

Computational Prediction of miRNA Regulations in Mouse Whole Brain Using *in situ* Hybridization Data

Risa Kawaguchi, Hisanori Kiryu

Abstract

Recent research for microRNA (miRNA) expression has revealed its importance in vertebrate animals especially for gene silencing at the synapse. Many researchers have measured miRNA expressions to investigate such miRNA activity by microRNA array, RT-PCR, RNA-seq, and so on. It is, however, still impossible to investigate an extent of miRNA regulation directly. Therefore, we tried to predict it by examining relationships between mRNA expressions and an existence of some 7-mer (seed of miRNA) sequences in its 3'UTR using the data from Allen Brain Atlas database.

At first, we calculated p-values using hyper geometric distribution (HGD) about strength of mRNA expression and seed existence. Shuffled mRNA 3'UTR sequences were used as a control. We also applied a FDR control for p-values obtained from HGD.

We were able to predict many regulations of miRNA specifically to brain organs by HGD and some of which agreed with previous research, miR-200 family and let-7 family, for example. After applying the false discovery rate (FDR) control method, however, many local regulations were veiled and the influence of seed sequence specific bias was observed. It supposed to be because of differences of seed frequency after 3'UTR shuffling.

To avoid such bias, it is necessary to account for a frequency of seed appearances in shuffled 3'UTRs. So we calculated each seed frequency compared with shuffles to get "wordscore", which means Z-score of mean-adjusted cumulated score of logarithms of seed frequency in order of mRNA expressions. We obtained the wordscores and their ranks at each position. The distributions of wordscore rank showed some seeds, for example miR-124 and miR-137 previously reported to relate with brain and neuron, had high orders of wordscore at many positions and inferred to have relationships with mRNA expressions.

In conclusion, our method is useful to predict miRNA or some signatures regulating broadly and also has a potential to predict local regulations combining the method to detect localization of high wordscores.