Study on the variation of RNA Secondary Structure prediction as a function of Thermodynamic Parameters

Rodolfo Vieira Maximiano[12], Gerald Weber[2]

[1]Centro Federal de Educação Tecnológica de Minas Gerais - unidade Contagem,
[2]Departamento de Física - Universidade Federal de Minas Gerais

In this work we evaluated how the accuracy of RNA secondary structure prediction softwares is modified as a function of the thermodynamic parameters. The parameters in use today are obtained from DNA and RNA melting experiments and each measurement is subject to an intrinsic experimental error. These experiments are modeled using the nearest-neighbour technique, which is one of the most commonly used theoretical models in this field of study. Here we used the RNAFold software from the Vienna package, which is in frequent use for bioinformatics applications. Our purpose was to verify the impact of the measured experimental error in the software ability to predict the final RNA secondary structure of a diverse database of known RNAs, which have folded into structures determined experimentally by other methods. When comparing both the experimentally known and predicted structures we can evaluate the prediction quality. The indicators of prediction quality used in this work are often used in literature, namely, the positive predict value (PPV) and the sensitivity. We created additional sets of nearest-neighbor thermodynamic parameters taking into account the reported experimental error. For each new set, we performed the folding of all known RNA sequences, calculating a general PPV and sensitivity quality indicator for that particular set. We are also performing verifications of the variation of prediction accuracy when subdividing the known RNA sequences by size, with the intent of detecting how, and when, does the prediction ability of the software cease to be as precise as it is for small sequences. So far our results indicate a stronger modification in the prediction accuracy when modifying the hairpin and bulge formation parameters. In comparison, modifying the base pair stacking parameters has generally a lesser impact on the folding prediction quality.

Funding Agencies: CNPQ, FAPEMIG, CAPES