Supplemental Material to:

Ivan Dotu, Gloria Lozano, Peter Clote,
and Encarnacion Martinez-Salas

Using RNA Inverse Folding to identify IRES-like structural subdomains

2013; 10(11)
http://dx.doi.org/10.4161/rna.26994

www.landesbioscience.com/journals/rnabiology/article/26994/

Supplementary information

Supplementary Table 1. Hits retrieved using RNA Inverse Folding following pipeline filters

GenBank accesion number	Organism	Gene	Region	Nt length	E-value	Nt Id* (%)	Function
NM_079437.3	*Drosophila melanogaster*	TAF6	Coding region	36	2.7	79	Transcription factor TFIID assoc. protein 6
NM_214436.1	*Sus scrofa*	ST6GAL NAC-V	Coding region	43	0.016	91	Adds sialic acid in α2,6 linkage to GalNAc
NM_001205770.1	*Bos taurus*	TNFSF11	Coding region	38	0.97	88	Tumor necrosis factor ligand superfamily member 11
NM_007496.2	*Mus musculus*	Zfhx3	5’ UTR	38	0.97	84	Zinc finger homeobox protein 3
NM_177289.1	*Mus musculus*	Cbfa2t3	5’ UTR	36	2.7	85	Core-binding factor, runt domain, alpha subunit 2, translocated to 3
AY005471.1	*Rattus norvegicus*	bZip TF c-maf	5’ UTR	36	2.7	76	bZip transcription factor c-maf

* Nt Id: nucleotide identity

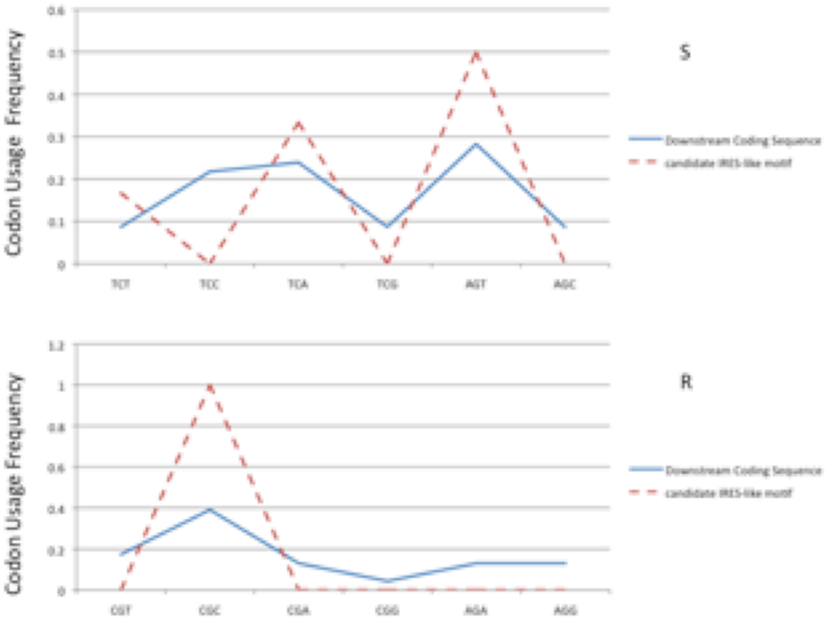
> Sus scrofa beta-galactosamide α2,6-sialyltransferase (ST6GALNAC-V) NM_214436.1

