

MC-Cons: RNA 2D structural assignment

by

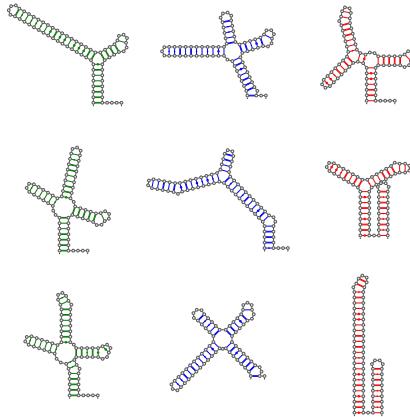
Gabriel C-Parent

outline

- ✦ what?
- ✦ why?
- ✦ how?
- ✦ what's next?
- ✦ Q&A

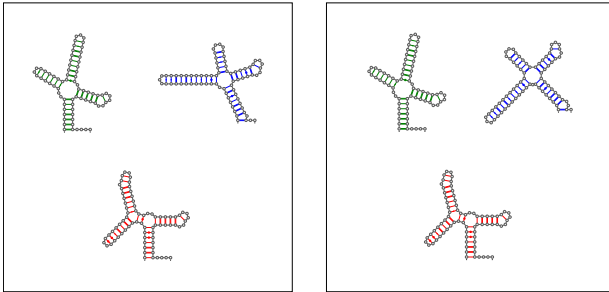
what?

so you have structures...



Given secondary structures of functionally related RNAs, which ones might explain the function?

now you can have consensus!

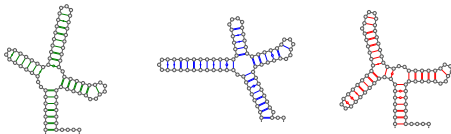


Let's suppose that their function is explained by a subset of similar structures (consensus).

In this case, 2 consensus make sense. Sometimes you get more, sometimes you get less.

why?

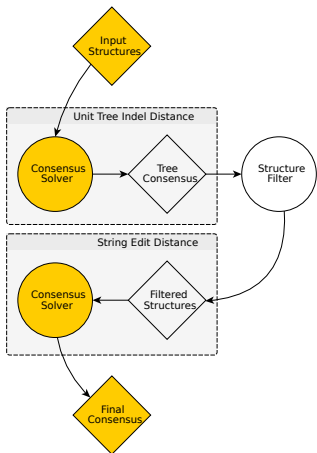
what's the point?



- goal: finding common/similar structures within suboptimal secondary structures
- useful for
 - improving secondary structure prediction for functionally related RNAs
 - exploring shared alternative structures

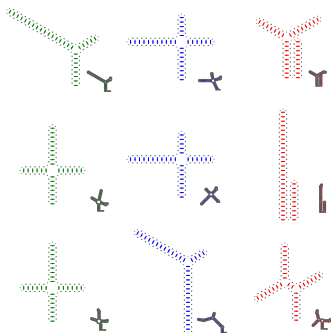
how?

workflow



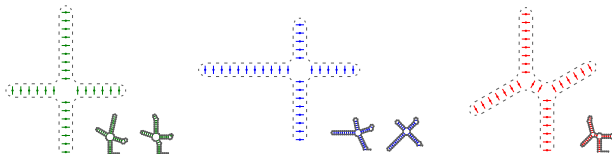
- ❖ a consensus minimizes the sum of pairs of distances between objects
- ❖ optimization steps
 1. base pair structural similarity (modified tree edit distance)
 2. string edit distance on Vienna dot-bracket representation
- ❖ two solver in C++ (branch & bound and genetic algorithm)

first match base pairs



- ❖ base pair trees created (many-to-one function)
- ❖ unit tree indel distance calculates number of base pairs to add or remove to transform a tree into another

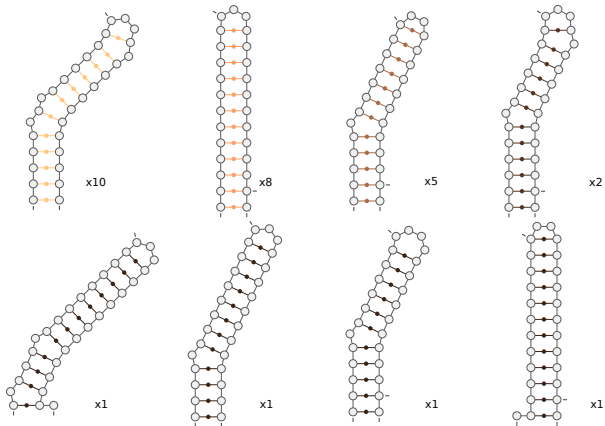
then match unpaired



- filtered by tree structure, now refine
- placing the bulges around the skeleton...
- string edit distance on dot-bracket

examples

IREs



IREs structural assignment given 10 suboptimal for 29 different IRES.

tRNAs

what's next?

future work

- ✚ MC-Cons improvements
 - ✚ parallelization
 - ✚ better distance functions?
 - ✚ heuristic improvement?
- ✚ derivative works
 - ✚ base pair tree constraints/masks
 - ✚ 3D to 2D tree distance

recap

- MC-Cons creates RNA 2D structural assignment
- useful for finding structures explaining function
- available at
<https://github.com/major-lab/MC-Cons>

Q&A

`https://github.com/major-lab/MC-Cons`