGGATAAACTCACTTTGATTCTTGATCCTACTG
 T  R  H  V  R  I  Y  A  A  Y  I  D  P  E  T  C  E  F  D  Q  S  Q  M  D  K  L  T  L  I  L  D  P  T

AAGAATTTGTGTGGGATGATGAAAGCTGTAACAAGGTCTATTCTTATTTCCAAGAACTCGTCGATCACTACGAGGTAACGATAACAATAAACGAACGAAA
AAGAATTTGTGTGGGATGATGAAAGCTGTAACAAGGTCTATTCTTATTTCCAAGAACTCNTCGATCACTACGAG..........................
E  E  F  V  W  D  D  E  S  C  N  K  V  Y  S  Y  F  Q  E  L  V  D  H  Y  E [V  T  I  T  I  N  E  R  N

CATTCATCTAAAAGTTAGACTCATTTCAAGAAGACCTTATAATATTGTTTTTGAACAGGGAGCACCATTAACGGAGTACACATTGAGATTGATAGGATCG
.......................................................... GGAGCACCATTAACGGAGTACACATTGAGATTGATAGGATCG
  I  H  L  K  V  R  L  I  S  R  R  P  Y  N  I  V  F  E  Q] G  A  P  L  T  E  Y  T  L  R  L  I  G  S

GACGTAGAACATTACATAAGGAAGATGTTGTTCGATGGAGAAATACAATACAATATGGATGCAAGAGTTCTTAATTTCAGTATGGGAAAGCCTCGTGTAC
GACGTAGAACATTACATAAGGAAGATGTTGTTCGATGGAGAAATACAATACAATATGGATGCAAGAGTTCTTAATTTCAGTATGGGAAAGCCTCGTGTAC
 D  V  E  H  Y  I  R  K  M  L  F  D  G  E  I  Q  Y  N  M  D  A  R  V  L  N  F  S  M  G  K  P  R  V

AATTCAATACCAGTAACATTGAAGGTGGTGGAGATGgtcagcctcaagaggacgcttagtgaaacaaactctatttataatcatgtaatgggaatataa
AATTCAATACCAGTAACATTGAAGGTGGTGGAGATGGTCAGCCTCAAGAGGACGCTTAGTGAAACAAACTCTATTTATAATCATGTAATGGGAATATAA
Q  F  N  T  S  N  I  E  G  G  G  D  G  Q  P  Q  E  D  A  *

atatattgtagtatgcttttaattttgtctatcacaaagtaacatatgtatatatcatgacatatgctataattattttgtaataatcaatcttaagt
ATATATTGTAGTATGCTTTTAATTTTGTCTATCACAAAGTAACATATGTATATATCATGACATATGCTATAATTATTTTGTAATAATCAATCTTAAGT

ttctttgtcttttcttgttaatgtaaggtcttaaggtcatgtgaatatgtaatgtcaaagaaactcaattctggtcttgctcattcatgttccttactcc
TTCTTTGTCTT

ttgcctctttttaaaaagatttcaaactgatgcaaaacttgttagttttaacatttttacccaaaaagttatattaagtttttgaatttgtcaacgttta

At4g37920
ctaattatttacaattaaaaccctgttcttcttatcctctggtgtttctcgttgttcgtcaccaagACGAACTTTGTGAGGGAGAAAAAATGGCGAATTT
RAFL15-26-D09                    ATCCTCTGGTGTTTCTCGTTGTTCGTCACCAAGACGAACTTTGTGAGGGAGAAAAAATGGCGAATTT
                                                                                          M  A  N  L

ACTGGAAACATCAATCTTTTTCTCCTCTGCTGATAAATTACTGTCTTTCCCGCCTAAGAATTCTCAAACCCATCATCTCCCTTTCTCTGCTTTCATCAAT
ACTGGAAACATCAATCTTTTTCTCCTCTGCTGATAAATTACTGTCTTTCCCGCCTAAGAATTCTCAAACCCATCATCTCCCTTTCTCTGCTTTCATCAAT
  L  E  T  S  I  F  F  S  S  A  D  K  L  L  S  F  P  P  K  N  S  Q  T  H  H  L  P  F  S  A  F  I  N

GGAGGAAGAAAGATACGCAAAAGCTCAACTATCACCTTTGCTACCGACACTGTCACTTACAATGGCACTACAAgtaagctcataaaccctaaatctctaa
GGAGGAAGAAAGATACGCAAAAGCTCAACTATCACCTTTGCTACCGACACTGTCACTTACAATGGCACTACAA..........................
 G  G  R  K  I  R  K  S  S  T  I  T  F  A  T  D  T  V  T  Y  N  G  T  T ...........................

tctttctgtctaattaccaataaattctatttctttagattccaatttgtattaattgtcaccattagataagctaaaccctttttcatagactaggattc
....................................................................................................