cagttttaattctgtctctaatctgcaacagcagtgcttcctgggtcccgcggctcccctgcctactcag
ccgccgctggccaccgggcgaaaatcagagccgcctccgccccatcacccaccatggaaaccctccaaga
aaaagtgaccccggacgcgtgtgcctgaggattccgcacaaaagaggtgcccaaa**atg**$_1$aagaccctggtg
cgccatggtctggcagtgtgtttagcgctcaccaccatgtgcaccagcttgttgctcgtatacagcagcc
tcggcggccagaaggagcggcccccgcagcagcagcagcagcagcagcaggcagcg**gcga**ctggcagcgc
gcagcggtcggcg**gaga**gcagcc*ccccgccgcgccccga****aacccc****gcgggaccg*cggccgctggaggga
taccttggcgtggcggaccacaagcccctgaaa**atg**$_2$cactgcagggactgtgccctggtgacgagctcag
ggcatctgctgcgcagccaccaaggcccccagattgacaagacggagtgtgtcatccgc**atg**$_3$aacgatgc
cccgacgcggggctacgggcgagacgtgggcaaccgcacgagcctgagggtcatcgcgcattccagcatc
cagcgcatcctgcgcaaccgccacgacctgctcaacgtcagccagggcaccgtgttcatcttctggggcc
ccagcagctacatgcggcgcgacggcaagggccaggtctacaacaacctgcagctcctgagccaggtgct
gccccggctgaaggccttcatgatcacgcgccacaagatgctgcagtttgatgagctcttcaagcgcgag
accggcaaagacaggaagatctccaacacttggctcagcactggctggttcaccatgacaattgccctgg
agctctgtgacaggatcaatgtgtatggcatggtgcccccagacttctgcagggatcccaatcacccttc
agtaccttatcactactacgaacctttggacccgatgaatgcacaatgtacctctcccatgagcgaggc
cggaagggcagtcatcaccgcttcatcacagagaaacgcgtctttaagaactgggcacggacattcaata
ttcactttttcaaccagactggaaaccagactcactggctgtaaatcatcctgagaacaaacccgtgtt
ctgaggnatgtgtgtagtcaggctgaattcttggcctcga

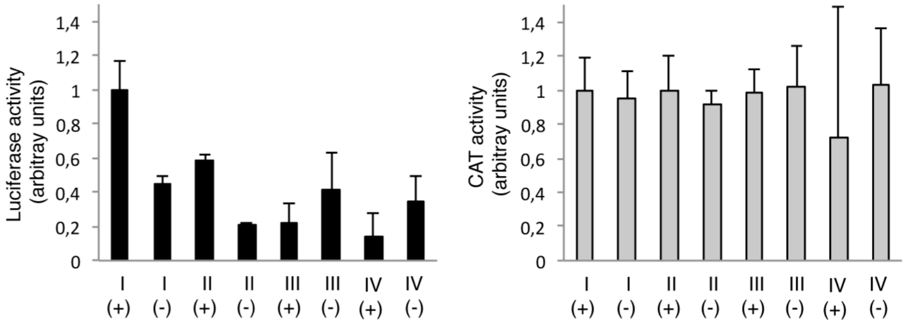
> Bos taurus tumor necrosis factor (ligand) superfamily member 11 (TNFSF11) NM_001205770.1

atg$_1$cgccgcgccagcagagactacaccaagtacctgcgcggctccgaggagatgggcggcgcggcgccc
cgcacgagggcccgctgcacg*ccccgccgccgcccgcaacccccgcggccgcc*gccgcctctcggtccac
gttcgtggccctcctggggctgggactgggccaggtcgtgtgcagcgttgccctgttcctctacttccga
gcgcag**atg**$_2$gatcctaacagaatatcagaagatgacactcactgcattaatagaattttcaagctccatg
aaaacacggatttgcaagacacaactctggagagtcaagactcaaaattaatacctgattcgtgtaagag
aattaaacaggcttttcaagcagctgtgcagaaggaa**atg**$_3$caacatattgttaaatcacagcacattaga
gcagaaaaagcc**atg**$_4$gtggaaggttcgtggttagatctggccagaaggaacaaacctgagattcagccct
ttgcccatctcacgattaacgccactgacatcccatcaggttcccacaaagtgactctgtcctgctggta
ccatgaccgaggttgggccaagatctccaacatgaccttcagcaatgggaaactgatagttaaccaagac
ggcttctattacctgtacgccaacatctgctttcgacatcatgagacttcgggaaacctcagtgcgaatt
atcttcagctgatggtatatgtcaccaaaaccagcatcaaaatccccagttctcacactctgatgaaagg
aggaagcaccaagtactggttggggaattctgaattccatttttactccataaacgtgggaggattttt
aagctacggtctggcgagaaaataagcattgaggtctccaacccttcactactggatccagatcaagatg
caacatactttggggcttttaaagttcgagatatagattga

