Computational search for functional structures in viral RNA genomes

Alexander P. Gultyaev

Erasmus Medical Center, Department of Viroscience, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands,

a.goultiaev@erasmusmc.nl

René C.L. Olsthoorn

Leiden Institute of Chemistry, Leiden University, P.O.Box 9502, 2300 RA Leiden, The Netherlands,

olsthoor@chem.leidenuniv.nl

Many viruses contain their genetic information in RNA genomes. Replication of RNA genomes is influenced by higher order structures, formed by virus RNA's at the different steps of virus replication cycle. For instance, such structures can modulate virus RNA replication, transcription, translation, processing and packaging into virion particles. Thus, apart from coding for viral proteins, RNA genomes encode multiple functional RNA structures. In many cases, the structure-function relationships of viral regulatory RNA structures are confirmed experimentally by structure determination and/or functional assays. On the other hand, important functional RNA motifs can be predicted using computational approaches for RNA structure prediction and comparative sequence analysis. However, despite availability of multiple RNA folding prediction methods, a reliable theoretical prediction of biologically important structures remains a difficult task. Here, various strategies of computational searching for functional RNA structures in viral genomes will be analysed in applications for different types of viruses and virus RNA functions. Several case studies demonstrate how computational strategies can be adjusted for solving particular problems in the prediction of functional viral RNA structures.

One of the typical problems in RNA folding predictions is existence of metastable structures that makes a direct application of free energy minimisation less reliable. Furthermore, a biologically relevant model usually requires the prediction of structures conserved in a group of related viruses. In particular, nucleotide covariations yield a reliable support for the functional importance of predicted structures, being an evidence for significant RNA

structure constraints in evolution. Thus, the presence of covariations is considered as one of the best theoretical criteria for RNA model evaluations.

In case of thermodynamically stable local secondary RNA structures and availability of a dataset of related RNA sequences with certain degree of diversity it is possible to derive a reasonable structural model using direct comparison of the lowest free energy conformations. For instance, such an approach revealed the existence of group-specific packaging signals in coronavirus genomic RNAs (1).

An intriguing example of functional RNA structure evolution is provided by influenza viruses. Despite the availability of massive sequence data obtained from thousands of virus strains, a task of finding functional structures in the influenza virus genome is not straightforward. Nevertheless, using various strategies of RNA structure analysis, based on the folding simulations and predictions of consensus structures, it is possible to show the presence of conserved RNA motifs in the influenza virus genome. Furthermore, these motifs undergo conformational transitions, apparently relevant for the origin of new phylogenetic lineages (2, 3). A number of structures, predicted in various RNA segments of influenza genome, are supported by nucleotide covariations, demonstrating functional importance of influenza virus RNA folding (3).

Viral RNA genomes also demonstrate a remarkable structural diversity of RNA motifs, such as complex pseudoknotted structures (e.g. 4). Some computational approaches allowing one to cope with the complexity of viral RNA folding and examples of applications for animal and plant viruses will be discussed.

1. S.C.Chen, R.C.L.Olsthoorn (2010) Group-specific structural features of the 5'-proximal sequences of coronavirus genomic RNAs, *Virology,* **401:**29-41.

2. A.P.Gultyaev et al. (2007) An RNA conformational shift in recent H5N1 viruses, *Bioinformatics,* **23:**272–276.

3. A.P.Gultyaev et al. (2010) Influenza virus RNA structure: unique and common features, *Int Rev Immunol,* **29:**533–556.

4. A.P.Gultyaev, R.C.L.Olsthoorn (2010) A family of non-classical pseudoknots in influenza A and B viruses, *RNA Biology,* **7:**125–129.