

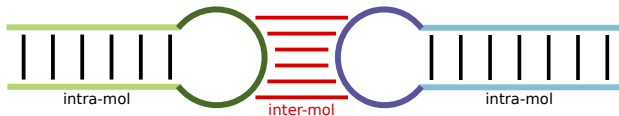
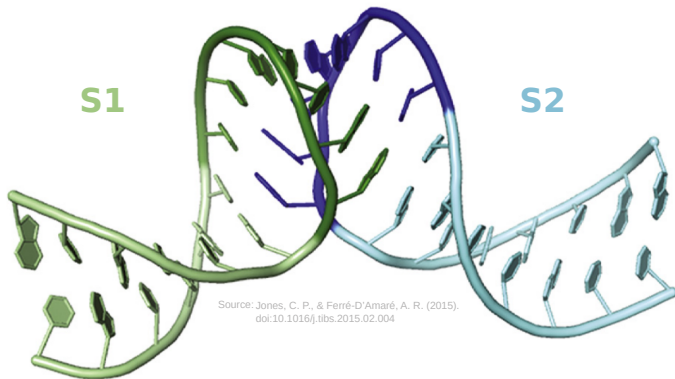
Constraint Maximal Inter-molecular Helix Lengths within RNA-RNA Interaction Prediction Improves Bacterial sRNA Target Prediction

Rick Gelhausen Sebastian Will Ivo L. Hofacker Rolf Backofen Martin Raden

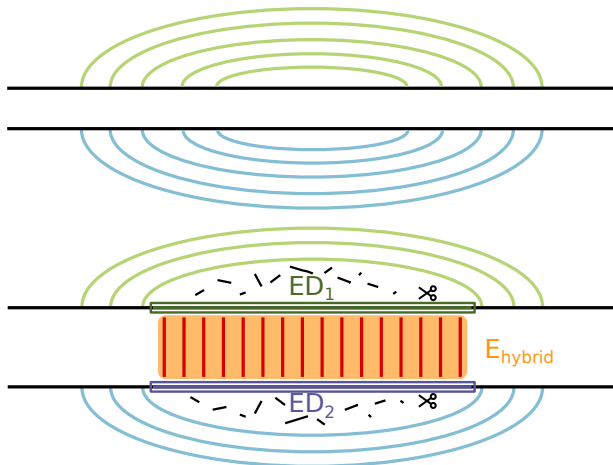
15.02.2019



RNA-RNA Interaction

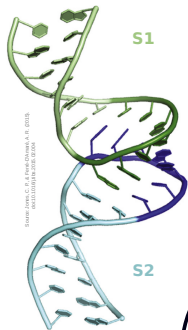


The underlying energy model

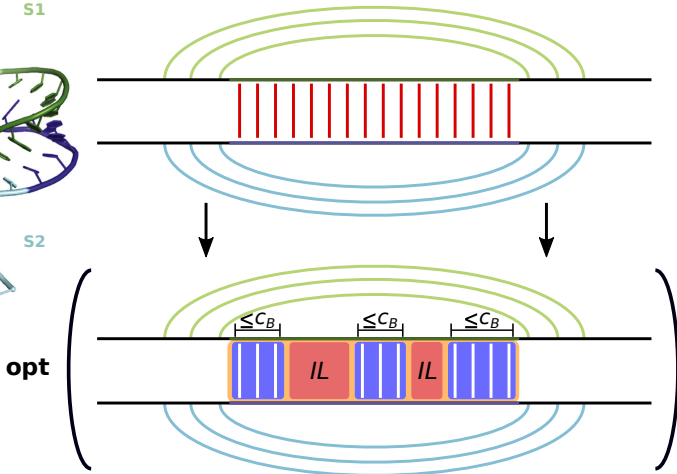


$$E_{\text{total}} = E_{\text{hybrid}} + ED_1 + ED_2 + \dots$$

Length restriction increases flexibility



Source: Jones, C. P. & Freese, D. W. A. R. (2015)
 DOI:10.1038/nrn3756



c_B = maximum allowed base-pairs per helix
 IL = interior loop



New method constrains maximum helix length

Step 1:

helix precomputation:

$$\begin{array}{c} i \quad k \\ \boxed{\text{helix}} \\ j \quad l \end{array} = \min \left\{ \begin{array}{l} +\infty \\ \text{opt. energy of helix} \leq C_B \end{array} \right.$$

