

Fine-Grain Matrix Graph Representation for Predicting Mutations Leading to Conformational Rearrangements in Small RNAs

Assaf Avihoo, Nir Dromi, and Danny Barash
Department of Computer Science
Ben-Gurion University
84105 Beer-Sheva, Israel
Corresponding Email: dbarash@cs.bgu.ac.il

Abstract

Previously, it was shown that predicting selective mutations leading to topological transitions in the secondary structure of RNAs can be achieved by a coarse-grain Laplacian matrix tree graph representation using its second eigenvalue. When applying the coarse-grain tree graph representation, introduced by Shapiro and coworkers in the 80's, it is possible to predict mutations leading to conformational rearrangements in RNAs of around 50 nt and higher. However, for small RNAs, such representations at the level of stems, bulges, and loops become ineffective. Recently, there is an interest in investigating secondary structure rearrangements in small RNAs, following their structural probing by comparative imino proton NMR spectroscopy. For computational predictions of mutations leading to the structure rearrangements of small RNAs, it is necessary to use a fine-grain graph representation as introduced by Waterman in the 70's at the level of nucleotides. Each nucleotide becomes a node in the graph and its equivalent Laplacian matrix is of the size $N \times N$ for a sequence of N nucleotides. Conformational rearrangements caused by mutations can be studied using measures to assess the differences between Laplacian matrices of fine-grain graph representations. The second eigenvalue of the Laplacian matrix can be used to filter mutations that lead to a structure similar to the wildtype but additional measures are needed. Image analysis techniques, by moving a sliding window over Laplacian matrices, can facilitate in differentiating between local rearrangements and global rearrangements.

1. Background

The computational prediction of mutations that lead to conformational rearrangements in the secondary structure

of RNAs can be used for various applications. First, it can assist in RNA evolution simulations, as suggested in [1]. Second, it can be used to predict deleterious mutations that interfere with RNA regulation (riboswitch examples in [1]) and assist in the design of bi-stable RNAs [2]. An earlier study proposing mutation predictions in RNA secondary structure can be found in [3].

For predicting topological changes in the secondary structure of RNAs, a coarse-grain tree graph representation that was initiated by Shapiro and coworkers at the level of stems, bulges, and loops [4,5,6] was used in the aforementioned works. In [1,2] the second eigenvalue of the Laplacian matrix corresponding to the coarse-grain tree graph representation was shown to provide a good estimate in practice for measuring distances between secondary structure mutants and the wildtype. Using theorems that specify the upper and lower bounds of the algebraic connectivity of tree-graphs [7,8] it was possible to detect candidate mutations leading to conformational rearrangements in RNA sequences of around 50nt and higher. This was performed by successive folding predictions using *mfold* [9] or the Vienna package [10] for the wildtype sequence and the various layers of mutations inserted (single-point, double-point, etc). By arranging the data in eigenvalue tables [1,2] it is possible to discard the majority of possible combinations that are not suspected for causing global rearrangements in the structure.

A challenging task is to extend the methodology for predicting topological changes in small RNAs at a finer resolution. For RNA sequences that are less than 50nt long, coarse-grain tree graph representations are less effective. Fine-grain graph representations at the level of nucleotides introduced in [11] (see also [10,12] for various other possibilities) are needed. A matrix representation corresponding to the fine-grain graph (see Figure 1) can be used in the same manner as with the coarse-grain graph [1]. However, other distance measures between the wildtype and mu-

tant structures besides the second eigenvalue of the Laplacian matrix are needed. We are currently investigating image analysis techniques as a consequence of viewing the Laplacian matrix as an image, for differentiating between local rearrangements and global rearrangements to predict mutations leading to structural transitions in small RNAs.

2. Biological Relevance

Recently, there is a renewed interest in studying bi-stable RNAs. Their successful structural probing by comparative imino proton NMR spectroscopy [13,14] prompts the design of artificial RNA secondary structures possessing bi-stability properties. For this purpose, various computational predictions to assist in the design and detection of bi-stable RNAs [15,16] can be used. We propose a complementary approach for such predictions based on mutation simulations. As was validated by comparing to the experimental results of LeCuyer and Crothers [17] our computational method with the coarse-grain tree graph representation is able to predict the single-point mutation that drives the RNA molecule into a bi-stable structure [2]. However, for smaller RNAs such as the ones examined in [13] and depicted in Figure 2, the coarse-grain graph representation becomes ineffective. The fine-grain graph representation can potentially enable predicting mutations leading to bi-stability in small RNAs. Details concerning the developed methodology and its application will be given elsewhere.

