

Supplementary file 6.

Gene structure of carnitine palmitoyltransferase 1 genes (*cpt 1*) in Nile tilapia. Exons are shown in red font capital letters and deduced amino acid sequence coded by each exon is shown below the exon in blue font capital letters. The transcription initiation site is indicated by +1. Lowercase letters indicate introns. Putative TATA boxes and polyadenylation signals are shown in underlined letters and highlighted in yellow. An asterisk indicates the stop codon in the coding sequence of each gene. Forward and reverse primers for RT-qPCR are highlighted in yellow and green, respectively. RT-PCR primers used to amplify the retinoid-binding gene transcripts are shown in green and turquoise, respectively.

a) *cpt 1a1a*

5' ctgataaaagatgagggatagttctgatattttgccttacaccaatcattcacagccttat
ctagagatcagtatgctgtgtgaaccttgtagtgttgaccaatcttgaagttaatatc
tgtgtaaagccttgtgaagaagagattgtataggaacaagaatgaagatttagatcaggt
gttgattgacgtgctccttaaatctttaatgtttggcataaactttattcagagtttcac
TATA Box
tttgctaataactactatttgccttactgtcctcagtgaccctatacatcagtagctacc
cgaaaattgagatatgtgacgtagcaaaatgggcagtcctgtgtatacaagtctgcggtct
+1
cattcataagGAGTGTAAACGTTAGCAGTCCCCGCCGACAGCCAGCCAGAGTCACATCC
TCCTCACAACtgaagttttg-N₃₂₆₆-ctgtcctcagTGACCCTATACATCAGTAGCTA
CCATGGCAGAAAGCCACAGGCGGTGGCCTTTTCAAGTTCACAGTCACCCCTGAAGGCATTG
-M--A--E--A--H--Q--A--V--A--F--Q--F--T--V--T--P--E--G--I--
ATCTACACCTGTCCCATGAGGCTCTTCGCCAGGTCTACCTTTCTGGCATCCACTCCTGGA
D--L--H--L--S--H--E--A--L--R--Q--V--Y--L--S--G--I--H--S--W--
AGAAAAGATTTGTTTCGCTTCAAGgtgggaggca-N₄₁₄-catttccacagAATGGGATAAT
K--K--R--F--V--R--F--K-- -N--G--I--M
GACTGGTGTGTACCCAGGACAGCCAGCTGGGTCTCCATAGTGGTTGGGTCCTACATGGC
--T--G--V--Y--P--G--S--P--A--G--F--S--I--V--V--G--S--Y--M--A
TTACAATAAATTTAAACAGCTGGATCCATCTTTGGGCTGTGTTAACAAGCTGGGTCAACA
--Y--N--K--F--Q--L--D--P--S--L--G--L--F--T--K--L--G--Q--H
TATACCCATCAGgtaagctgat-N₁₀₈₀-ctgtccacagCCGATATATGTCCACAGACAG
--I--P--I--S --R--Y--M--S--T--D--S
CCAGAAGATAGTCGGAGGAATTCCTAGTGGGCACAAGCCTGTGGGTTACCATAATAATGAT
--Q--K--I--V--G--G--I--L--V--G--T--S--L--W--V--T--I--I--M--I
TATGAGAAGTGTCTCAAGTCCTTGCTGTCTATGGCATGGATGGATGCAAGAACGTCATGG
--M--R--T--V--L--K--S--L--L--S--W--H--G--W--M--Q--E--R--H--G
CTCCTTGTCTGGTTCGAGTCGCATTTGGATGgtgagacaat-N₄₇₁-gcttgtccagGTT
--S--L--S--W--S--S--R--I--W--M-- -V--
CTAGTGAAGGTGTTCTCTGGCCGAAAACCAATGCTGTACAGTTTCCAGAATCCCTGCCA
-L--V--K--V--F--S--G--R--K--P--M--L--Y--S--F--Q--N--S--L--P--
CGACTACCTGTTTCCTTCTATCAAGGACACCTGCGAAAGGgtaatttctt-N₃₈₆-gtggtt
-R--L--P--V--P--S--I--K--D--T--C--E--R--
ttcagTACTTGGAGTCAGTGCGCCCACTGATGGATGATCAGCAGTTTGAACGAATGAAAG
-Y--L--E--S--V--R--P--L--M--D--D--Q--Q--F--E--R--M--K--
GCCTGACTCAAGACTTTGAGAAGAATCTGGGTCCCAGGCTGCAGTGGTACCTTAAGCTCA
G--L--T--Q--D--F--E--K--N--L--G--P--R--L--Q--W--Y--L--K--L--
AGTCTGGTGGGCGTCCAACATgtgagttctc-N₉₆-cctgtcacagGTCAGTGACTG
K--S--W--W--A--S--N--Y-- -V--S--D--W
GTGGGAAGAGATATATTTACCTGAGAGGCCGTGGACCTATCATGGTGAACAGTAACATATTA
--W--E--E--Y--I--Y--L--R--G--R--G--P--I--M--V--N--S--N--Y--Y
TGCCATGgtaagacaac-N₂₀₆₆-gtttctccagGATTTCTGTACGTGTTCCCCACCTC
--A--M-- -D--F--L--Y--V--F--P--T--S
CATCCAGGCTGCCAGGGCAGGTAATGCCATTCTATCATGCTGTACAGGAGAAAAC
--I--Q--A--A--R--A--G--N--A--I--H--S--I--M--L--Y--R--R--K--L