Step 2:

interaction optimization:

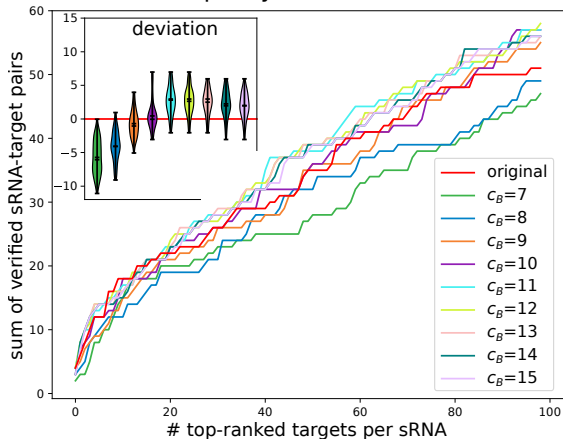
$$\begin{array}{c} i \quad k \\ \boxed{H} \\ j \quad l \end{array} = \min \left\{ \begin{array}{l} \begin{array}{c} i \quad k \\ \boxed{\text{helix}} \\ j \quad l \end{array} + E_{\text{init}} \\ \min_{pqr s} \begin{array}{c} i \quad p \quad r \quad k \\ \boxed{\text{helix}} \quad \boxed{E_{IL}} \quad \boxed{H} \\ j \quad q \quad s \quad l \end{array} \end{array} \right.$$

=> additionally a seed constraint can be incorporated

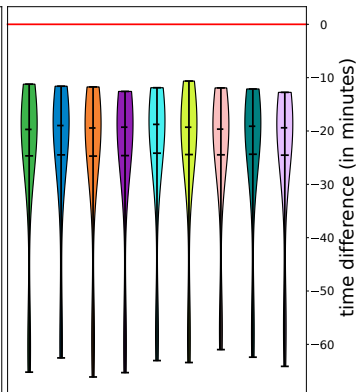


Prediction quality for different maximum helix lengths

quality assessment



time consumption

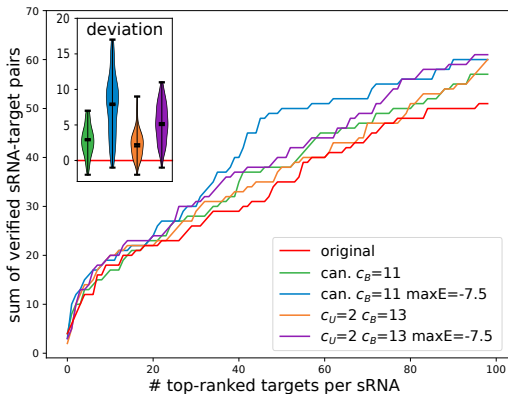


overall best results achieved for $c_B = 11-13$



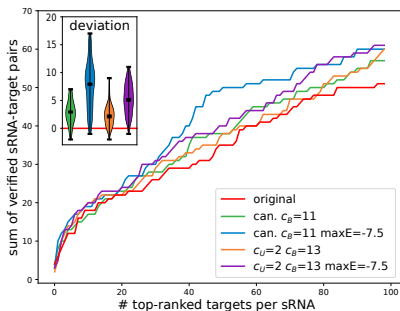
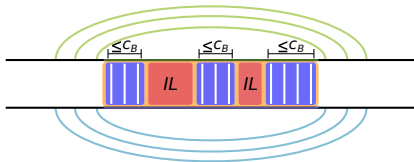
Further improvements are possible

- c_U : maximum size of bulges and interior loops
- maxE: maximum allowed energy of a helix



Summary

- restriction on the maximum helix length
- investigation of different parameterisations
- quality \uparrow runtime \downarrow
- Next: investigate a possible correlation between c_B and E_{max}



