Title: Physically-informed machine learning for modelling the dynamics of plant-pathogens molecular interactions

Context

Plants live in a constantly changing environment that happens to be unfavourable or even hostile. Therefore, plant defense against biotic threats requires multiple signaling processes responsible for surveillance, perception, and immune response activation that are influenced by varying spatial and temporal factors. On the top of that, those interactions are based on a molecular dialogue between the pathogen and its host that occur at different time frames, altogether concurring for a successful or unsuccessful infection.

The flourishing of omics techniques has led to the possibility of studying complex biological systems, through the analysis of its content at the molecular level. Transcriptomics is by far the most used omics providing the quantification of change in gene expression. Time-course transcriptomics data are analyzed considering each time point as independent and using approaches based on profile analysis in which the temporal continuity of the data is not fully appreciated. An alternative approach consists in the use of mathematical models including regression and spline models. However, these models generally fail to provide mechanistic interpretations. On the top of that, transcriptome analysis with high resolution in time is challenging, particularly with plant tissues, resulting in longitudinal experiments usually composed of few time points. Those are insufficient to robustly infer statistically significant changes over the infection.

Despite often longitudinal experiments are performed, current available analysis methods do not explicitly consider the time dependency of successive observations.

Tips for the research project:

A novel class of models which is growing in popularity is the physics-informed neural networks (PINN), a novel class of deep learning algorithms that can integrate observational data and physical or mathematical understanding. Those are particularly suitable for high-dimensional, noisy data and time-series, although their application in this domain is very recent. However, all those methods have never been used on omics data, certainly because the physical system is generally unknown in this case. In this project we will investigate the potentiality of these methods and their suitability to analyse information contained in longitudinal multi-transcriptomics data.

Bibliography:

Karniadakis, G. E. *et al.* Physics-informed machine learning. *Nat. Rev. Phys.* 3, 422–440 (2021).

Glyn-Davies, A. *et al.* Φ-Dvae: Physics-Informed Dynamical Variational Autoencoders for Unstructured Data Assimilation. *J. of Comp. Phys.* 115, 113293 (2024).

Liu, X. *et al.* Decomposition of dynamic transcriptomic responses during effector-triggered immunity reveals conserved responses in two distinct plant cell population*s. Plant Communications*, 5, 100882 (2024).

Contact:

Silvia Bottini, Junior Professor Chair INRAe/UniCA – PI team SMILE, silvia.bottini@inrae.fr