Universitat Pompeu Fabra Barcelona | Master in Bioinformatics for Health Sciences

Master project 2020-2021

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

Project

Computational genomics

Project Title:

A pan-cancer computational study of the interplay between transcription factor binding, DNA methylation, and enhancer activity

Keywords:

DNA methylation, Transcription Factor, Enhancer, Quantitative Trait Loci, Machine Learning

Summary:

Methylation of DNA is a prominent DNA modification linked to gene expression alteration in cancers [1,2]. While DNA methyltransferase (DNMT) enzymes de-methylate DNA, Ten-Eleven Translocation (TET) proteins are involved in demethylation. As DNMTs and TETs do not bind DNA in a sequence-specific manner, how these proteins are recruited to their specific sites of action is still an open question. Further, our understanding of the cascading effect of these aberrant DNA methylation patterns on gene expression deregulation is still limited. With a better characterization of the cascading effects of DNA methylation in cancer patients, we could reveal key regulatory networks critical for an improved molecular understanding of the diseases, as we recently showed for breast cancer [3]. In this project, the Master student will perform pan-cancer computational analyses to study the interplay between TF binding, DNA methylation, and enhancer activity. Specifically, the project will aim at (1) unravelling the interplay between DNA methylation and TF-binding in cancer types with limited statistical power and (2) unravelling the interplay between DNA methylation and enhancer activities. 1. We recently computed expression-methylation quantitative trait loci (emQTL) between TF expression and methylation at high-confidence TF-DNA interaction information stored in our UniBind database [4]. emQTL highlighted an interplay between DNA 5mC marks and TF-binding and showed that the binding of key pioneer TFs at their binding sites are likely to trigger local DNA demethylation that could lead to carcinogenesis (unpublished). Unfortunately, the small sample size for some cancer types prohibited the identification of the TFs involved, due to reduced statistical power. The student will use Generative Adversarial Networks (or alike machine-learning approaches) to simulate synthetic data for both methylation and gene expression from available patient data. This approach will alleviate the statistical power bottleneck currently observed. The generated data will be used to perform emQTL analyses and highlight key TFs modulating DNA methylation landscape in these cancer genomes. 2. In the second part of the project, the emQTL framework will be extended to investigate the relationship between DNA methylation and enhancer activity. Specifically, we will use RNA-seq data mapped at enhancers annotated by the FANTOM5 consortium with DNA methylation information from both normal and cancer tissues. The results will be used to investigate how the interplay between DNA methylation, TF binding, and enhancer activity mimics cell fate transition. Indeed, recent reports found that, during cell fate transition, pioneer TFs prime inaccessible enhancers, leading to increased chromatin accessibility and loss of DNA methylation [5]. This project will equip the student with computational biology skills employed in studying gene regulation and cancer genomics. She/he will build computational workflow using Snakemake and scripts in Python, R, and bash. The master student will be introduced to and learn to handle large public cancer genomics data (from ICGC, TCGA,

and BASIS) and gene regulation resources (e.g. UniBind, JASPAR).

References:

1. Suzuki T, Maeda S, Furuhata E, Shimizu Y, Nishimura H, Kishima M, et al. A screening system to identify transcription factors that induce binding site-directed DNA demethylation. Epigenetics Chromatin 2017;10:60. 2. Suzuki T, Shimizu Y, Furuhata E, Maeda S, Kishima M, Nishimura H, et al. RUNX1 regulates site specificity of DNA demethylation by recruitment of DNA demethylation machineries in hematopoietic cells. Blood Adv 2017 3. Fleischer T, Tekpli X, Mathelier A, Wang S, Nebdal D, Dhakal HP, et al. DNA methylation at enhancers identifies distinct breast cancer lineages. Nat. Commun. 2017 4. Gheorghe M, Sandve GK, Khan A, Chèneby J, Ballester B, Mathelier A. A map of direct TF–DNA interactions in the human genome. Nucleic Acids Res. 2019 5. Barnett KR, Decato BE, Scott TJ, Hansen TJ, Chen B, Attalla J, et al. ATAC-Me Captures Prolonged DNA Methylation of Dynamic Chromatin Accessibility Loci during Cell Fate Transitions. Mol. Cell 2020

Expected skills::

Proficiency in Python, R, and/or bash, previous experience in genomics data analysis, team spirit, English proficiency, Exposure to gene regulation and cancer biology will be a plus but not a strict requirement

Possibility of funding::

To be discussed

Possible continuity with PhD: :

To be discussed