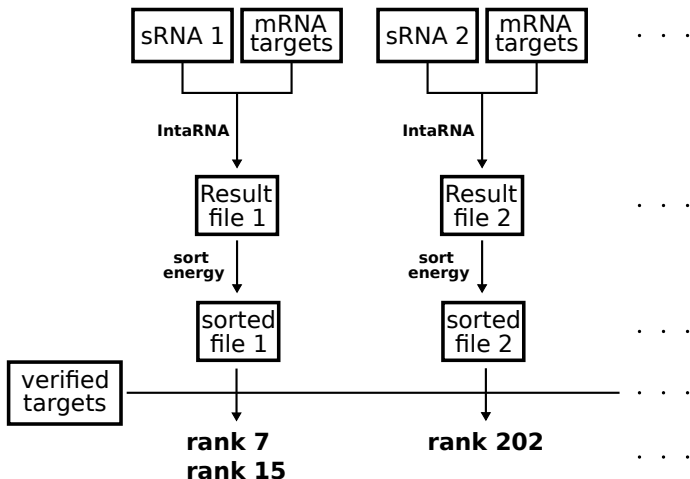
Appendix

	sRNA queries	target regions
E.coli	15	4319
Salmonella typh.	15	4552

- 133065 potential sRNA-target interactions
- 149 experimentally verified interactions
- genomic subregions of the target mRNA, including 200nt upstream and 100nt downstream



Benchmark



Original IntaRNA recursion

$$\begin{array}{c} i \\ \text{---} \\ \boxed{H} \\ \text{---} \\ j \end{array} \begin{array}{c} k \\ \text{---} \\ \boxed{H} \\ \text{---} \\ l \end{array} = \min \left\{ \begin{array}{l} \begin{array}{l} i = k \\ j = l \\ E_{init} \end{array} \\ \begin{array}{c} i \quad i+1 \quad k \\ \boxed{E_S} \quad \boxed{H} \\ j \quad j+1 \quad l \end{array} \\ \begin{array}{c} \min_{p,q} \begin{array}{c} i \quad p \quad k \\ \boxed{E_{IL}} \quad \boxed{H} \\ j \quad q \quad l \end{array} \end{array} \right.
 \end{array}$$

Canonical Helix-length Restricted Prediction

$$\begin{array}{c} i \\ \text{helix} \\ j \quad l \\ k \end{array} = \min \left\{ \begin{array}{l} \begin{array}{c} i \quad i+1 \quad k \\ E_S \quad \text{helix} \\ j \quad j+1 \quad l \end{array} \\ \begin{array}{c} i \quad k \\ E_S \\ j \quad l \end{array} \end{array} \right.$$

$$\begin{array}{c} i \\ H \\ j \quad l \\ k \end{array} = \min \left\{ \begin{array}{l} \begin{array}{c} i \quad k \\ \text{helix} + E_{init} \\ j \quad l \end{array} \\ \min_{pqr s} \begin{array}{c} i \quad p \quad r \quad k \\ \text{helix} \quad E_{IL} \quad H \\ j \quad q \quad s \quad l \end{array} \end{array} \right.$$



Enforcing seed constraints

$$\begin{array}{c}
 \begin{array}{|c|c|} \hline i & k \\ \hline \text{helix}_S & \\ \hline j & l \\ \hline \end{array} = \min_{pqr s} \begin{array}{|c|c|c|} \hline i & p & r & k \\ \hline \text{helix} & \text{seed} & \text{helix} & \\ \hline j & q & s & l \\ \hline \end{array}
 \end{array}$$

$$\begin{array}{c}
 \begin{array}{|c|c|} \hline i & k \\ \hline H_S & \\ \hline j & l \\ \hline \end{array} = \min \left\{ \begin{array}{l}
 \begin{array}{|c|c|} \hline i & k \\ \hline \text{helix}_S & \\ \hline j & l \\ \hline \end{array} + E_{init} \\
 \min_{pqr s} \begin{array}{|c|c|c|} \hline i & p & r & k \\ \hline \text{helix}_S & E_{IL} & H & \\ \hline j & q & s & l \\ \hline \end{array} \\
 \min_{pqr s} \begin{array}{|c|c|c|} \hline i & p & r & k \\ \hline \text{helix} & E_{IL} & H_S & \\ \hline j & q & s & l \\ \hline \end{array}
 \end{array} \right.
 \end{array}$$



Allowing only stable helices

