Propuesta de Trabajo Fin de Máster
Año académico 2021-2022
MÁSTER EN MÉTODOS COMPUTACIONALES EN CIENCIAS

Project Nº 25 ASIGNADO

Título: A Bioinformatic Methodology to Identify small Circular RNAs.

Departamento/ Laboratorio: TERAPIA GÉNICA Y REGULACIÓN DE LA EXPRESIÓN GÉNICA. CIMA

Director: Tomás ARAGÓN AMONARRIZ
Correo electrónico: taragon@unav.es
Codirector: Idoia OCHOA ÁLVAREZ
Correo electrónico: iochoal@unav.es

Resumen

The unfolded protein response (UPR) is an intracellular signaling pathway that, in response to deficiencies in protein folding at the endoplasmic reticulum (ER), builds a gene expression program to adjust the protein folding capacity to need. The most core, conserved mechanism is the non-canonical splicing of a small, 26-nucleotide intron from the mRNA encoding the transcription factor XBP1. XBP1 splicing is catalyzed by the ER-resident, stress-induced endonuclease IRE1, and the tRNA ligase RTCB that joins the exons after IRE1-mediated cleavage, and allows the translation of XBP1s protein, a potent driver of the transcriptional response to ER stress. In the last years, we found indirect, but compelling evidence that RTCB not only sticks the XBP1 exons together, but also ligates the intron with itself, thereby producing a small, circular RNA molecule.

Circular RNA (circRNAs) molecules have been identified as an abundant and very diverse group of RNAs in the last years, and arise as a new type of molecule that may participate in cell regulation through new, unexplored mechanisms. Most of the circRNAs discovered to date were found in standard RNAseq experiments, and correspond to molecules generated by the nuclear spliceosome. While studying the regulation of XBP1 splicing we realized that current methods to isolate and make sequencing libraries have a very compromised capacity to detect small (<60 nucleotide-long RNAs) circular RNAs, like the XBP1 circIntron.

Currently we are developing a method to identify small circRNAs. In this project we aim to develop the bioinformatic tools needed to analyze our RNAseq libraries to a) identify RNAseq reads corresponding to circRNAs of small size, to then b) determine their abundance, their site of ligation, and c) identify their genomic coordinates. This tool will be of invaluable use to investigate the expression, regulation and biological role of XBP1 circIntron, but also as a general method to uncover the diversity and relevance of this type of RNAs.

In parallel, we will use/adapt other bioinformatic tools to identify small circular RNAs in public, available RNAseq datasets.

OPTATIVAS RECOMENDADAS

· Análisis e interpretación de datos de alto rendimiento
· Aprendizaje automático (*Machine learning*)
· Análisis de secuencias y bioinformática estructural
· Minería de datos (*data mining*) y biología de sistemas