....................................................................................................

gtctagttagagaagctttttttattgaagatagaaatttgtgtttagGTGCGGAAGTAAAAAGCTCTGTTGAGGATCCAATGGAAGTTGAGGTTGCTG
................................................ GTGCGGAAGTAAAAAGCTCTGTTGAGGATCCAATGGAAGTTGAGGTTGCTG
.................................................S  A  E  V  K  S  S  V  E  D  P  M  E  V  E  A

AAGGTTACACTATGGCTCAGTTTTGTGACAAGATTATTGATCTCTTCTTGAATGAGAAACCTAAAGTTAAACAATGGAAGACGTATTTGGTTTTAAGAGA
AAGGTTACACTATGGCTCAGTTTTGTGACAAGATTATTGATCTCTTCTTGAATGAGAAACCTAAAGTTAAACAATGGAAGACGTATTTGGTTTTAAGAGA
E  G  Y  T  M  A  Q  F  C  D  K  I  I  D  L  F  L  N  E  K  P  K  V  K  Q  W  K  T  Y  L  V  L  R  D
```

```
TGAATGGAACAAGTACAGTGTTAATTTCTATAAGCGTTGCAGAATCCGAGCTGATACCGAAACTGATCCTATATTGAAGCAGAAACTGGTATCTTTGGAA
TGAATGGAACAAGTACAGTGTTAATTTCTATAAGCGTTGCAGAATCCGAGCTGATACCGAAACTGATCCTATATTGAAGCAGAAACTGGTATCTTTGGAA
  E  W  N  K  Y  S  V  N  F  Y  K  R  C  R  I  R  A  D  T  E  T  D  P  I  L  K  Q  K  L  V  S  L  E

AGTAAAGTCAAGAAGgtatttatagtgttgtattttgtgtttattagtgaactcaggttagtttattgttgttgttgtgtacttataggagtgagtattg
AGTAAAGTCAAGAAG.......................................................................................
 S  K  V  K  K.........................................................................................

ttgcagATAGATAAGGAAATGGAGAAACACAATGACCTTTTGAAGGAGATTCAAGAAAACCCAACAGACATAAACGCAATCGCTGCCAAGAGACGCAGAG
..... ATAGATAAGGAAATGGAGAAACACAATGACCTTTTGAAGGAGATTCAAGAAAACCCAACAGACATAAACGCAATCGCTGCCAAGAGACGCAGAG
.......I  D  K  E  M  E  K  H  N  D  L  L  K  E  I  Q  E  N  P  T  D  I  N  A  I  A  A  K  R  R  R

ACTTTACGGGAGAGTTTTTCCGCTATGTTACTTTGTTATCAGAAACTCTTGATGGCTTGGAAGACCGAGATGgtacttaaaacattaatcatgttgctta
ACTTTACGGGAGAGTTTTTCCGCTATGTTGCTTTGTTATCAGAAACTCTTGATGGCTTGGAAGACCGAGATG..............................
D  F  T  G  E  F  F  R  Y  V  T  L  L  S  E  T  L  D  G  L  E  D  R  D...............................

cttattgattcccacagtttcatgattgttttgttattttgagcagCTGTTGCGAGGCTTGCAACTAGATGTTTATCTGCTGTGAGTGCTTATGACAAC
..............................................CTGTTGCGAGGCTTGCAACTAGATGTTTATCTGCTGTGAGTGCTTATGACAAC
..............................................A  V  A  R  L  A  T  R  C  L  S  A  V  S  A  Y  D  N