GGACAGAGCTCAGATCAAACCTgtgagtatag-N₁₁₈₂-tgcttttcagCTCATGCTCCT
--D--R--A--Q--I--K--P-- -L--M--L--L--
ACACACTATTCCCATGTGCTCAGCCCAATACGAGCGTATGTTCAACACGACTCGTGTCCC
--H--T--I--P--M--C--S--A--Q--Y--E--R--M--F--N--T--T--R--V--P
AGGTGTAGAAACAGgtagagtctt-N₂₅₆₈-tctccaatagACACCCCTTCAGCATGTGAA
--G--V--E--T-- -D--T--L--Q--H--V--N
CGAGAGCAAACACATAGTTGTGTACCATAAGGGACGTTACTTCAAGGTGTGGATGTTTTTA
--E--S--K--H--I--V--Y--H--K--G--R--Y--F--K--V--W--M--F--Y
TGATGGGCGGCTGCTGTTGCCACGGGAGATCGAGCAGCAGATGGAGCGGATCTTGGCAGA
--D--G--R--L--L--L--P--R--E--I--E--Q--Q--M--E--R--I--L--A--D
CAAGTCGGAGCCCTTGCCAGGAGAGGAGAGACTAGCAGCACTTACTGCAGGGGATAGgtg
--K--S--E--P--L--P--G--E--E--R--L--A--A--L--T--A--G--D--R
agtttca-N₆₇₅-aatttgttagAACACCGTGGGCAAAGGCCCGTGAGTCGTTCTTCAGC
--T--P--W--A--K--A--R--E--S--F--F--S--
CGTGGAAGAACAAGCAGTCTCTGGATGCCGTAGAAAAGGCAGCCTTCTTCTTGACCCTC
-R--G--K--N--K--Q--S--L--D--A--V--E--K--A--A--F--F--L--T--L--
GATGACACAGAGCAGCGTTACGATACAAAAACCCAGTGAAGTCTCTGGACATTTATGCC
-D--D--T--E--Q--R--Y--D--T--K--N--P--V--K--S--L--D--I--Y--A--
AAGTCCCTGCTCCATGGAAAATGCTACGACAGgtggactaaa-N₇₈-aatccttcagGTG
-K--S--L--L--H--G--K--C--Y--D--R -W
GTTTGATAAGTCCCTTAATATGATAGTTTACAAGAATGGTACGATGGGACTGAATGCAGA
--F--D--K--S--L--N--M--I--V--Y--K--N--G--T--M--G--L--N--A--E
GCATTCTTGGGCTGATGCTCCGATTGTTGGCCATCTGTGGGAGgtaagtcact-N₁₆₉₉-
--H--S--W--A--D--A--P--I--V--G--H--L--W--E--
tctcttgcagCATGTACTTTCCATGGATCCCAAGCTGGGATACACTGAGGATGGTCACTG
-H--V--L--S--M--D--P--K--L--G--Y--T--E--D--G--H--C
CGTTGGGAAGCCTCATCCTAACCTGCCCGGCCCTCAAAGGCTGCAGTGGACCATCCCTGC
--V--G--K--P--H--P--N--L--P--G--P--Q--R--L--Q--W--T--I--P--A--
TGAGgtagaaaatc-N₂₄₃-attcctctagTGCCAGGAGGCCATCGAGAGTTCTTTTGACA
--E-- -C--Q--E--A--I--E--S--S--L--T--
GTAGCCAGAGCTCTGGCAGACGATGTGGACTGTCACATTATCCCCCTTCACTGACTTTGGA
-V--A--R--A--L--A--D--D--V--D--C--H--I--I--P--F--T--D--F--G--
AAGGGGCTGATTAAGAAGTGCCGCACCACTCCCGATGCCCTTCATCCAGATTGCCCTGCAG
-K--G--L--I--K--K--C--R--T--S--P--D--A--F--I--Q--I--A--L--Q--
CTGGCCCATTTACAGAgtaaattccc-N₅₄₉-ctctcaacagGACAAAGGGAAGTTTTGCC
-L--A--H--Y--R-- -D--K--G--K--F--C--
TGACTTATGAAGCCTCCATGACTCGTATGTTTCAGGGAAGGTGCAACAGAACTGTGCGCT
L--T--Y--E--A--S--M--T--R--M--F--R--E--G--R--T--E--T--V--R--
CCTGCACCATAGAGACTTGTGCCTTTGTTCGTGCCATGGTCGGAGATGAGACAgtagtc
S--C--T--I--E--T--C--A--F--V--R--A--M--V--G--D--E--T--
ctg-N₁₂₀-ctttctacagAGAGAAGAGCGTCTCAGGTGCTCAAGCTGGCAGCAGAAAA
-R--E--E--R--L--R--L--L--K--L--A--A--E--K
ACACCAAAACTTGTACCGTTTGGCAATGACAGGACAGGGCATCGATCGCCACCTTTTCTG
--H--Q--N--L--Y--R--L--A--M--T--G--Q--G--I--D--R--H--L--F--C
CCTCTATGTGGTTTCCAAATACCTCGGAGAGGATTACCCCTTCTGAAGGAGgtatgagc
--L--Y--V--V--S--K--Y--L--G--E--D--S--P--F--L--K--E--
tt-N₁₃₉-ttacatttagGTCCGTGCTGAGCCATGGAAGCTGTCCACCAGCCAGACTCCCT
-V--L--S--E--P--W--K--L--S--T--S--Q--T--P--
CTGCAGCAGGTTGAGCTGTTTGATCTGGTCAGACACCCAGAATACGTGTCCTCTGGAGGT
-L--Q--Q--V--E--L--F--D--L--V--R--H--P--E--Y--V--S--S--G--G--
GGTTTTGGTCCAgtagttgtc-N₁₇₄-tgcccacaagGTGGCTGATGATGGTTATGGCG
-G--F--G--P-- -V--A--D--D--G--Y--G--
TCTCCTACATCATTTCTTGGTGAGAACCACATCAACTTCCACATCTCCAGCAAACGCTCAA
V--S--Y--I--I--L--G--E--N--H--I--N--F--H--I--S--S--K--R--S--
GCCCAGAAACTgtaagttcac-N₈₄-tttttcttagGACTCACACCGTTTTGGGACCAAC
S--P--E--T-- -D--S--H--R--F--G--T--N--
ATCAGACAGGCCATGCTGGACATTCTGGGTCTCTTTGAGCTCAACAACAAAGCTGCCAAG

