Computational methods for comparing and integrating multiple probing assays to predict RNA secondary structure
Afaf Saaidi, Delphine M Allouche, Yann Ponty, Bruno M Sargueil, Mireille M Regnier

To cite this version:
Afaf Saaidi, Delphine M Allouche, Yann Ponty, Bruno M Sargueil, Mireille M Regnier. Computational methods for comparing and integrating multiple probing assays to predict RNA secondary structure. Doctorial journey Interface, Ecole polytechnique, Palaiseau, Nov 2016, Palaiseau, France. 10. hal-01558227

HAL Id: hal-01558227
https://hal.science/hal-01558227
Submitted on 7 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Computational methods for comparing and integrating multiple probing assays to predict RNA secondary structure

Afaf Saaidi1,2,3, Delphine Allouche3,4
Supervisors: Yann Ponty1,2,3, Bruno Sargueil3,4, Mireille Regnier1,2
1-AMIB team, Inria Saclay 2-LIX 3-CNRS 4-Laboratoire de cristallographie et RMN Biologiques, Paris Descartes

1-Introduction
- RNA is key to understanding many biological processes.
- RNA maintains a stable tertiary structure.
- The determination of the structure allows understanding its operating mechanism.
- We study the 444nt long HIV Gag-IRE.

RNA Structure determination
- 3D structure can be resolved experimentally but remains expensive and time-consuming.
- Computational methods allow to have accurate secondary structure predictions (PPV > 75%). Less accurate predictions for long RNA.
- Experimental Data [Chemical/SHAPE \ Enzymatic] improve predictions.

2-Material & Methods
2-1 Experimental data
SHAPE-Map experiments
High Throughput Sequencing

SHAPE reaction calculation

Reactivity(n) = \frac{\text{max}_{\text{Shape}(n)} - \text{min}_{\text{Control}(n)}}{\text{max}_{\text{Operational}(n)}}

Boltzmann probability to observe a structure S:

P(S) = \frac{e^{-\beta E(S)}}{Z}

with Z the partition function: Z = \sum_{S} e^{-\beta E(S)}.

2-2 Sampling/Clustering workflow
- Experimental data from different conditions
- Data processing
- Structure sampling
- Set of ensemble structures
- Clustering[Affinity propagation]

Optimal clusters?

3-Results
Optimal centroid structures from 140 [8000 structures]

4-Conclusion & perspectives
- We have obtained a set of models supported by our integrative approach, those models are subject to validation.
- Some of centroid structures have shown high compatibility with existing proposed structural models.
- We will extend the approach to the simultaneous analysis of probing data for a set of RNA variants.

References

Acknowledgments
PhD funded by the "Fondation pour la Recherche Médicale".

Contact address
saaidi.afaf@polytechnique.edu