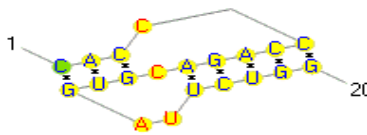


Covariant/Substitution Structures of a Ribosomal-Like Pseudoknot

October 2010 RNAParse.com
James F. Lynn

The 20 nucleotide rna pseudoknot CACCCCAGACGTGATTCTGG (((: [[[[(:)) : :]]]]) was determined by context-sensitive parsing techniques of the genome: AY350569.1 Avian leukosis virus, as described in www.rnaparse.com. Structural energy verifications were performed using KnotSeeker <http://knotseeker.csse.uwa.edu.au/>. A NCBI BLAST search revealed several exact matches.

CACCCCAGACGUGAUUCUGG RNAParse# RP0640
(((: [[[[(:)) : :]]]])



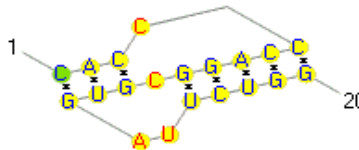
RP0640

Found in genomes:

>gi|493011|gb|M10455.1|AC SUR2CG UR2 sarcoma virus, complete genome
>gb|AY350569.1| Avian leukosis virus LR-9 gag protein, gag/pol polyprotein, and envelope protein genes, complete cds

Covariant matches are found at position 9 and 16 (Highlighted below) at the end of stem 2 that indicate non-Watson Crick pairing with a uridine in the corresponding position of stem 2. The pseudoknot's structure maintains its original structure according to KnotSeeker with many homologs found via NCBI BLAST.

CACCCCAG^GCGUGAU^UCUGG RNAParse #RP0641
(((: [[[[(:)) : :]]]])



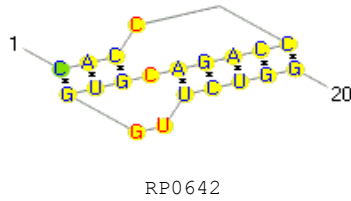
RP0641

Found in genomes:

>gb|M21954.1|HAMRSVPR1 Hamster proviral 5' LTR protein kinase (pp60v-src), 5' end
>emb|Y00302.1| Avian sarcoma virus CT10 genomic sequence
>emb|X52822.1| M.auratus H-19 proviral sequence (LTR,v-3scr1,LTR)
>emb|X51863.1| Avian Sarcoma Virus PR2257T Length=4503
>emb|X68524.1| Rous Sarcoma Virus (duck adapted) sequence of gag, pol, env and src genes Length=9312
>emb|X15345.1| Hamster H-19 proviral DNA (LTR- v-src -LTR)
>emb|V01197.1| Rous sarcoma virus genome, strain Prague C (Pr-C)
>emb|X51860.1| Duck adapted Rous sarcoma virus genes for gag, pol, (p19, p10, p27, p12, and p15 proteins) and 5' LTR
>gb|AF033810.1|AF033810 Fujinami sarcoma virus, complete genome
>gb|L21974.1|ACSCSRC Avian sarcoma virus proviral c-src gene, complete cds, and env gene, untranslated region
>gb|M21526.1|ALRSRCAC Rous sarcoma virus defective mutant PR2257 src gene encoding p66-src protein, complete cds

A third KnotSeeker-stable homolog was discovered via NCBI BLAST consisting of a single nucleotide substitution in position 14, loop 3 (Highlighted below.)

(((:[[[[[:]])]::]]]])
CACCCAGACGUGGUUCUGG RNAParse# RP0642



Found in genomes:

>gb|K03377.1|ALRGENVM Rous sarcoma virus (transformation defective B77 strain) env mRNA, slice-junction region

Discussion

This experiment was intended to test the effectiveness of context-sensitive grammars written by the author to parse the possible structures of certain small ribosomal pseudoknots. Therefore, there is no guarantee that the pseudoknots discovered and mentioned here actually exist; grammars are not physical methods of structure verification but strongly determine if a structure *can* exist in a given string of nucleotides.

Further work will include modifications of the grammars to include slightly-more variation of the pseudoknots mentioned here as well as greatly different nucleotide configurations of the same structure discovered at parse time such as exemplified below.

(((:[[[[[:]])]::]]]])
GAGCTGCCTACTCAGAGGCA (Additional pseudoknot included in a loose parse of the original genomes.)

Found in:

- >gi|297718719|gb|HM104174.1| Enzootic nasal tumour virus of goats strain ENTV-SC, complete genome RNA linear
- >ref|XM_002749140.1| PREDICTED: Callithrix jacchus abhydrolase domain-containing protein 2-like mRNA
- >gb|EF680318.1| Ovis aries strain enJSRV-24 endogenous virus Jaagsiekte sheep retrovirus, complete sequence
- >gb|EF680308.1| Ovis aries strain enJSRV-14 endogenous virus Jaagsiekte sheep retrovirus, complete sequence
- >gb|EF680302.1| Ovis aries strain enJSRV-20 endogenous virus Jaagsiekte sheep retrovirus, complete sequence