-I--R--Q--A--M--L--D--I--L--G--L--F--E--L--N--N--K--A--A--K--

TGA tggctttgctggtggaagactatgagaaacatatattc aaaaaa aggttgtacaga

-*-

atcacttggttggttagaggttatatttaatcacttggttggttagaggttatatttatggt
aattaatgttacaaaaataagaatgattgaaagtcttttggaagcattatcgttaatatggt
tgtgaggcacttatgtaagctgtatatttaggcatagtgagattccggtggtggttggtaca
ggaatgaacagatcaaagtt -3'

b) *cpt 1a2a*

5' gcctcatcccaagggttaacagagtctcgccgatcctgttcatctggaaaacctgttagat
TATA Box
attdaaacaggtgtctgtttctgttgcctgttttaaaagtggcacctggaatgttaaa
+1
tgcTTTCACCTGATGTCTACTTTTCGCCTGAGGTTAGTAATTTTCCACACGTGTTCTTGTA
TCTAAACACTTGTTTTCTCTGGCCTCTAGGGTGTCCGATGACTTTTCTACATTTTGGTTA
AAGAAATCCTTTCTCTCGTCTGCTCCATCTTTATTTTAGGTTCACTGTGCACGGCAGTC
ATGGCAGAAGCCCACCAAGCCGTCGCTTTTCAATTCACCGTTACCCCAGATGGCATTGAT
-M--A--E--A--H--Q--A--V--A--F--Q--F--T--V--T--P--D--G--I--D--
CTGCAGCTGTCCCACAGGCCCTCACTGAGATCTACCTCTCTGGCATGCGTTCATGGAAG
-L--Q--L--S--H--Q--A--L--T--E--I--Y--L--S--G--M--R--S--W--K--
AAGCGTATCGTCAGAGTTAAGgtatctttgt-N₅₆₈-actctaccagAACAGTGTGATAAC
-K--R--I--V--R--V--K--N--S--V--I--T
AGGAGTGTATCCTGCTAGTCCGTCACTCCGCTGTTTGTGGTCAATTGCAATCCTGGCTAC
--G--V--Y--P--A--S--P--S--S--W--L--F--V--V--I--A--I--L--A--T
TATGTACACTCGCTCAGACCCCTCCATGGGACTCATAGCCAAGATACAGGAGCACC TGCC
--M--Y--T--R--S--D--P--S--M--G--L--I--A--K--I--Q--E--H--L--P
AGTCAGgtaacacacc-N₁₂₃₀-tcattttccagCCAGTCGATGAGTACGCAGTGCCAAGCAG
--V--S--Q--S--M--S--T--Q--C--Q--A--
TGGTGTGACGAGTGTCTTTCAGCACCATGCTTTGGCTAATGCTCATTTTCACTATGCGCC
V--V--S--A--V--L--F--S--T--M--L--W--L--M--L--I--F--T--M--R--
TGTGCCCTCAAGCAGCTCCTCTCCTACCACCGCTGGATGTTTCGAGCAGCATGGCAAGATGT
L--C--L--K--Q--L--L--S--Y--H--R--W--M--F--E--Q--H--G--K--M--
CCACCAACCAAAATCTGGGTGgtcagtgctt-N₃₄₁-tctatgtcagGCGCTGGTGC GC
S--T--T--T--K--I--W--V--A--L--V--R--
ATCTTTTCTGGCAGAAAGCCTCTGCTCTACAGCTATCAGGGTTCGCTGCCAAATCTCCCT
-I--F--S--G--R--K--P--L--L--Y--S--Y--Q--G--S--L--P--N--L--P--
GTGCCCACCATCAAGGACACAGTCAAGAGgtgagatgtc-N₂₄₁-ttttatttagTACTT
-V--P--T--I--K--D--T--V--K--R--Y--L--
GGAATCAGTCCGTCGCTGATGGACGATAAAGAGTATGAACGCATGACGAAGCTGGCCGC
--E--S--V--R--P--L--M--D--D--K--E--Y--E--R--M--T--K--L--A--A--
AGAGTTTGTAGAGCAGCCTTGGTAACCGCCTGCAGTGGTACCTCAAACCTCAAAGCTCTCTG
--E--F--E--S--S--L--G--N--R--L--Q--W--Y--L--K--L--K--A--L--W--
GGCTTCAAACATATgtaagtgggt-N₁₂₁-ttgctttcagGTCAGTGACTGGTGGGAGGAAT
--A--S--N--Y--V--S--D--W--W--E--E--
ATGTCTACCTCCGGGGTCGAGGCCCAATCATGGTCAACAGTAACACTACTATGGCATGgtaa
Y--V--Y--L--R--G--R--G--P--I--M--V--N--S--N--Y--Y--G--M--
gaaaat-N₅₁₁-ttctcctcagGACTTCTTGTATGTGACTCCAACCTCCATCCAGGCAGCC
-D--F--L--Y--V--T--P--T--P--I--Q--A--A--
AGGGCCGGCAACAGCATCCATTCTTCTCTTACC GCCGCAAACCTCAACAAAGAAGAG
-R--A--G--N--S--I--H--S--F--F--L--Y--R--R--K--L--N--K--E--E--
CTCAAACCTgtgagtttaag-N₁₀₉₄-ttctactcagTGGTTGCTGAGGTCTGCAGTTCCTT
-L--K--P--W--L--L--R--S--A--V--P--
GCTGTTCTTATCAGTTTGTAGCGGATGTTTGTGACTTGTGCAATCCCTGGAATACTAACAG
C--C--S--Y--Q--F--E--R--M--F--D--T--C--R--I--P--G--I--L--T--
gtagtattta-N₁₄₆₆-tcttttttatccaaagATACAGTCCAGCATTGGCAGGACAGCGA
D--T--V--Q--H--W--Q--D--S--D--
CTACATTGTGGTTTACCACAAAGGGCGCTACTTTCGTCTCAGGGTGTACCAGGCAGGAG
--Y--I--V--V--Y--H--K--G--R--Y--F--R--L--R--V--Y--Q--A--G--R--
ACTTCTGTCTCCGAGGGAGATTGAATTCCAGATTTCAGAGGATCCTAGATGACCTTTCACC
--L--L--S--P--R--E--I--E--F--Q--I--Q--R--I--L--D--D--P--S--P--
TCCTTCCAAAGGGGAGGCCAAACTTGGGGCCCTGACTGCTGGAGACAGgtcagtttga-N₅₄-
--P--S--K--G--E--A--K--L--G--A--L--T--A--G--D--R--
taatcaccagAATTCCTGGGCTAAGGCCAGGACCAAGTATTTTCAGCAGCGGGGTCAATA
--I--P--W--A--K--A--R--T--K--Y--F--S--S--G--V--N--

