The evolution of 5' untranslated regions' structure in Bacilli and Clostridia genomes

A.S.Obraztsova, Z.S.Chervontseva, E.D.Stavrovskaya, A.A.Mironov

Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, 119991, Moscow, GSP-1, Leninskiye Gory, MSU, 1-73, Moscow, Russia, `chezoya@gmail.com`

Noncoding RNAs (ncRNAs) were shown to have a wide variety of functions in cells. The common feature of this molecules is a stable secondary structure which often determine their functional activity. There are RNA regulators of gene expression among them and riboswitches particularly. Riboswitches are located in 5' untranslated regions of mRNA and could change their conformation in the presence or absence of some ligand. This change affect transcription and translation of downstream genes which often results in changes of cell metabolism so investigation of new riboswitches is important for metabolic reconstruction.

Modern sequencing technologies produce a huge amount of genomic data, hence computational prediction of noncoding RNAs is an actual problem. Most existing methods are based on searching for evolutionarily conserved RNA secondary structures, but they don't use information about evolutionary distances between species. A big disadvantage of such methods is that they are very sensitive to noise level in the data. In this work we suggest a novel method for finding new RNA regulators based on single sequence structure analysis with taking phylogenetic surrounding into account. Furthermore this method allows to reconstruct the most likely evolutionary scenario of gains and losses of ncRNA-regulation.

Our method is based on two programs previously developed in our group. We use RNASurface [1] to get a z-score - a numerical measure of RNA structure stability based on sequence free energy minimization. Then we apply tHMM [2] which takes as input prior probability distribution of obtained z-scores together with a phylogeny tree of species of interest. As a result we obtain the posterior probability of some functional structure presence for each tree node.

1. R.A.Soldatov, S.V.Vinogradova, A.A.Mironov (2014) RNASurface: fast and accurate detection of locally optimal potentially structured RNA segments, *Bioinformatics,* **30(4):**457-63.

2. N.A.Bykova, A.V.Favorov, A.A.Mironov (2013) Hidden Markov models for evolution and comparative genomics analysis, *PLoS ONE*, **8(6):**e65012.