Non-coding RNAs (ncRNAs) include different classes with regulatory functions. Among them, the most studied are miRNAs that act directly inhibiting mRNA expression or protein translation through the interaction with a miRNAs-response element (MRE). Different pieces of evidence show their potential role both in physiological than in pathological processes as cancer. Other RNA molecules participate in the complex network of gene regulation, as long non-coding RNAs (lncRNAs), circular RNAs (circRNAs), pseudogenes, and mRNA. They behave as competitive endogenous RNA (ceRNA), acting as natural miRNA sponges to inhibit miRNA functions and modulate mRNA expression, through their de-repression. It became evident that understanding the ceRNA–miRNA–mRNA crosstalk would increase the functional information across the transcriptome, contributing to identify new potential biomarkers for translational medicine. We aim to improve our miRTissue web service, combining validated miRNA-mRNA target interactions and predicted ceRNA-miRNA target interactions with statistical correlation among expression profiles of ceRNAs, miRNAs, mRNAs and proteins in different human tissues. Our web service has three main advantages: 1) to deep understand mechanisms and networks involved in gene regulation; 2) to focus the network analysis on a specific tissue type; 3) to include protein expression values for the analysis of miRNA-target interactions, to provide a novel insight into the type of interactions.

Methods

The main computing pipeline is shown in Figure 1. Its purpose is to obtain a prediction of ceRNA interactions, that are ncRNA-miRNA-mRNA triplets meaning that the ncRNA and the mRNA are putative competitors for the miRNA interaction. Following a similar approach with the original miRTissue workflow, input data are a set of validated miRNA-mRNA interactions, a set of predicted ncRNA-miRNA interactions and a set of matched expression profiles of all the involved molecules, considering different normal and tumor human tissues. For each RNA pairs, we compute the hypergeometric statistical test to find the significance of sharing of miRNAs between these RNAs. This filtering step allows us to obtain a preliminary set of ceRNA-ceRNA interactions. Then, for each tuple ncRNA-miRNA-mRNA, we use the global test to compute the statistical correlations among expression profiles of: ncRNA-miRNA; ncRNA-mRNA; mRNA-miRNA (already available in current miRTissue release). According to recent literature, a tuple ncRNA-miRNA-mRNA represents a likely ceRNA-ceRNA interaction if ncRNA-miRNA are negatively correlated, miRNA-mRNA are negatively correlated, ncRNA-mRNA are positively correlated. Moreover, taking into account mRNA-protein expression correlation, already available, it is possible to update as ceRNA-ceRNA interactions, those miRNA-target interactions labeled as "degradation" in case mRNA and coded protein are positively correlated.

Results

We are going to extend current miRTissue release with data concerning ceRNA-ceRNA interactions, and in particular with lncRNA for now. In this work, we consider lncRNA-miRNA predicted interactions from StarBase, miRCode, Diana LncBase v2, LncACTdb 2.0. We take into account experimentally validated miRNA-mRNA interactions from the miRTarBase repository. With regards to the former release of miRTissue, already containing the expression values of miRNA, mRNA, and proteins belonging to 15 normal and cancer tissues, we download from TCGA the expression profiles of lncRNA (long intergenic non-coding RNA). This way miRTissue can offer improved views and interaction results involving not only miRNAs and their targets, but also the cross-talk interactions between lncRNA and mRNA. In the near future, other ceRNAs such as circRNA and pseudogenes will be included as soon as their expression profiles are available on TCGA.