AACGCTCTTTGGACTGCATTGAGAAGGCCGCCCTTCTTTGTGACACTGGATGACGAAGAAC
K--R--S--L--D--C--I--E--K--A--A--F--F--V--T--L--D--D--E--E--
AGGGCATGATGGGAGATGACCCAGCAGCCAGTTTAGACCGCTATGCCAAATCTCTATTGC
Q--G--M--M--G--D--D--P--A--A--S--L--D--R--Y--A--K--S--L--L--
ACGGGAAATGTTATGACAGgtactttttac-N₉₅-cgccatgcagGTGGTTTGACAAGTCTT
H--G--K--C--Y--D--R--W--F--D--K--S--
TCACAGTTGTGTACTACAAGAACGGGAAAAATGGCATCAATGCAGAGCACTCGTGGGCTG
F--T--V--V--Y--Y--K--N--G--K--N--G--I--N--A--E--H--S--W--A--
ATGCGCCAGTGTGGCACACGTGTGGGAGgcaaggaata-N₆₁-gcggtttccagTACACTC
D--A--P--V--L--A--H--V--W--E--Y--T--
TGGCCAATGACAGCTTCCAGCTCGGTTACAACGCAGAAGGTCACGTGTAAAGGAGACGTGG
L--A--N--D--S--F--Q--L--G--Y--N--A--E--G--H--C--K--G--D--V--
ATCCCTCCCTGCCCCGACCAGTGAAGCTCAGCTGGGAGATCCCTCCAGAAgttagttctg
D--P--S--L--P--R--P--V--K--L--S--W--E--I--P--P--E--
-N₃₃₉-tgtgtgacagTGTGAAGAGCAGATCGCTCAGTCTCTGGCTGTGGCCAGGCGTTG
C--E--E--Q--I--A--Q--S--L--A--V--A--Q--A--L--
GCTGATGATGTGGACTTCCATGTTTTCTCCTTCCAAGAATTTGGCAAAGGAAAAGTCAAG
A--D--D--V--D--F--H--V--F--S--F--Q--E--F--G--K--G--K--V--K--
AAGTGTGAGTCACTCCAGACGGCTTCATTCAAGATGGCTCTTCAGTTGGCCTACTTCAGG
K--C--R--V--S--P--D--G--F--I--Q--M--A--L--Q--L--A--Y--F--R--
gtgagtggttc-N₅₅₆-gtggttgacagGAACGGGGGACTTTCGTGTTTGACATATGAAGCCTC
E--R--G--T--F--C--L--T--Y--E--A--S--
CATGACCCGCTCTGTTTCAGGGAGGGCAGGACCGAGACCGTTCGCTCCTGCACCAACGAGTC
M--T--R--L--F--R--E--G--R--T--E--T--V--R--S--C--T--N--E--S--
TAGCGCCTTCATCCGAGCGCTGGAGGGCGGAGAGgtgagagcat-N₆₈-tctccaccagGC
S--A--F--I--R--A--L--E--G--G--E--A--
AACAGACGTGTGCAAATGTTTGTTCGAGTTGCCCTCAGAAAAGCACCAGCTACTTTACCG
T--D--V--C--K--C--L--F--R--V--A--S--E--K--H--Q--L--L--Y--R--
CTTGGCCATGACTGGAGCCGGCATCGACAGACACCTGTTCTGCCTCTACGTGGTGTCCAA
L--A--M--T--G--A--G--I--D--R--H--L--F--C--L--Y--V--V--S--K--
ATACCTCGGAGTGGAGTCTCCTTTCTCAAGAGgtttgtttcc-N₆₇-tgtgtttcagG
Y--L--G--V--E--S--P--F--L--K--E--
TTCTGTCTGAGCCGTGGAGGCTGTCCACAAGTCAGACTCCCATCCAGCAGGTGGAGCTGT
V--L--S--E--P--W--R--L--S--T--S--Q--T--P--I--Q--Q--V--E--L--
ttgacatagagaaccaccctgagtatgtctcctgtggaggaggctttggaccggtcagta
F--D--I--E--N--H--P--E--Y--V--S--C--G--G--G--F--G--P--
cca-N₁₃₂-tgtctttcagGTGGCCGATGATGGATATGGGGTGTCTACTGCGTTTTGGGA
V--A--D--D--G--Y--G--V--S--Y--C--V--L--G--
GAGAACATGATTAACCTCCACATCTCATGCAAACACTCATGTCCAGACACTgtaagtcta
E--N--M--I--N--F--H--I--S--C--K--H--S--C--P--D--T--
-N₁₅₃-ctccttgaagGACGCTCATAAGTTTGGTACTCAGATCAGACAAGCCCTCCGAGAT
D--A--H--K--F--G--T--Q--I--R--Q--A--L--R--D--
CTGCTAAAGCTGCTGAGCCCCAACCAAACAGAGCCAGCAAAACGGAAGAGAGCCAACCC
L--L--K--L--L--S--P--N--Q--T--E--P--S--K--T--E--E--S--Q--P--
AAGGTCAAGAAAGAGCAGTAGatattcaaatcagaagatgtgatgtcgaagggaacatgt
K--V--K--K--E--Q--*-
gcaggcattacaatatgaagtaaagaaaggcggtttttgttcttaatttagtctgagactt
ttaaagccaggtagatgatattcgggagacagctccactggaaaatgataacaacggag
tga**tggtttt**tactgacatgcataatgtttgtgcgtttatattagctgtgcgatgggtgc
ggagcgagaaagggtccacagagctggttttaacagtgaacaacttgaaatgtgcacat
gtgtaaatgaaaagtacagtg -3'

