



PREDICTIVE MODELS FOR EXPLORATION AND ANNOTATION OF METHYLATED SITES ACROSS A WIDE EPIGENETIC LANDSCAPE



PROMOTORS

Prof. dr. Willem Waegeman
Dr. ir. Gerben Menschaert

TUTOR

Jim Clauwaert

MASTER

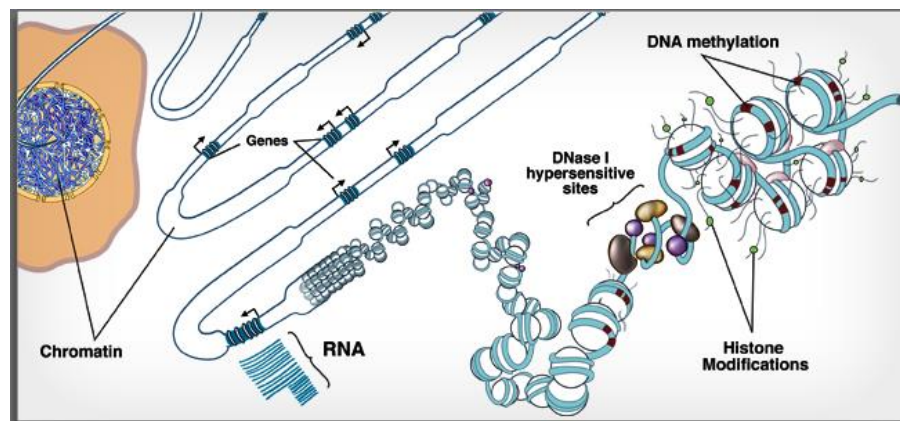
C&G, Bioinformatica

MORE INFO

willem.waegeman@ugent.be
gerben.menschaert@gmail.com
jim.clauwaert@ugent.be

Introduction

Methylation of DNA has proven to have a multitude of roles in prokaryotic genomes. The regulation of gene expression rates and protection of enzymatic activity belong to their best known functions. Blow et. Al. [1] have mapped the methylation sites of > 230 prokaryotic organisms using Single Molecule Real Time (SMRT) sequencing, resulting in > 830 methylation patterns observed. The study is the first to expose the methylation patterns across prokaryotes on this scale, and can be applied for the creation of predictive models w.r.t. the annotation of methylated sites across a multitude of genomes. For this, a wide variety of techniques from the field of artificial intelligence can be adapted, ranging from unsupervised techniques for data clustering to advanced state-of-the-art methods applied for natural language processing (e.g. transformer units).



Aim of the thesis

With the help of an interested and persisting student, this ambitious project aims to gather enough information for publishing, and is therefore perfectly suited for those aiming to lay the groundworks for a future academic position.

The thesis is going to utilize applied methods for data integration and machine learning (given as Predictive Modeling during the 1st semester), for which a strong engagement from the start is required.

[1] <https://doi.org/10.1371/journal.pgen.1005854>

