

Efficient calculation of exact probability distributions of integer features on RNA secondary structures

Ryota Mori 1, Michiaki Hamada 2,3 and Kiyoshi Asai 1,2

1 Graduate School of Frontier Sciences, The University of Tokyo, Japan

2 Computational Biology Research Center, National Institute of Advanced Industrial Science and Technology (AIST),
Japan

3 Faculty of Science and Engineering, Waseda University, Japan

Abstract

Background

Although the needs for analyses of secondary structures of RNAs are increasing, prediction of the secondary structures of RNAs are not always reliable. Because an RNA may have a complicated energy landscape, comprehensive representations of the whole ensemble of the secondary structures, such as the probability distributions of various features of RNA secondary structures are required.

Results

A general method to efficiently compute the distribution of any integer scalar/vector function on the secondary structure is proposed. We also show two concrete algorithms, for Hamming distance from a reference structure and for 5'-3'distance, which can be constructed by following our general method. These practical applications of this method show the effectiveness of the proposed method.

Conclusions

The proposed method provides a clear and comprehensive procedure to construct algorithms for distributions of various integer features. In addition, distributions of integer vectors, that is a combination of different integer scores, can be also described by applying our 2D expanding technique. The web server for the distributions of the Hamming distances is available at <http://www.ncrna.org/rtools>.