The SL RNA delimitation based on sequence conservation is not a trivial approach and the structure prediction of several known SLs can vary dramatically by the inclusion/removal of as little as a single base. Particularly among the Cestode sequences. To Illustrate this effect, we picked three SLRNAs to showcase these effects. Each section shows an illustration of the sequence, followed by the unpaired bases constraints (the SM-like site) and the predicted structure by Maximum Expected Accuracy (MEA), Minimum Free Energy (MFE) and Ensemble Free Energy (EFE). Added bases are highlighted in green, removed in red, the SM-like site that was constrained as unpaired is underlined.

1. Addition of 1 base at the 5’ of SL-RNA Uniqe\_SL-41, locus Hdiminuta\_Def\_Loci-7

Imagen que contiene collar

Descripción generada automáticamente

A

>Unique\_SL\_41 -- Hdiminuta\_Def\_Loci-7 (plus 1 base 5')

ACCGAGUAUUCGGUCUUACUAUACGAACUUGUAUGGUGAGUAUCGUGACAGCUCUCGUUUUGCUUGAGCUGUAGAGUUUUUGGCUGGUCCGCGAGGGCC

(((((....)))))(((((((((((....))))))))))).......(((((((..((...))..))))))).............(((((....))))) MEA

(((((....)))))(((((((((((....))))))))))).......(((((((..((...))..))))))).............(((((....))))) MFE

{((((....))))}(((((((((({....})))))))))).......(((((((..({...})..))))))).............(((((....))))) EFE

B

>Unique\_SL\_41

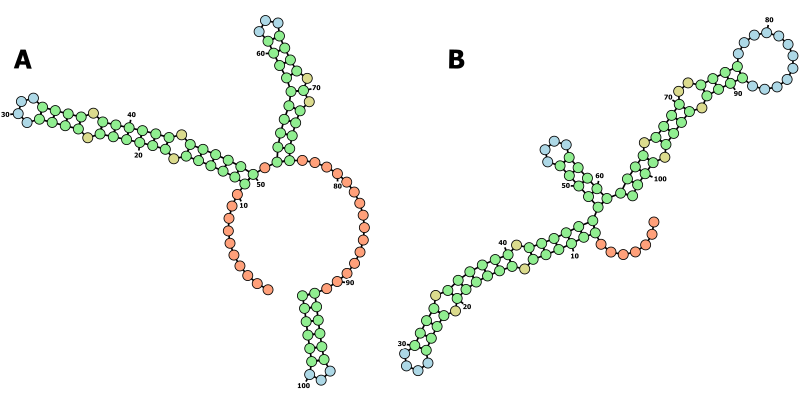
CCGAGUAUUCGGUCUUACUAUACGAACUUGUAUGGUGAGUAUCGUGACAGCUCUCGUUUUGCUUGAGCUGUAGAGUUUUUGGCUGGUCCGCGAGGGCC

.........(((((((((((((((....))))))))))).))))..(((((((..((...))..))))))).............(((((....))))) MEA

.........(((((((((((((((....))))))))))).))))..(((((((..((...))..))))))).............(((((....))))) MFE

,,,,....,|||{(((((((((({....})))))))))).}}}}..(((((((..({...})..))))))).............(((((....))))) EFE

This modification induces that portion of the leader into a small hairpin that compromises the first hairpin (B), like SL-4 described in our previous work (Calvelo et al., 2023). Matches Outlier5 conformation.

1. Addition of 3 base at the 5’ of SL-RNA Uniqe\_SL-8, reference sequence AJ292365.1

A

>Unique\_SL\_8 -- Reference SL Echinococcus multilocularis AJ292365.1 (plus 3 bases 5')