ACGTTGGAGAGTGTTGAGACTCTAGATACTGCTCAAGCAAAGTTTGAAGATATTTTAAACTCTCCATCCGTGGATTCTGCTTGCGAAAAGATCAGAAGTT
ACGTTGGAGAGTGTTGAGACTCTAGATACTGCTCAAGCAAAGTTTGAAGATATTTTAAACTCTCCATCCGTGGATTCTGCTTGCGAAAAGATCAGAAGTT
 T  L  E  S  V  E  T  L  D  T  A  Q  A  K  F  E  D  I  L  N  T  L  E  S  V  E  T  L  D  T  A  Q  A

TAGCCAAGGCAAAGGAGCTTGATTCCTCATTGATCTTATTGATCAACAGTGCATACGCTGCTGCAAAAGAGTCTCAAACTGTTACGAACGAGgtaatagc
TAGCCAAGGCAAAGGAGCTTGATTCCTCATTGATCTTATTGATCAACAGTGCATACGCTGCTGCAAAAGAGTCTCAAACTATTACGAACGAG.........
K  F  E  D  I  L  N  D  S  S  L  I  L  L  I  N  S  A  Y  A  A  A  K  E  S  Q  T  V  T  N  E.........

cattggattacaatcttctttaatattcttcacaatcttctcaaatcgatatcgtatttttgaacagGCTAAAGACATAATGTATCATTTATACAAAGC
...................................................................GCTAAAGACATAATGTATCATTTATACAAAGC
.................................................................... A  K  D  I  M  Y  H  L  Y  K  A

TACAAAGAGCAGTCTTAGAAGCATCACACCGAAAGAGATCAAACTTTTAAAGTACTTACTCAACATAACAGACCCCGAAGAACGATTCTCTGCCCTAGCT
TACAAAGAGCAGTCTTAGAAGCATCACACCGAAAGAGATCAAACTTTTAAAGTACTTACTCAACATAACAGACCCCGAAGAACGATTCTCTGCCCTAGCT
  T  K  S  S  L  R  S  I  T  P  K  E  I  K  L  L  K  Y  L  L  N  I  T  D  P  E  E  R  F  S  A  L  A

ACAGCATTCTCTCCTGGTGATGACCACGAAGCTAAGGACCCTAAAGCATTATACACgtaagtaacataaagatgcacactttagtgtttcatctttaacc
ACAGCATTCTCTCCTGGTGATGACCACGAAGCTAAGGACCCTAAAGCATTATACAC............................................
 T  A  F  S  P  G  D  D  H  E  A  K  D  P  K  A  L  Y  T............................................

acaaccattgattctaaccaagctgatgacacagGACGCCAAAGGAGCTGCATAAGTGGATAAAGATAATGCTTGATGCTTACCATCTCAACAAAGAAGA
..................................GACGCCAGAGGAGCTGCATAAGTGGATAAAGATAATGCTTGATGCTTACCATCTCAACAAAGAAGA
.................................... T  P  K  E  L  H  K  W  I  K  I  M  L  D  A  Y  H  L  N  K  E  E

AACAGACATTAAAGAAGCTAAACAGATGAGCCAACCTATCGTTATTCAGCGGCTTTTCATTCTCAAGGACACTATCGAAGACGAGTATTTAGACAAGAAA
AACAGACATTAAAGAAGCTAAACAGATGAGCCAACCTATCGTTATTCAGCGGCTTTTCATTCTCAAGGACACTATCGAAGACGAGTATTTAGACAAGAAA
  T  D  I  K  E  A  K  Q  M  S  Q  P  I  V  I  Q  R  L  F  I  L  K  D  T  I  E  D  E  Y  L  D  K  K

ACAATCGTGGCTGATGAAACTCCAAAGAAGGAAGAAGAAGATACGACTATTGAAGATTTTTTGAACTGAaaattgtgaaatccatgaacacactctgttt
ACAATCGTGGCTGATGAAACTCCAAAGAAGGAAGAAGAAGATACGACTATTGAAGATTTTTTGAACTGAAAATTGTGAAATCCATGAACACACTCTGTTT
 T  I  V  A  D  E  T  P  K  K  E  E  E  D  T  T  I  E  D  F  L  N  .

acaattagtaccttttatgtaatggggaaaagagtaattttgtagaattctcagctataattgcaattacgactaaatgtgaggtctaatttgtaattaa
ACAATTAGTACCTTTTATGTAATGGGGAAAAGAGTAATTTTGTAGCAAAAAAAAAAAAAAAAA
```
