Epigenetic time series analysis

Kossi Julien Kowou*[1], Eynard Sonia[1], Servin Bertrand†[1], and Pitel Frédérique[1]

[1]UMR 1388 GenPhySE – INRAE – France

Abstract

Understanding the impact of environmental changes through epigenetic modifications, particularly DNA methylation, is essential to understand the regulation of gene expression. To do so, we can rely on a third-generation sequencing technologies, which provides both the base composition of a DNA sequence and its methylation status.

Our study focuses on developing a statistical framework to analyze temporal changes in methylation profiles through the genome. We aim to understand how methylation evolves in the context of selection and if it can potentially impact the regulation of genes associated to trait of interest. To answer this question, we analysed fifteen generations of a breed, each generation being a pool of individuals sperm sample, sequenced using Oxford Nanopore Technology (ONT).

Our statistical framework goes as follows : i) a Poisson log normal model (PLN) (Chiquet and al., 2021) was used to normalize the methylation rate of the CpG island, ii) functional principal component analysis (FPCA, Ramsay and al.,2005) was applied to discriminate CpG islands based on their temporal methylation trajectories. This allowed to identify CpG islands exhibiting outlying patterns of temporal variation. And iii) we clustered temporal trajectories for the outlier CpG islands with a non-parametric clustering method based on Gaussian processes, MAGMACLUST (Leroy and al., 2023). Clusters derived from this model were used to perform an enrichment analysis to identify biological functions associated to each cluster of temporal trajectories.

Knowing the strong link between genetic and epigenetic changes, the observed methylation change patterns will be evaluated under the light of genetic selection signatures identified on the same dataset for the pig breed of interest.

Keywords: Epigenetic, time series, PLN, fonctional data analysis, Gaussian process, Gene Set Enrichment Analysis.

*Speaker

†Corresponding author: bertrand.servin@inrae.fr