Instituto de Medicina Molecular João Lobo Antunes

Master Project Proposal

Title: Using computational biology tools to identify long noncoding RNAs involved in telomere stability

Synopsis:
Cancer cells need to maintain the ends of their chromosomes, the telomeres, in order to achieve immortality and thus propagate indefinitely. In most cancerous cells, where division is unlimited, telomeres are maintained through the activity of telomerase, but ~15% of cancers use the Alternative Lengthening of Telomeres (ALT) mechanism. The ALT pathway is found to be active in particularly aggressive tumors often developed during childhood, such as juvenile osteosarcomas and glioblastomas. Long noncoding RNAs (lncRNAs) are transcripts over 200 base-pair long that are not translated into proteins. We previously reported that ALT telomeres are heavily transcribed by RNA polymerase II into the lncRNA TERRA, which engages in DNA:RNA hybrids with the telomeric DNA duplex (telomeric R-loops or telR-loops). The ARRET and ARIA lncRNAs are also transcribed from subtelomeres and telomeres. Like TERRA, other lncRNAs are involved in chromosome conformation and maintenance. These include Xist, responsible for female X-chromosome inactivation in mammals, and Hox Antisense Intergenic RNA (HOTAIR), which helps regulate the Hox clusters of genes through long-range interactions. We intend to use publicly available and in-house transcriptomic data to identify new lncRNAs that are involved in telomeric and chromosomal integrity. We will compare the expression of wild-type and perturbed ALT and telomerase-positive cells and use differential gene-expression techniques to select candidate lncRNAs. We will then validate the presence of the candidates in cells using molecular and cell biology. This project will potentially open new avenues to fully understand ALT and possibly to develop novel therapeutic approaches.

A successful Masters applicant will learn how to process and analyze RNA-seq data and will interpret the results. Time-permitting, the student will also learn the bench techniques necessary to validate their presence. Previous experience with computational biology tools is preferable, but not essential.

Supervisor: *Claus M. Azzalin, CMAzzalin Lab, cmazzalin@medicina.ulisboa.pt*
Co-Supervisor:*Filipe Tavares-Cadete, CMAzzalin Lab, filipe.tavarescadete@medicina.ulisboa.pt*
Webpage of the group

Bibliography:
Telomeric Repeat-Containing RNA and RNA surveillance factors at mammalian chromosome ends. C.M. Azzalin, P. Reichenback, L. Khoriauli, E. Giulotto and J. Lingner (2007). *Science* 318: 798-801.

The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Derrien, T., et al. (2012). *Genome Research* 22: 1775-1789.

Remunerated or volunteer training: *volunteer*