c) *cpt 1a2b*

5' ctgctgaacggtgtgcagtcacagactttttcacatttgaagtatcatgaaagaaaaatggc
gactaaagaaaaccagaactgctgagctgctacaatcctgtatcggacaaattattaag
TATA Box
ttagtcagcaaatacacatacaaatattgtcagttgaagattgaagcatc **ataaaaa**aaaa
atTTTTtctcacaaagtattattatctaagcaagcatgaacctccttactaatgacatcc
+1
atcactatggggcgtttttaacaaggctgttaaacctcgtttgggttAACACATTATGGTT
ATGGATGAATCCAGATCATTTCATCGGTCCCTCCCGACAACACAACAGGTCTGTGTCCTTGG
GCAAAGTGTCAACAACCTGCACTCATTGTGTGACAGTGTCCCTTGCCACATCCACCCCT
CCCACGTTTACACACACCAGCAGAGAAATAGTTGTTGGAAGTGAGAAATGTGTGGCAATC
AGTCGCTCGGCTGCGCAGCGAGAAGAGGAAGCAATAACACAAACACGCTCCCTGCAACGT
GACTTCAAAGATGTATGTTGACTACTTTCTGTAAGTGTATGGGTGTTACGTACCGTCTT
TCACCTGAACTTTTAGTCTGCAACAGACGTGGATTATGCTGCATTTTAACTATGTGGCC
CCTCCAGGCTCTTCTGCCCTCAACAGAACCCTTACTGCAAAAACCACGAGAGAAACCAG
TGATGTAATGTTCCGCAAACAAGGAACAGTACAGCGATGACTAAGATTATTTAGTGAAAC
CGCTGTGAGCTTGTGTTCCCCAGTTAAGCCTTGTCTGTCTGCATACATACTTCAGTG
TTGAAACAGACTGTGCAACTGTAGGTTTAGTCAAACAGCGAGCAGCATGTGCTCTGTATT
CATGGGGCTGCCGTGAGACAACAACATTCCAATGACGCATTAAATTAGTCTAACAGTATT
GCGTAGAGAGCTGCGTCGAGCCGTATGAAATTGGCTGTTCACTCATAAAGCAAGTCAAAA
ATGTTTCAAGTTATGAAATCTTATCTGTAATCTCTCTCGTTTTTCTCCACAATCACAGTCA
AAGTCATGATGGTTTTCCAGGCTGCAGATATCGAAGTTTAGTTTTGGATCAGATAAGAAGT
GAACAAAACACAGCTTCCTCTTGGTGTTTACATTTGTGCTTTACTTGTGTTTCCAAAGAA
ATGGCCGAGGCCACCAGGCAGTGGCATTCCAGTTCACCATAACACCAGAGGGCATCGAT
-M--A--E--A--H--Q--A--V--A--F--Q--F--T--I--T--P--E--G--I--D--
CTGCAGCTGTCTTACCAGGCGCTGAACCAGATCTACCTGTCTGGAGTACGATCCTGGAAG
-L--Q--L--S--Y--Q--A--L--N--Q--I--Y--L--S--G--V--R--S--W--K--
AAGCGAGTCAGCCGCGTGAGGgtgagcatca-N₃₅₃₆-ctgctaacag**AACAGTCTGATTAA**
-K--R--V--S--R--V--R-- -N--S--L--I--K
GGGAGTGTACCCCTGCCAGCCCGTCATCCTGGCTCTTCGTTGCCATAGGAATCCTGGCCAC
--G--V--Y--P--A--S--P--S--S--W--L--F--V--A--I--G--I--L--A--T
CATGTACATGCAGTCGGATCCCAGCATGGGGCTCATCACCAGATACAGCAGCACCTGCC
--M--Y--M--Q--S--D--P--S--M--G--L--I--T--K--I--Q--Q--H--L--P
GCTAAGgtacacaacg-N₂₈₄₁-ctctctgtgcag**TCTCCACATGTCCCTCAGCACCCAGGG**
--L--S -L--H--M--S--L--S--T--T--Q--G
TCAGACTATGGTGTGACGCTGGTCTTCAGCACTCTGCTGTGGCTCTCTCTCATCTGGC
--Q--T--M--V--S--A--L--V--F--S--T--L--L--W--L--S--L--I--L--A
CCTCAGGTTCTGCCTCAAACCTGCTGCTGTCTTACCACCAGTGGATGTTTGAAAAGCACGG
--L--R--F--C--L--K--L--L--L--S--Y--H--Q--W--M--F--E--K--H--G
CCGGATCTCCAACACCACC**AAAGTCTGGGTG**gtaagccacg-N₈₈₀-tttctctcag**CCTC**
--R--I--S--N--T--T--K--V--W--V- -P--
TGGTGAGGTTGCTGTCAAGCAGGAAGCCTTTGCTGTACAGCTATCAGTCGTCTCTACCTC
L--V--R--L--L--S--S--R--K--P--L--L--Y--S--Y--Q--S--S--L--P--
ACCTTCCTGTTCCAGCCATCCAAGACACAGTGAGCAGGgtgagtgatga-N₁₀₃-ctgttatc
H--L--P--V--P--A--I--Q--D--T--V--S--R-
ag**TACTTGGAGTCGGTGCGCCCTCTGCTGACGGACCTGGAGTTTAAACGCATGACAGACC**
-Y--L--E--S--V--R--P--L--L--T--D--L--E--F--K--R--M--T--D--
TGGCCAATGAGTTTGAAGTCGAATCTGGGAAACC**DOCTODAAOGCTAOGTGA**ACTTAAAG
L--A--N--E--F--E--S--N--L--G--N--R--L--Q--R--Y--L--K--L--K--
CTCTGTGGGCCACCAACTATgtaggtattc-N₁₅₉₃-ttggctctag**GTGAGTGACTGGTGG**
A--L--W--A--T--N--Y- -V--S--D--W--W-
GAGGAATACGTCTATTTAAGGAGCCGAGGTCCCATAATGGTCAACAGTAACACTACTATGGA
-E--E--Y--V--Y--L--R--S--R--G--P--I--M--V--N--S--N--Y--Y--G-

