Universitat Pompeu Fabra Barcelona | Master in Bioinformatics for Health Sciences

Master project 2021-2022

Personal Information

Supervisor	Josep Vilardell
Email	josep.vilardell@ibmb.csic.es
Institution	ICREA / IBMB
Website	https://www.ibmb.csic.es/en/department-of-molecular-genomics-dmg/molecular-mechanisms-of-pre-mrna-splicing/
Group	RNA Splicing

Project

Computational systems biology

Project Title:

A Genomics Perspective of pre-mRNA Splicing

Keywords:

spliceosome, ribosome, pre-mRNA, ribosomal protein, genome

Summary:

The information content of genomes can be greatly expanded by pre-mRNA splicing. Virtually all human pre-mRNAs need to be spliced to become mRNAs. Furthermore, most pre-mRNAs can be spliced into different mRNAs by alternative splicing. Therefore, it is hardly surprising that perturbations in splicing are linked to disease. However, we know little on how the splicing of particular sets of RNAs may be affected, and even less on how a number of splicing changes are coordinated during development or disease. To start addressing this question, we are analyzing WGS and RNASeq data from a number of cancer datasets. Although we are interested in all events of regulated splic-ing, we pay special attention to those related to the biosynthesis and function of the ri-bosome. A cycling cell depends on a suitable set of ribosomes to provide the necessary amount of structural and functional proteins before mitosis; paradoxically, making this machinery requires most of the cell's energy (as an illustrative example, a growing HeLa cell is making 1.6 x105 ribosomal proteins (RP) per minute). Thus, we expect that fast-growing cell, subjected to a strong selection (such as a tumor cell), will tweak this process trying to get any advantage. However, the analysis of the RP transcriptome presents specific challenges because (a), it includes the mRNAs that are amongst the cell's most abundant, and while the ribosome has one copy of each RP, the corre-sponding mRNAs are in variable amounts; (b), the RP pre-mRNAs undergo little alterna-tive splicing; and (c), the majority of human pseudogenes originate from them, which introduces ambiguity when mapping reads to the genome. Our initial results suggest that processing of this set of transcripts is altered in cancer in unexpected ways, and we plan on strengthening our conclusions by expanding our analyses. In this context there are many opportunities for those with a strong motivation to document genomic strategies that control the transcriptome of specific gene families, like those related to the ribosome or the spliceosome. The tasks involve quality analysis of raw RNASeq data, mapping using standard tools (for example, Hisat, STAR, and those related to direct sequencing of RNA), statistical analysis (Ballgown, Salmon, Vast-tools, DexSeq, or others), and modeling. Subject to progress, we would explore the use of transcriptomics data as a disease prognosis tool; namely, is a distinct distribution of transcripts indicative of a particular disease out-come?

References:

* Hussain, S. (2018) "Native RNA- Sequencing Throws its Hat into the Transcriptomics Ring" TiBS 1434. https:// /doi.org/10.1016/j.tibs.2018.02.007 * Guimaraes, JC. and Zavolan, M. (2016) "Patterns of ribosomal protein expression specify normal and malignant human cells" Genome Biol. 17:236-248 * Gupta, V. and J. R. Warner (2014). "Ribosome-omics of the human ribosome." RNA 20: 1004-1013. * Bitton, D. A., et al. (2014). "LaSSO, a strategy for genome-wide mapping of intronic lariats and branch points using RNA-seq." Genome Res 24(7): 1169-1179. * Acuna, L. I. and A. R. Kornblihtt (2014). "Long range chromatin organization: a new layer in splic-ing regulation?" Transcription 5. * Kawashima T et al (2014) Widespread use of non-productive alternative splice sites in Saccharo-myces cerevisiae. PLoS Genet. 2014 Apr 10;10(4):e1004249. * Zhang, J. and J. L. Manley (2013). "Misregulation of pre-mRNA alternative splicing in cancer." Cancer Discov 3(11): 1228-1237. * Fu, R. H., et al. (2013). "Aberrant alternative splicing events in Parkinson's disease." Cell Trans-plant 22(4): 653-661. * Plass, M., et al. (2012). "RNA secondary structure mediates alternative 3'ss selection in Saccha-romyces cerevisiae." RNA 18(6): 1103-1115.

Expected skills::

In addition to a good Bioinformatics background, an ability to work independently and to integrate diverse data (mutational analyses, epigenetics, transcriptomics, and others) is desirable. Knowledge of bash and R will be helpful.

Possibility of funding::

No

Possible continuity with PhD: :

To be discussed

Comments:

Our group provides an field with significant questions to be addressed by a bioinformatician developing a career. We can steer the student in the right direction to use her/his skills, helping to make sure that the effort is relevant. We cannot, however, provide refined Bioinformatics background. These questions relate to the study of the spliceosome from a Systems Biology point-of-view: is there a limit in the number of exons that can be processed into the same transcript? What determines the splicing order of the introns? What may be the role of the pre-mRNA structure? Has the number of RNA Pol II molecules transcribing a gene any relation with its splicing? Is splicing connected to translation in more ways than we presently know?