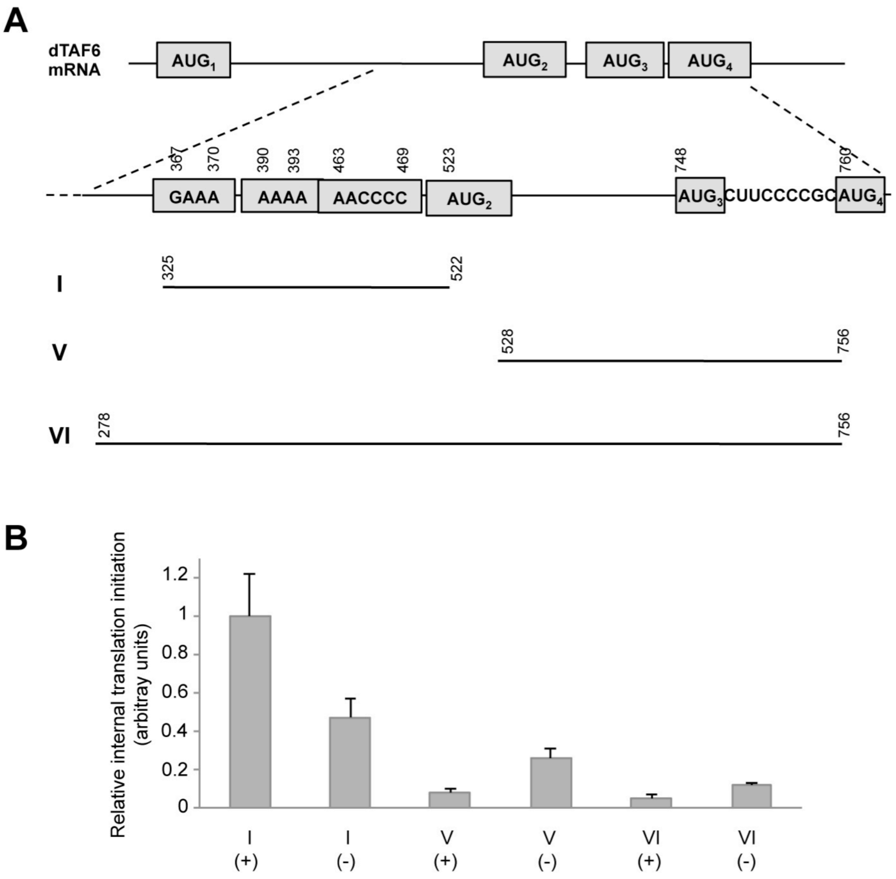
Supplementary Figure 1. Sequence of hits found in Sus scrofa β-galactosamide-sialyl transferase and Bos taurus tumor necrosis factor retrieved by the RNA Inverse Folding search. The sequence retrieved with RNA Inverse Folding is indicated by italics letters; motifs conserved with the IRES subdomain are indicated by bold letters; ATGs are depicted by yellow (ATG1), blue (ATG2), green (ATG3), or pink (ATG4) boxes.



Supplementary Figure 2. Codon usage distributions for region 325 to 522. X axis correspond to each synonymous codon and Y axis plots relative frequency of the codons in the region. IRES is used to indicate the candidate IRES-like motif and Protein for the corresponding downstream coding sequence. Upper graph is for serine (S), lower graph is for arginine (R).



Supplementary Figure 3. Luciferase and CAT activities measured with bicistronic constructs bearing regions I, II, III, or IV in either sense (+) or antisense (-) orientation observed in BHK-21 cells transfected with the corresponding bicistronic constructs. Values (mean±sd) were made relative to that of region I(+). Experiments were done in triplicate wells in at least three independent assays.



Supplementary Figure 4. (A). Schematic representation of *Drosophila melanogaster* TAF6 mRNA, with indications of regions I, V and VI. Nucleotides numbers are referred to the A of AUG1. (B). Relative internal initiation of translation of regions V and VI in either sense (+) or antisense (-) orientation relative to region I observed in BHK-21 cells transfected with the corresponding bicistronic constructs. Values represent the mean±sd obtained in at three independent assays.