ATGgtaaccactg-N₂₉₆-gtccttccagGATTTTCTGTACGTGACACCAACCTCAGTGCA
-M- -D--F--L--Y--V--T--P--T--S--V--Q
GGCGGCCAGGGCGGGGAACACGATCACTGCCCTGCTGCTCTACCGCCGCAAGGTGAACCG
--A--A--R--A--G--N--T--I--T--A--L--L--L--Y--R--R--K--V--N--R
GGAGGAACGAAGCCGgtgaggctgc-N₃₈₈₀-tgtgctgtagAGCCCGCTTCCGGGTACCG
--E--E--L--K--P- -S--R--V--P--G--T--
TCATCCCCCTTTGTGCTGCTCAGTGTGAGAGGATGTTCAACACAACACGCACACCAGGAG
V--I--P--L--C--A--A--Q--C--E--R--M--F--N--T--T--R--T--P--G--
TCGAGACTGgtaacaaaca-N₃₈₂-ttttccacagATGTGCTCCAGCACTGGCTAGACAGT
V--E--T-- D--V--L--Q--H--W--L--D--S-
GAGTTTGTGGCAGTGTACCACAGGGGGCGCTATTTTCGCCCTGTGGGTATACCGGGCAGGC
-E--F--V--A--V--Y--H--R--G--R--Y--F--R--L--W--V--Y--R--A--G-
CGCCTGCTGTCCCCCAGGGAGATTGAATACCAGATCCAGAGGATCCTGGATGACCTTTCA
-R--L--L--S--P--R--E--I--E--Y--Q--I--Q--R--I--L--D--D--P--S-
CCTGCTCTTCTGGAGAGGAGAACTGGGAGCTCTCACTGCTGGAGACAGgtacgtacaa
-P--A--L--P--G--E--E--K--L--G--A--L--T--A--G--D--R
-N₃₃₈₃-ccccctccagAATCCCTTGGGCTCAGATGAGGAAGCAGTACTTCAGCTCTGGGGT
--I--P--W--A--Q--M--R--K--Q--Y--F--S--S--G--V
CAACAAGAGATCGCTGGATGCCATTGAGAGGGCAGCCTTCTTTGTAACCTTGGATGACGA
--N--K--R--S--L--D--A--I--E--R--A--A--F--F--V--T--L--D--D--E
GGAGCAAGGCATGAGGGGAGATGACCCAGAGGGGAACTTGGACAGCTACGCCAAGTCTCT
--E--Q--G--M--R--G--D--D--P--E--G--N--L--D--S--Y--A--K--S--L
GCTTCATGGGAAATGTTATGACAGgtgagaggctc-N₃₀₄₇-cctctcccagGTGGTTTGATA
--L--H--G--K--C--Y--D--R --W--F--D--
AATCCTTCTCCGTTGTGATCTACAAGAATGGGAAGAGCGGGCTGAATGCAGAGCACTCCT
K--S--F--S--V--V--I--Y--K--N--G--K--S--G--L--N--A--E--H--S--
GGGCTGATGCACCAACAGTGGCCACCTTTGGGAGgtaatccttg-N₁₉₅₃-ttccctccag
W--A--D--A--P--T--V--A--H--L--W--E-
TACACACTGGCCACAGATGCCCTTTCAGCTAGGCTACACTGAGGACGGCCACTGTAAAGGT
-Y--T--L--A--T--D--A--F--Q--L--G--Y--T--E--D--G--H--C--K--G-
GAGGTGGAGCCTTCTCTACCCCAACCTCAGAGACTTGTCTGGAACATCCCGGCAGAGgta
-E--V--E--P--S--L--P--Q--P--Q--R--L--V--W--N--I--P--A--E-
ggaatgc-N₂₀₇₆₇-tgtgtttcagGTTTCAGGCTCAAGTCAGCAGTTCCTAGCTGTTGCAC
-V--Q--A--Q--V--S--S--S--L--A--V--A--
AGGCGCTGGCAGATGATGTTGACTGTACGTGTTTCCATTTCAGAGAGTTTGGCAAAGGAC
Q--A--L--A--D--D--V--D--C--H--V--F--P--F--R--E--F--G--K--G--
GAATCAAGAAATGCCGGATCAGCCCAGACGCCTTCATCCAGATCAGCTTACAACCTGGCAT
R--I--K--K--C--R--I--S--P--D--A--F--I--Q--I--S--L--Q--L--A--
ACTACAGGgtaacacaca-N₂₅₃₉-ctccttccagGATCGCAGTGGCTTTTGTCTGACTTAC
Y--Y--R- -D--R--S--G--F--C--L--T--Y-
GAGGCATCAATGACCCGCCTGTTTCAGGGAGGGCCGGACAGAGACTGTGCGTTCCCTGCTCC
-E--A--S--M--T--R--L--F--R--E--G--R--T--E--T--V--R--S--C--S-
AACGAGAGCTGCGCTTTCATCAAAGCTCTGGAGAGTGGAGAGgtactgaagt-N₅₅₀-gctt
-N--E--S--C--A--F--I--K--A--L--E--S--G--E-
tgacagACAGAGGAAGAATGCAGACGCCTTTTCCAACCTGGCTTCTGAGAAGCACCAGAAC
-T--E--E--E--C--R--R--L--F--Q--L--A--S--E--K--H--Q--N
CTTTATCGGATGGCTATGACTGGAGCAGGAATCGACAGACATCTCTTCTGCCTTTACGTG
-L--Y--R--M--A--M--T--G--A--G--I--D--R--H--L--F--C--L--Y--V-
GTCTCCAAATACTTGGGAGTGGAGTCGCCCTTTCCTCAAAGAAgtaagccttg-N₁₂₂₂-ctc
-V--S--K--Y--L--G--V--E--S--P--F--L--K--E-
atthaagGTTCTGTCTGAGCCTTGGCGTCTCTCCACCAGTCAGACTCCAGTTCAGCAGAT
-V--L--S--E--P--W--R--L--S--T--S--Q--T--P--V--Q--Q--M
GAATCTGTTTCGACCTGAAGAACCACCCAGATTTTCATCTCGCTTGGTGGGGGCTTTGGACC
--N--L--F--D--L--K--N--H--P--D--F--I--S--L--G--G--G--F--G--P
Tgtgagtgggc-N₉₀-gtcatgtcagGTTGCTGACGATGGTTATGGAGTTTCATACATCAT
- -V--A--D--D--G--Y--G--V--S--Y--I--I
CGTGGGTGAAGATATGGTCAACTTCCACGTGTCTTCCAAATACTCCTGCAGTCAGACTgt

--V--G--E--D--M--V--N--F--H--V--S--S--K--Y--S--C--S--Q--T--
aagtaacg-N₃₉₂₁-tccttaccagGACTCTCACAGGTTTGGAAATTCAGATAAGTAAGGCTA
-D--S--H--R--F--G--I--Q--I--S--K--A--
TGCAGGACATCATGGCTCTTCTCTCAGCTGACCCCAAACCTCTTCGTCATCCAAAGGCA
M--Q--D--I--M--A--L--L--S--A--D--P--K--T--S--S--S--S--K--G--
GAGGCCATCCGCAGACCAAGAAGCTCGAGTAAaatcaaatgtcggccgactatttcctta
R--G--H--P--Q--T--K--K--L--E--*-
atgattccgaaccacaaatggttcaagcacaggtcaacgagacattaattgacggagggaa
agtcggggccaaacaacagagagcttggttcgtagatacagatctttcgtcttctgcggt
ctttagacacaacaaacattaacagcagccttttacaggccaacagattaacataactac
tggtgtgtgcctttttttaaaacctcccactcagtgacaaaagataagccacgacgctcc
cgtacctcatccattcgggaagcctcttgtgca -3'

