DEFINING RNA-RELATED REGULATORY EVENTS: IMPLEMENTATION OF A PLATFORM TO INTEGRATE HIGH-THROUGHPUT RNA DATA

Felipe E. Ciamponi¹; Michael T. Lovci¹; Mario H. Bengtson²; Katlin B. Massirer¹

1- Center for Molecular Biology and Genetic Engineering, University of Campinas - CBMEG-UNICAMP, Campinas, São Paulo, Brazil;
2 - Department of Biochemistry and Tissue Biology, University of Campinas - UNICAMP, Campinas, São Paulo, Brazil.

RNA binding proteins (RPBs) have multiple regulatory roles, acting primarily in post-transcriptional regulation of several cellular mRNA targets, many of which are related to diseases. Thus, the process of predicting and selecting specific mRNA targets for experimental validation is very challenging. We implemented a local Galaxy environment coupled to R-based packages to accelerate answers to common questions associated with the study of RBP function. We are constructing our pipeline to handle protein interaction data, high-throughput RNA sequencing (RNA-seq) and RNA immunoprecipitation followed by sequencing (RIP-seq). We integrated the analysis of these short-read sequencing methods to evaluate measures of RNA abundance, differential exon usage, intron retention events and associations of RNA to their targets. As proof-of-principal, we are studying two stress granule associated RBPs: the first is CAPRIN1, thought to repress translation for dendritically localized ribosome complexes. We have identified a preference for CAPRIN1 to form protein-protein interactions with other RBPs and we discovered a preference for the regulation of transcriptions factors and mRNA stabilization. The second is TARDBP, which has been shown to regulate RNA splicing and mutations of this gene are linked to amyotrophic lateral sclerosis (ALS). We found that TARDBP can regulate a set of intron retention events in neurological-related mRNAs and we are now following up on validation of these events. The pipeline we are developing is a consistent and user-friendly means to track and process data and will be essential as we move to larger-scale sequencing-based screens for RBPs functions and to design more accurate experiments for experimental validation.

Keywords: Bioinformatics; RNA-seq; RNA-binding Proteins

Acknowledgments: FAPESP, FAEPEX-Unicamp e CNPq