RNAMute: RNA Secondary Structure Mutation Analysis Tool

Alexander Churkin, Ornit Elgad, Oshrat Petgho and Danny Barash
Department of Computer Science
Ben-Gurion University
84105 Beer-Sheva, Israel
Corresponding Email: dbarash@cs.bgu.ac.il

Abstract

RNAMute is an interactive web tool written in Java that calculates the secondary structure of all single point mutations, given an RNA sequence, and organizes them into categories according to their distances from the wildtype predicted structure. The secondary structure predictions are performed using the Vienna RNA package. Several alternatives are used for the categorization of single point mutations: Vienna's RNADistance based on dot-bracket representation, tree edit distance, and second eigenvalue of the Laplacian matrix based on Shapiro's coarse grain tree-graph representation. Selecting a category in each one of the tables lists all single point mutations belonging to that category. Selecting a mutation displays a graphical picture of the single point mutation and the wildtype, and includes basic information such as associated energies, representations, distances. RNAMute is a user friendly tool that can be used to predict single point mutations leading to conformational rearrangements in the secondary structure of RNAs.

1. Introduction

RNAMute is a user friendly computer tool that analyzes point mutations in the secondary structure of RNAs. Initial ideas can be found in [7] and associated works in the late 80's [10, 5, 9]. Since then, much progress has been made in the field RNA secondary structure prediction, with the gradual development of sophisticated energy minimization folding prediction packages (most notably, Zuker's *mfold* [13] and the Vienna RNA package [3, 4]). The possibility of reliably predicting conformational rearranging point mutations in the secondary structure of RNAs has been revisited in [1], suggesting a coarse-grain tree graph representation of the RNA secondary structure [10] and the use of mathematical theorems that relate to eigen-decomposition of the Laplacian matrix [2, 8] corresponding to the coarse-grain tree graphs. Both fine-grain and coarse-grain graph representations, including distance measures between the graphs, have been implemented in the Vienna RNA package [3]. We use the Vienna RNA package as the core of RNAMute, attaching to it the mutation prediction procedure described in [1].

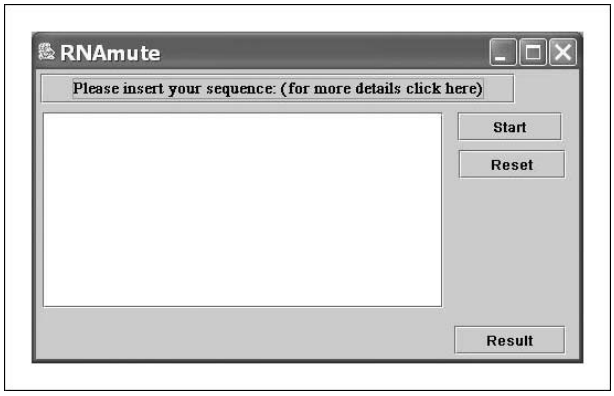

Figure 1. RNAMute Input Screen.

The input to RNAMute is simply an RNA sequence (see Figure 1). Subsequently, RNAMute scans all possible single point mutations in that sequence and computes their folding prediction using Vienna's RNAfold program. The analysis of point mutations is illustrated in Figures 2-5 and will be described in detail elsewhere. Such analysis is capable of predicting conformational rearranging single point mutations, for example the point mutation that is responsible for switching between FORM 1 WT RNA to FORM 2 M3 RNA, as described and examined experimentally in [6].

2. Biological Relevance

An example to a potentially beneficial use of RNAMute is the examination of phenotypic data available from Hepatitis C Virus (HCV) experiments [11, 12]. The example input sequence in Figures 2–5 is from the 5BSL3.2 structure that was studied experimentally by mutagenesis.

Dot-Bracket Distance range	Frequency
0.0 - 0.0	60
4.0 - 4.0	22
8.0 - 24.0	53
28.0 - 28.0	1

*** the average of Dot-Bracket Distances is: 7.5407**
Clustering Resolution : 4

Figure 2. Mutation Grouping Table According to Vienna's RNADistance Ranges.

Shapiro-Distances range	Frequency
0.0 - 30.0	115
38.0 - 45.0	14
49.0 - 54.0	7

*** the average of Shapiro's Distances is: 13.755**
Clustering Resolution : 4

Figure 3. Mutation Grouping Table According to Shapiro's Tree-Graph Distance Ranges.

