



# Novel representation of RNA secondary structure used to improve prediction algorithms

Q. Zou<sup>1</sup>, C. Lin<sup>1</sup>, X.-Y. Liu<sup>2</sup>, Y.-P. Han<sup>3</sup>, W.-B. Li<sup>3</sup> and M.-Z. Guo<sup>2</sup>

<sup>1</sup>School of Information Science and Technology, Xiamen University, Xiamen, China

<sup>2</sup>School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

<sup>3</sup>Key Laboratory of Soybean Biology, Chinese Education Ministry, Soybean Research Institute, Northeast Agricultural University, Harbin, China

Corresponding author: M.-Z. Guo

E-mail: maozuguo@hit.edu.cn

Genet. Mol. Res. 10 (3): 1986-1998 (2011)

Received November 29, 2010

Accepted July 24, 2011

Published September 9, 2011

DOI <http://dx.doi.org/10.4238/vol10-3gmr1181>

**ABSTRACT.** We propose a novel representation of RNA secondary structure for a quick comparison of different structures. Secondary structure was viewed as a set of stems and each stem was represented by two values according to its position. Using this representation, we improved the comparative sequence analysis method results and the minimum free-energy model. In the comparative sequence analysis method, a novel algorithm independent of multiple sequence alignment was developed to improve performance. When dealing with a single-RNA sequence, the minimum free-energy model is improved by combining it with RNA class information. Secondary structure prediction experiments were done on tRNA and RNase P RNA; sensitivity and specificity were both improved. Furthermore, software programs were developed for non-commercial use.

**Key words:** RNA secondary structure; Dot plots; Stem; Comparative sequence analysis