GUCACCGUUAAUCGGUCCUUACCUUGCAGUUUUGUAUGGUGAGUAUCGAUGCAGCUGAGGCUGUGCCUACGAGCUGACCCAGUAUUUGGCUGGUCCUUCGAGGGCC

..........((((((.((((((.((((....)))).)))))).)))))).((((((((((...)))).).)))))...............((((((...)))))) MEA

..........((((((.((((((.((((....)))).)))))).))))))(((((....)))))((((.((((..((((............)))).)))).)))). MFE

..........((((((.((((((.((((....)))).)))))).)))))),(((({,{{||,,,||||.|{||}}|,{,............||||{|,,,}))))} EFE

B

>Unique\_SL\_8

ACCGUUAAUCGGUCCUUACCUUGCAGUUUUGUAUGGUGAGUAUCGAUGCAGCUGAGGCUGUGCCUACGAGCUGACCCAGUAUUUGGCUGGUCCUUCGAGGGC

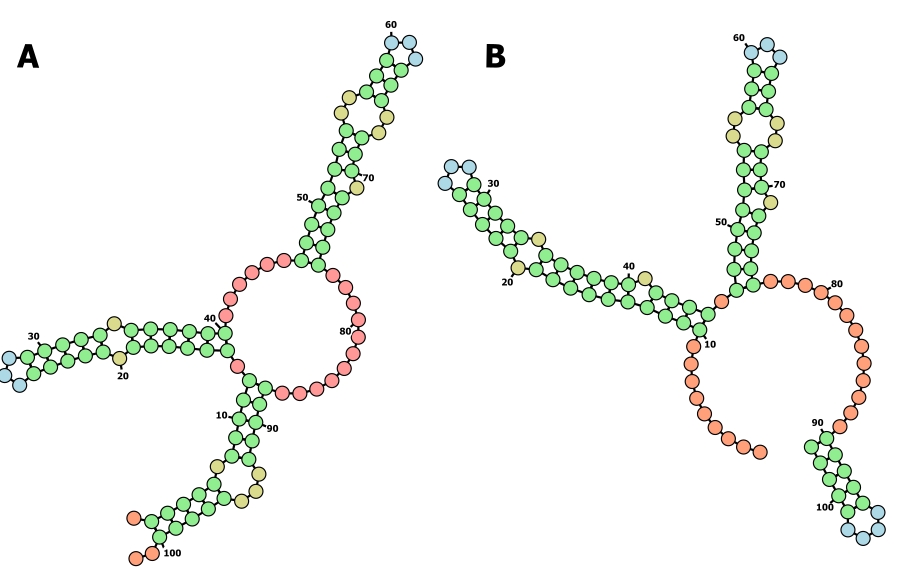
.......((((((.((((((.((((....)))).)))))).))))))(((((....)))))((((.((((..((((............)))).)))).)))) MEA

.......((((((.((((((.((((....)))).)))))).))))))(((((....)))))((((.((((..((((............)))).)))).)))) MFE

.......((((((.((((((.((((....)))).)))))).))))))(((((....)))))((((.((((..((((............)))).)))).)))) EFE

The addition of 3 bases on the 5’ of Uniqe\_SL-8, as reported in the reference sequence AJ292365.1 (A), restores the structure of Uniqe\_SL-8 (B). However, the MEA and MFE predictions remain contradictory. This matches Outlier1, 3 and 4.

1. The removal of a single base at the 3’ of Uniqe\_SL-44, locus Hmicrostoma\_Def\_Loci-6.



A

>Unique\_SL\_44 (minus 1 base 3')

ACCGUUUACCGGUCUUACCUUGCAAUUUUUGUAUGGUGAGUAUCGAUACAGCGCUUGGCUUUGCCUUGUGUCUGUAAAAUUCUUUGGCUGGUCCUAACGGGC

.(((((.(((((.((((((.(((((...))))).))))))......((((((((..(((...)))..))).)))))...........)))))...))))).. MEA

.(((((.(((((.((((((.(((((...))))).))))))......((((((((..(((...)))..))).)))))...........)))))...))))).. MFE

.(((((.(((((.((((((.(((({...})))).))))))......((((((((..(((...)))..))).)))))...........)))))...))))).. EFE

B

>Unique\_SL\_44

ACCGUUUACCGGUCUUACCUUGCAAUUUUUGUAUGGUGAGUAUCGAUACAGCGCUUGGCUUUGCCUUGUGUCUGUAAAAUUCUUUGGCUGGUCCUAACGGGCC

.........((((((((((.(((((...))))).)))))).)))).((((((((..(((...)))..))).))))).............(((((....))))) MEA

.........((((((((((.(((((...))))).)))))).)))).((((((((..(((...)))..))).))))).............(((((....))))) MFE

.........((((((((((.(((({...})))).)))))).)))).((((((((..(((...)))..))).))))).............(((((....))))) EFE

This single removal of a single base at the 3’ end causes the circularization of the SL-RNA (A) when compared to the original sequence (B), leading to the compromise of the first in the structure prediction. This matches Outliers2, 7, 8 and 9. And arguably Outlier6.

Calvelo, J., Brehm, K., Iriarte, A., & Koziol, U. (2023). Trans-splicing in the cestode Hymenolepis microstoma is constitutive across the life cycle and depends on gene structure and composition. *International Journal for Parasitology*, *53*(2), 103–117. https://doi.org/10.1016/j.ijpara.2022.11.006