d) *cpt 1β*

5' agggctttcaataacagactaaaaatcttaagatatctgcaagatthttgaaatthttgtat
aataacaaagctgcagagctthttcttaattaaaattataaccagtaggtthtgcacacattt
cgctgttgatattctgtgtttatatttccatgcccgctctgcggattgcgcacatgctcagc
TATA Box +1
cgggtgtaccactatatctggccactgaactthtgcacacttGGTGACTGTTTCAGTCATAT
GACATGGAAAAAGAAATACTAGACCTGCTGCCCCAACCATTTATATGGACTGGTTTGGGA
AGTGCCGATATTTTTGTTAAGAATTACAGAAGGTCATCGTATTTCTCAGACAAAGACTGA
ACCTTTTCAACCTTCAGTGTATTTATTTTTCATTTAGACACCTGTTAACTTCTGTCAGT
GACCCACCTTATCACATTATGAGCAAACAGGACAAATAGAGCCACTTTCTTTGAAACTGC
CGTCAGTCACATCTGTGGAAGATTGTTTTTTTTCTTGAGGTTAACTGTTGAATGATCAGC
CTATTTGTTTGTATAACAACAGTGAATGAAATACCCAGTGTTTAGTCAAACCTGAGCATG
GCCAATGGGGAAAGAAATCATCTGGCATGATGAGGGAGTTCAGTGGCAAGTGTGAAAA
TCAGTCTGTAAGCCTCTCAGTGACGACCTTTCAATGGACCAGTAATAAGTAGGCATGATT
CAAAATTTGGATCTGTACAGATGTAGGTTATTTGTTCTGTGGAAAATAATTTAGAGTGTG
TTATCACAATGAAGTCGATACCCGTGACAACCGTAAATCTCTGAAACCTTTAAAAACCACT
ATTCAGATTTTCATGGTTTGGTGTCTGAATCCGGTGCATGAAACGTACATTTTTTACCAG
GCATCTGATGGTTGCCACGGGTGCAGACTCCCTCTAGGCCTGTGTGGGACTTTAAACAAC
AAAAGgtatgcattt-N₄₄₄-ttcatgttagCAGAACACAGAGATGGCTGAGGCACATCAG
-M--A--E--A--H--Q--
GCAGTAGGCTTCCAGTTCACCTGTGCGCCCAGATGGTGTGGACTTTAACTGAGCCAAGAA
-A--V--G--F--Q--F--T--V--R--P--D--G--V--D--F--K--L--S--Q--E--
GTCATCAAAAATATCTACCTATCAGGAGTGACAGCATGGAAGAAGAAAGCCATTCAATTC
-V--I--K--N--I--Y--L--S--G--V--T--A--W--K--K--K--A--I--Q--F--
AAGgtacaaaaac-N₆₉-atatccacagAATGGTGTATTGGCAGGAGTCTACCCTGCCAGT
-K--N--G--V--L--A--G--V--Y--P--A--S--
CCATCCAGTTGGCTAATAGTTGTGATTGCCATGATGAGCTCTTTATATATCCACATAGAC
-P--S--S--W--L--I--V--V--I--A--M--M--S--S--L--Y--I--H--I--D--
CCGTCTCTGGGAATGATGGATGCCATCAAGGAAAACCTACCATATCGgtcagtaatc-N₁₂₁
-P--S--L--G--M--M--D--A--I--K--E--N--L--P--Y--R--
-ttctgagcagAGACTGCATGTACAGTACAGACACGAGCAGTCCTGAGTGCCATCCTGTTT
--D--C--M--S--V--Q--T--R--A--V--L--S--A--I--L--F--
GCCACAGGACTGTGGCTGTTCCTCATCTACCTCCTGAGATACACTCTTAAAGCTTTGCTG
-A--T--G--L--W--L--F--L--I--Y--L--L--R--Y--T--L--K--A--L--L--
TCCTATCATGGATGGATTTTTTGAGTCCCATGGAAAAATGAGTACATCAACTAAAGTGTGG
-S--Y--H--G--W--I--F--E--S--H--G--K--M--S--T--S--T--K--V--W--
CTGgtatgttcgt-N₁₀₆-ttgttaccagTGCTTGGTAAAGATGTTTTCTGGACGCAGAC
-L--C--L--V--K--M--F--S--G--R--R--
CCTTGCTCTACAGTTTCCAGGCCTCCTTACCCAGGCCTCCTGTGCCCAGTGTGGATGATA
P--L--L--Y--S--F--Q--A--S--L--P--R--L--P--V--P--S--V--D--D--
CAATTCACAGGgtatgtagtt-N₁₄₆₅-ctthttcacagTACCTTGAGTCAGTTCGACCTC
T--I--H--R--Y--L--E--S--V--R--P--
TGCTGGACAATGAGCAGTACAATAAAATGGAGCTTTTGGCCAGTGAATTTAAAGAAAACA
L--L--D--N--E--Q--Y--N--K--M--E--L--L--A--S--D--F--K--E--N--
AAGCAGCCCAACTGCAGAGATGCCATACTTAAATCATGGTGGGCAACAAATTATgtaa
K--A--A--Q--L--Q--R--C--L--I--L--K--S--W--W--A--T--N--Y--
gtagtg-N₆₄-ttttctccagGTAAGTGACTGGTGGGAGGAGTACATTTACCTCAGGGGC
-V--S--D--W--W--E--E--Y--I--Y--L--R--G--
AGGAGTCCTATCATGGTCAACAGCAACTTCTATATAATGgtatgtaaaa-N₁₇₆-ttgtg
-R--S--P--I--M--V--N--S--N--F--Y--I--M--

