A systematic analysis of alternative 3’ UTR of human transcripts: improvment of detection of microRNA targets

Sales G[(1)], Biasiolo M[(1)] and Romualdi C[(1)]

[(1)] Department of Biology, University of Padova, Padova

Motivation

MicroRNAs are small noncoding RNAs that serve as post-transcriptional regulators of gene expression in higher eukaryotes. Their widespread and important role in animals is highlighted by recent estimates that 20% to 30% of all genes are microRNA targets. DNA microarray technology can be used to detect the interaction between microRNAs and their targets, supplementing sequence based methods that search for seed complementarily and for energetically favourable folded structures. Given that both messenger RNAs and microRNAs are measured on the same biological samples the resulting profiles are searched for anti-correlation patterns. Independently from the method used to detect such anti-correlation, current approaches suffer from a major problem. Strong experimental evidences indicate that miRNAs interact with their targets by direct pairing to their 3'UTR regions; each gene may have multiple such UTRs corresponding to alternatively spliced transcripts. Microarray annotations, on the other hand, usually associate (meta) probes with gene symbols or EntrezGene, neglecting any information about single transcripts. The net result is that anti-correlation studies combine contributes from different transcripts. This could potentially lead to opposite behaviours with respect to the same microRNA.

Methods

We have developed a fully automated software pipeline (based on the BioinfoTree system: http://www.bioinfotree.org) that re-annotates microarray experiments using transcript information, such that at least 4 different probes will be exclusively linked to a transcript, filtering out those pointing to multiple transcripts. We use this custom annotation to extract expression levels from experimental data, averaging the measures by different probes in the same metaprobe. Two different prediction tools, miRanda and RNAhybrid, have been used for the prediction of microRNA targets Finally, we integrate all the information we have gathered in the previous steps, and search for anti-correlations between the expression profiles of transcripts and microRNAs that have been predicted to interact. We repeat the same analysis using gene-centered annotations (as opposed to transcriptcentered) for comparison purposes.

Results

We have analyzed the GEO datasets (GSE7055) using the procedure we have outlined. Our transcript based annotation was able to identify metaprobes for 6601 transcripts; 1766 of those belong to genes with multiple splicing variants, i.e. with different 3’UTR region that could cause different responses to the same microRNA. Restricting our attention to the alternatively spliced mRNA only, we found several cases particularly interesting for the purpose of studying the different types of microarray annotation. These cases, in fact, show that at least one transcript is more anti-correlated to the microRNA than the gene, that thus, seems to give a bias measure of the mRNA-microRNA correlation. Transcript-based annotation seems therefore to achieve higher sensitivity for this kind of integrative analysis.

Contact : gbrsales@gmail.com