A NOVEL PROFILE HIDDEN MARKOV MODEL TO PREDICT MICRORNAS AND THEIR TARGETS SIMULTANEOUSLY

Shu Yang, Kalpana Agrawal, Tak-Wah Lam1, Pak Chung Sham2, Kathryn SE Cheah and Junwen Wang
Department of Biochemistry, 1.Department of Computer Sciences, 2. Department of Psychiatry
The University of Hong Kong, Hong Kong SAR, China
Email:syang.2008@hku.hk

The microRNAs (miRNAs) are small non-coding RNAs of ~22 nucleotides, which play important roles in gene regulation. By binding to 3' UTR of the target mRNA, miRNAs can either speed up the degradation of mRNAs or slow down the translation of gene products. The identification of miRNA and its functional target sites becomes an essential issue. Although existing computational approaches have made significant progresses in predicting either miRNAs or targets, the underlying mechanism of miRNA-target interaction still remained largely unknown. To further understand this interaction, here we present a novel Profile Hidden Markov Model which can predict miRNAs and their targets simultaneously. Unlike current strategies which only use features from a group of miRNAs or a set of targets for one particular miRNA, our model integrates the features from both miRNA and target together and thus captured evolutionary relationships of both. Since the "seed region" is only ~8bp long, current models based on region like this short inevitably suffer from high false positive rate. Besides, recent studies also reported the prevalence of interactions beyond seed matching as supplementary binding. Our method provides an unbiased way to model the interaction. By considering miRNA and target features together, we double the size of modeling region to 22×2. As more and more miRNA-target pairs are discovered, this HMM could achieve higher specificity. In preliminary tests, it outperformed RNAhybrid and achieved satisfactory results for leave-one-out cross validation.