tgaagGACCTCTTGTATGTGACCCCAACACACCGTCAGGCTGCAAGAGCAGGAAATGTGG
-D--L--L--Y--V--T--P--T--H--R--Q--A--A--R--A--G--N--V--
TACATGCCATGTTGCAATACAGACGCAAACCTGGAACGTGGTGAACCTGCACCGgtatagc
V--H--A--M--L--Q--Y--R--R--K--L--E--R--G--E--L--A--P--
ttg-N₆₈-ttctgctcagCTGAGGGCTTTGGGGACTGTTCTATGTGTTCTACTCAGATGG
-L--R--A--L--G--T--V--P--M--C--S--T--Q--M--
AGAGGATGTTTAAACCACTCGTATCCCTGGCATTGAAACAGgtgagcaaat-N₂₁₈-at
E--R--M--F--N--T--T--R--I--P--G--I--E--T--
ctttacagATTTTGTGCAGCACCTGACTGATCGGAAACACCTGGTGTCTTCCACAAGGG
D--F--V--Q--H--L--T--D--R--K--H--L--V--V--F--H--K--G
TCGGTTCTTCCAGGTGTGGCTTTACACTGGAGGGCGTCACCTTTTGCCAGTGAACCTTGA
--R--F--F--Q--V--W--L--Y--T--G--G--R--H--L--L--P--S--E--L--E
GACACAGTTTCAAAGGATTCTCAATGATACATCAGAACCCTCAACCAGGAGAAGTCAAACCT
--T--Q--F--Q--R--I--L--N--D--T--S--E--P--Q--P--G--E--L--K--L
AGCTGCCCTGACAGCGGGATACAGgtatatgcaa-N₈₁-tcacccttagGATTCCATGGG
--A--A--L--T--A--G--Y--R-- --I--P--W--
CTCAGGCTCGGATAAAATATTTTAGCCAGGGCATAAACAAAGTATCCCTGGATGCAATCG
A--Q--A--R--I--K--Y--F--S--Q--G--I--N--K--V--S--L--D--A--I--
AGTCAGCTGCCCTTCTTCCCTGACACTGGATGATGAGCCGAGGGTTATGATCCTGCAAAGA
E--S--A--A--F--F--L--T--L--D--D--E--P--Q--G--Y--D--P--A--K--
GCAACTCACTTGATAGCTACGCCAAGTCCCTGCTGCATGGAAAATGTTATGACAGgtagg
S--N--S--L--D--S--Y--A--K--S--L--L--H--G--K--C--Y--D--R
caaag-N₅₆-tttcttgtagGTGGTTTGACAAATCTTTTACATTGATCTCTTATCCAAAT
--W--F--D--K--S--F--T--L--I--S--Y--P--N--
GGAAAAATGGGTGTAAATGTTGAACATTCTTGGGCTGATGCACCAATTGTAGGACACATG
-G--K--M--G--V--N--V--E--H--S--W--A--D--A--P--I--V--G--H--M--
TGGGAGgtaggaaatg-N₆₄-ttcttttttagTATATTCTAGCAACAGACTGCTTCCATCT
-W--E-- -Y--I--L--A--T--D--C--F--H--L
TGGATACACTGAGGAGGGACACTGTAAAGGGGATGTGAACAAAAACCTGCCTCATCCAC
--G--Y--T--E--E--G--H--C--K--G--D--V--N--K--N--L--P--H--P--T
TCGACTACAATGGCAGATTCCAAATGAGgtatagcaaa-N₁₀₀-cctctctaagTGCCAA
--R--L--Q--W--Q--I--P--N--E-- -C--Q--
AATGTCATTGAAACATCATACCTCTCAGCCAAGCAGATAGCTGACGATGTGGACTTCCAT
-N--V--I--E--T--S--Y--L--S--A--K--Q--I--A--D--D--V--D--F--H--
GGCTATCTGTTTGTGAGTTTGGGAAAGGCCGTGATCAAGAAGTGCAGGACAAGCCCTGAT
-G--Y--L--F--A--E--F--G--K--G--L--I--K--K--C--R--T--S--P--D--
GCCTTTATTTCAGCTGGCCTTGCAGCTGGCCCAGTTTAGGgtaggcctgc-N₆₇-tccttt
-A--F--I--Q--L--A--L--Q--L--A--Q--F--R--
atagGACCAGCGTGTGTTCTGTCTGACGTACGAGTCATCGATGACACGTATGTTTCAGAGA
-D--Q--R--V--F--C--L--T--Y--E--S--S--M--T--R--M--F--R--D
TGGGCGGACAGAGACAGTACGCTCCTGCATTCTGAGGCGGTTCGATTCTGTCAGAGCCAT
--G--R--T--E--T--V--R--S--C--T--S--E--A--V--A--F--V--R--A--M--
GGAGGACGCTGGTGGCAGgtaggcaaaa-N₄₃₉-atgttgtcagAATGCCAGAGACTT
--E--D--A--G--A--T-- -N--A--Q--R--L--
GCCCCTGTTTCGGAAAGCAGCAGAAAAGCACCAAAACATGTATCGTCTGGCCATGACTGGT
-A--L--F--R--K--A--A--E--K--H--Q--N--M--Y--R--L--A--M--T--G--
TCTGGGATCGATCGTCACCTTTTCTGTCTCTATATTGTATCTAAATACCTTGGTGTGCAC
-S--G--I--D--R--H--L--F--C--L--Y--I--V--S--K--Y--L--G--V--D--
TCACCATTCTTACAAAGgttagcacta-N₇₀-ttgcttaaagGTGCTTTCGGAGCCCTGG
-S--P--F--L--T--K-- -V--L--S--E--P--W--
AAGTTGTCCACCAGCCAGACTCCACAGCAGCAGCTTAATTTAGTTGACATCAACAAGTTC
-K--L--S--T--S--Q--T--P--Q--Q--Q--L--N--L--V--D--I--N--K--F--
CCGAAATATGTGGGTGGTGGAGGTGGATTTGGCCCTgtgagttaat-N₉₅-tccttttta
-P--K--Y--V--G--G--G--G--G--F--G--P--
gGTAGCTGACGATGGTTATGGTGTGTCTTATATCATTTGTTGGGGAAAACCTCATCACATT
-V--A--D--D--G--Y--G--V--S--Y--I--I--V--G--E--N--L--I--T--F--
CCACATCTCCAGCAAGTCTCCAGTCTGACACAgtaaagacaga-N₇₉-tgacctgcagG

--H--I--S--S--K--F--S--S--P--D--T--
ACTCATATCGGTTTGGTCAGCATATTCGAAAGGCCATGCTTGACATCCAAGCTCTTTTCA
D--S--Y--R--F--G--Q--H--I--R--K--A--M--L--D--I--Q--A--L--F--
AGCCTGAAAATGATAAGACAGCACAGAATGCAAAGTATGTCCATCTGGAAAATGGGAAA
K--P--E--N--D--K--T--A--Q--N--A--K--Y--V--H--L--E--N--G--K--
AGCACATATAAagagagtggtgggaggatatgttggatttggagcacatactgcctgtga
K--H--I--*--
ctcagtccttcattcataaaagtaaaacaaggttctggtgacaggatgggctgctgtggtg
agtcacagtacaggagcttggcattcctcagacaaaagcaatatttattacacatctatg
tttttcctgtttatatttgggggaaacaaatgtcatttgttttacacaataggtcctcac
atacacgtaccacactcactgtatataaaagccaatgtaatgtaataaaagctacaatatg
tatttattcatgatcttatgaaagctgtaaatatatctgc -3'