References

[1] D. Barash, "Second Eigenvalue of the Laplacian Matrix for Predicting RNA Conformational Switch by Mutation," *Bioinformatics*, Vol.20(12), pp.1861-1869, 2004.

[2] M. Fiedler,"Algebraic Connectivity of Graphs,"*Czechoslovak Math. J.*, Vol.23, pp.298-305, 1973.

[3] I.L. Hofacker, W. Fontana, P.F. Stadler, L.S. Bonhoeffer, M. Tacker, and P. Schuster, "Fast Folding and Comparison of RNA Secondary Structures", *Monatsh.Chem.*, Vol. 125, pp.167–188, 1994.

[4] I.L. Hofacker, "Vienna RNA Secondary Structure Server," *Nucleic Acids Res.*, Vol.31, pp.3429–3431, 2003.

[5] S.Y. Le, R. Nussinov, and J.V. Maizel. Tree graphs of RNA secondary structures and their comparisons. *Comput. Appl. Biosci.*, 22:461–473, 1989.

For more information about a specific eigenValue, press the required link

Second eigenvalue	Number of vertices	wild type	Frequency
0.585786	4	-	14
1.000000	3	WT	102
1.000000	4	-	16
2.000000	2	-	4

Figure 4. Mutation Grouping Table According to the Tree-Graph Laplacian Second Eigenvalues.

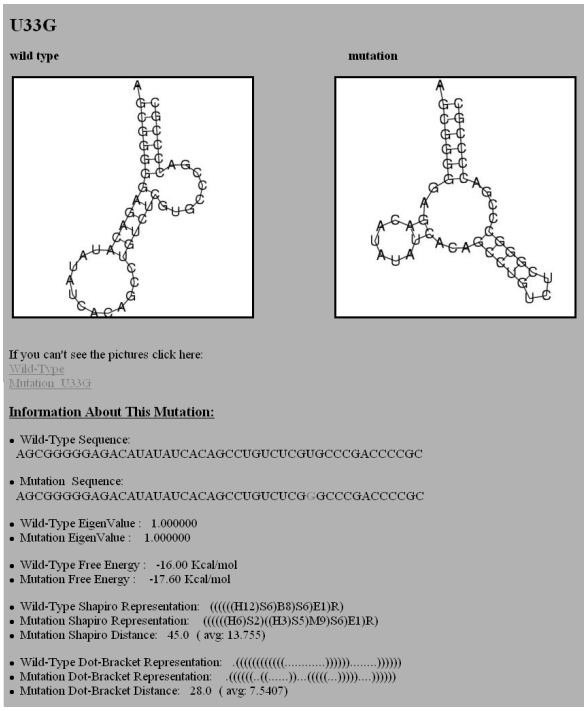

Figure 5. An Example of A Conformational Rearranging Mutation in the 5BSL3.2 Structure.

[6] K.A. LeCuyer, and D.M. Crothers, "Kinetics of an RNA Molecular Switch," *Proc. Natl. Acad. Sci.*, Vol.91, pp.3373–3377, 1994.

[7] H. Margalit, B.A. Shapiro, A.B. Oppenheim, and J.V. Maizel. Detection of common motifs in RNA secondary structures. *Nucleic Acids Res.*, 17:4829–4845, 1989.

[8] R. Merris. Characteristic vertices of trees. *Lin. Multi. Alg.*, 22:115–131, 1987.

[9] B.A. Shapiro and K. Zhang. Comparing multiple RNA secondary structures using tree comparisons. *Comput. Appl. Biosci.* 6:309–318, 1990.

[10] B.A. Shapiro, "An Algorithm for Comparing Multiple RNA Secondary Structures," *Comput. Appl. Biosci.*, Vol.4, pp.387-393, 1988.

[11] D.B. Smith and P. Simmonds, "Characteristics of Nucleotide Substitution in the Hepatitis C Virus Genome: Constraints on the Sequence Change in Coding Regions at Both Ends of the Genome," *J. Mol. Evol.*, Vol.45, pp.238-246, 1997.

[12] S. You, D.D. Stump, A.D. Branch, and C.M. Rice, "A *cis*-Acting Replication Element in the Sequence Encoding the NS5B RNA-Dependent RNA Polymerase Is Required for Hepatitis C Virus RNA Replication," *J. Virol.*, Vol.78, pp.1352-1366, 2004.

[13] M. Zuker, "Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction," *Nucleic Acids Res.*, Vol.31, pp.3406-3415, 2003.