

Leveraging scientific machine learning to decode the molecular response of plants exposed to multiple stresses.

Supervisors: Silvia Bottini (Institut Sophia Agrobiotech) – Régis Duvigneau (Inria)

Topic:

A major question in plant biology is to understand how plant growth, development, and environmental responses are controlled and coordinated by the activities of regulatory factors. In particular, infection triggers a dynamic cascade of reciprocal events between host and pathogen wherein the host activates complex mechanisms to recognise and kill pathogens while the pathogen often adjusts its virulence and fitness to avoid eradication by the host. The interaction between the pathogen and the host results in large-scale changes in gene expression. Timely and rapid plant response to these attacks is essential and can dramatically affect plants' fate. Uncovering how transcription factors (TFs) regulate their targets at different molecular levels over time is critical to define gene regulatory networks (GRNs) in normal and diseased states. While several different types of high-throughput experimental procedures are available to study systems in the cell, most only measure static properties of such networks. Despite decades of advancement, challenges remain in GRN inference, including dynamic rewiring.

With this project we propose to develop a hybrid model at the crossroad between mathematics and artificial intelligence to study how the complex GRN in plants is dynamically modulated during pathogens' attack.

Because of the large diversity of hazards and its important agro-economic interest, we will focus on tomato (*Solanum lycopersicum*) as biological model. In Dr. Bottini's team, we have selected and collected high-quality omics data publicly available, which are organized in an internal FAIR database, POMOdOROO (Pan OMics cOLlection of tOmato under biOtic and abiOtic stress). Another important resource developed in our team is TomTom (1), a knowledge graph for tomato gathering 11 databases which represents a fingerprint of a wide type of molecule interactions comprising 113 415 entities and 2 864 036 relationships. Leveraging those two databases, the needed omics data and the biological knowledge to include in the model are already available and ready to be used for the model development. During the last months Dr. Bottini and Dr. Duvigneau have been working together to set up the basis of a hybrid approach based on the formalism of Physics-Informed Neural Networks to model the dynamical changes of gene expression in tomato upon pathogen attack, as an ODE model fitting data, but without considering the interactions among different genes. Here, we wish to upgrade this initial model to study how the complex GRN is rewired in diseases versus non-diseased plants and in response to different stresses.

We wish to obtain at the end of the internship a novel model which combines the power of AI for modeling complex systems, of biological knowledge for the mechanistic explication and of the mathematical model for the predictive ability, to surpass the scale limitations of traditional modeling methods while being fully interpretable.

Scientific partnership:

Dr Bottini and Dr Duvigneau have started their collaboration one year ago, and one master internship between the two teams has just finished (2). The achievement of this project will foster their collaboration and set the basis to apply to other national and international grants.

The SMILE team is a novel team at the Institut Sophia Agrobiotech, with strong skills in computational biology, knowledge graph, omics and multi-omics analysis and integration. The ACUMES team at Inria is expert in ODE/PDE modeling and optimization for multi-disciplinary systems.

This proposal fits naturally the topics of the thematic semester ML+SIM by targeting the construction of a model that hybridizes mathematical models and data-driven ones, based on a complementary collaboration.

- (1) Multari M, Carriere M, Amoros-Gabarron X, Damy A, Lobentanzer S, Saez-Rodriguez J, Jaubert S, Dugourd A, Bottini S. A knowledge graph and topological data analysis framework to disentangle the tomato-multi pathogens complex gene regulatory network. *bioRxiv*; 2025. Available from: <https://www.biorxiv.org/content/10.1101/2025.04.09.647963v1>
- (2) Vidal M, Duvigneau R, Bottini S Modélisation de la dynamique de l'expression des gènes des plantes lors d'une attaque biotique par intelligence artificielle informée par la physique. Inria Research Report No 9595, August 2025 Available from : <https://inria.hal.science/INRIA-RRRT/hal-05210203v1>