Universitat Pompeu Fabra Barcelona

Master in Bioinformatics for Health Sciences

Master project 2021-2022

Personal Information

Supervisor	Anthony Mathelier
Email	anthony.mathelier@ncmm.uio.no
Institution	Centre for Molecular Medicine Norway, University of Oslo
Website	https://mathelierlab.com
Group	Computational Biology & Gene Regulation / Mathelier Group

Project

Computational genomics

Project Title:

PREDICTING EVOLUTIONARY CONSERVED PRIMARY MICRORNA TRANSCRIPTION START SITES ACROSS SPECIES

Keywords:

microRNA; evolution; gene regulation; transcription factors

Summary:

An internship for a research Master student in the field of Computational Biology / Bioinformatics is available at the Computational Biology & Gene Regulation group, Centre for Molecular Medicine Norway, University of Oslo, led by Anthony Mathelier. The group develops computational methods and tools to analyze the regulation of gene expression and the mechanisms by which it can be disrupted in human diseases such as cancers. See https://mathelierlab.com/ for further information. MicroRNAs (miRNAs) represent a class of small (~22nt) RNAs that post-transcriptionally regulate gene expression through mRNA degradation and/or translational repression. They are derived from primary miRNAs (pri-miRNAs), which are long RNAs (up to hundreds of kilobases) transcribed by RNA polymerase II. Expression of miRNAs must be accurately controlled as they are involved in key cellular processes and deregulation is associated with diseases such as cancer. Despite decades of active research on the identification of miRNAs, the understanding of their transcriptional regulation is very limited due to a lack of precise pri-miRNA annotations and the location of corresponding promoters driving their transcription. As part of the FANTOM consortium, we and others recently predicted transcription start sites (TSSs) of several pri-miRNAs, dedicated to controlling miRNA expression across hundreds of human and mouse samples. This work highlighted that pri-miRNA promoters seem to be more evolutionary conserved than protein-coding gene promoters. Moreover, distances between pri-miRNA TSSs and miRNAs seems overall conserved between the two species. Taking advantage of these findings, and the recently available high-quality miRNA complements of 45 species, stored in the state-of-the-art MirGeneDB (http://mirgenedb.org) database, the selected student will investigate how to use evolutionary conservation to predict pri-miRNA TSSs across species. The analyses will be complemented with the analysis of cis-regulatory conservation across species by looking at transcription factors binding at the determined promoter regions using the computational tools and resources developed in the Mathelier group. This work is part of a collaborative effort between the Mathelier (NCMM, UiO, Oslo, Norway) and the Fromm (UiT, Tromsø, Norway) groups for the annotation of primary miRNAs across major metazoan groups. The work will expose the trainee to the analysis of genomics data and the development of computational tools.

Expected skills::

Python, R, or bash

Possibility of funding::

To be discussed

Possible continuity with PhD: :

To be discussed