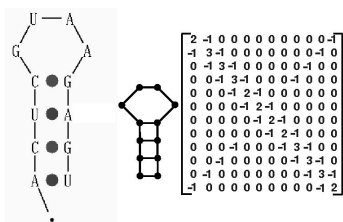


Figure 1. Fine-Grain Laplacian Matrix Graph Representation of the Hammerhead Ribozyme Hairpin.

References

[1] D. Barash, "Deleterious Mutation Prediction in the Secondary Structure of RNAs," *Nucleic Acids Res.*, 31, pp.6578–6584, 2003.
 [2] D. Barash, "Second Eigenvalue of the Laplacian Matrix for Predicting RNA Conformational Switch by Mutation," *Bioinformatics*, 10.1093/bioinformatics/bth157, 2004.
 [3] H. Margalit, B.A. Shapiro, A.B. Oppenheim, and J.V. Maizel, "Detection of Common Motifs in RNA Secondary Structure," *Nucleic Acids Res.*, Vol. 17, pp.4829-4845, 1989.

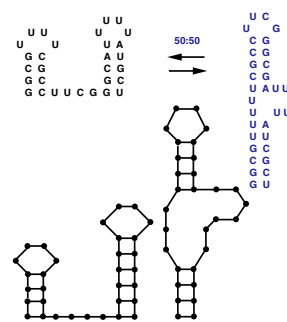


Figure 2. Fine-Grain Graph Representation of a Bi-Stable RNA taken from [13].

[4] B.A. Shapiro, "An Algorithm for Comparing Multiple RNA Secondary Structures," *Comput. Appl. Biosci.*, Vol.4, pp.387-393, 1988.
 [5] S. Y. Le, R. Nussinov, and J.V. Maizel, "Tree Graphs of RNA Secondary Structures and Their Comparisons," *Comput. Biomed. Res.*, Vol.22, pp.461-473, 1989.
 [6] B.A. Shapiro, and K. Zhang, "Comparing Multiple RNA Secondary Structures Using Tree Comparisons." *Comput. Appl. Biosci.*, Vol.6, pp.309–318, 1990.
 [7] M. Fiedler, "Algebraic Connectivity of Graphs," *Czechoslovak Math. J.*, Vol.23, pp.298-305, 1973.
 [8] R. Merris, "Characteristic Vertices of Trees," *Lin. Multi. Alg.*, Vol.22, pp.115-131, 1987.
 [9] M. Zuker, "Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction," *Nucleic Acids Res.*, Vol.31, pp.3406-3415, 2003.
 [10] I.L. Hofacker, "Vienna RNA Secondary Structure Server," *Nucleic Acids Res.*, Vol.31, pp.3429–3431, 2003.
 [11] M.S. Waterman, "Secondary Structure of Single-Stranded Nucleic Acids," *Studies on foundations and combinatorics, Advances in mathematics supplementary studies*, Academic Press N.Y., Vol.1, pp.167–212, 1978.
 [12] T. Jiang, G.-H. Lin, B. Ma, and K. Zhang, "A General Edit Distance Between RNA structures," *J. Comput. Biol.*, Vol.9, pp.371–388, 2002.
 [13] C. Höbartner, and R. Micura, "Bistable Secondary Structures of Small RNAs and Their Structural Probing by Comparative Imino Proton NMR Spectroscopy," *J. Mol. Biol.*, Vol.325, pp.421-431, 2003.
 [14] R. Micura, and C. Höbartner "On Secondary Structure Rearrangements and Equilibria of Small RNAs," *ChemBioChem*, Vol.4, pp.984–990, 2003.
 [15] C. Flamm, I.L. Hofacker, S. Maurer-Stroh, P.F. Stadler, and M. Zehl, "Design of Multistable RNA Molecules," *RNA*, Vol.7, pp.254-265, 2001.
 [16] B. Vofsi, C. Meyer, and R. Giegerich, "Evaluating the predictability of conformational switching in RNA," *Bioinformatics*, 10.1093/bioinformatics/bth129, 2004.
 [17] K.A. LeCuyer, and D.M. Crothers, "Kinetics of an RNA Molecular Switch," *Proc. Natl. Acad. Sci.*, Vol.91, pp.3373–3377, 1994.