PhD PROJECT PROPOSAL

CENTURI
TURING CENTRE
FOR LIVING SYSTEMS

Title of the project

Digital twins of organoids by multi-scale modeling

Lead supervisor: Anaïs Baudot, MMG anais.baudot@univ-amu.fr
Co-supervisor 1: Elisabeth Remy, I2M elisabeth.remy@univ-amu.fr

Abstract
This project aims to develop a digital twin model for muscle organoids. We will infer intra- and inter-cellular regulatory networks from diverse (single-cell, multi, spatial) omics data, and further integrate them in a multi-scale digital twin through advanced techniques like agent-based modeling. The digital twin can then be used to simulate the effects of perturbations induced by mutations or drugs. This project leverages the unique collaboration between Elisabeth Remy's expertise in mathematical modeling and Anaïs Baudot's skills in omics data analysis, further leveraging the MMG unit expertise in stem cell reprogramming and genetic diseases. We expect to improve our understanding of organoids, which are becoming very popular models in biology, and assist in studying muscle-associated genetic diseases, in particular to better predict the effects of genetic mutations and drug treatment.

Keywords
Digital twins; multi-scale modeling; organoids; genetic diseases; network inference; mathematical models

Objectives
Digital twins are mathematical models designed to replicate the dynamics and behavior of their physical counterparts. In cell biology, they have been mainly implemented to model embryos and tumors. We aim here to develop a multi-scale digital twin model of muscle organoids. We will infer intra-cellular and inter-cellular regulatory networks from (single-cell, multi, spatial) omics data and integrate these networks thanks to, e.g., agent-based modeling. We expect to improve our understanding of muscle organoids and assist in studying muscle diseases, gene mutations, and drug discovery.

Proposed approach (experimental / theoretical / computational)
We will first conduct an extensive survey of existing digital twins models and collect multi-modal data for muscle organoids, both from public sources and from the MMG teams and stem cell reprogramming facility. Multi-modal data encompass single-nuclei transcriptomics, bulk multi-omics, and spatial transcriptomics. We will infer intra-cellular gene regulatory models for the various cell types using the framework developed by E. Remy's team (Hérault et al. 2023). Simultaneously, inter-cellular networks will be inferred to model the communication between the different cell types, guided by methods reviewed in Dimitrov et al. (Nature Comm. 2022). The integration of these models into a multi-scale digital twin will finally be done with frameworks such as CompuCell3D (Swat et al. Methods in Cell Bio, 2012) or (agent-based) PhysiBOSS (Ponce de Leon. Nat Comm. 2023). Model calibration and validation will be undertaken, ensuring the digital twins' accuracy in simulating organoid growth, differentiation, response to stimuli, and disease progression, thus laying a foundation for future applications in disease modeling and drug discovery.

Interdisciplinarity

This project is a collaboration between two teams: MABioS (Mathematics and Algorithms for Systems Biology), led by Elisabeth Remy at the Mathematical Institute of Marseille (I2M) situated at Luminy, and the Systems Biomedicine team led by Anaïs Baudot at the Marseille Medical Genetics Center within the Timone Faculty of Medicine. Elisabeth Remy's team specializes in discrete mathematics, graph theory, and combinatorics, focusing on modeling and analyzing gene regulatory networks. Anaïs Baudot's team applies network approaches to extract insights from large-scale omics data, with a focus on investigating human disorders. The project and the accompanying Ph.D. student will have the opportunity to draw upon the expertise of numerous experimental biologists dedicated to monogenic diseases, as well as the resources of the MMG unit's stem cell reprogramming facility. We will also benefit from a collaboration with Laurence Calzone (Curie Institute, Paris), who contributed to the agent-based PhysiBOSS framework.

PhD student's expected profile

The proposed project will require strong analytical and computational skills. We hence seek for PhD students with either: i) Master's degree in mathematics and interest in biological modeling or ii) Master's degree in an area related to Computational Biology with interest for data analysis and mathematical modeling.

Is this project the continuation of an existing project or an entirely new one?
In the case of an existing project, please explain the links between the two projects

This project builds on the collaboration of Elisabeth Remy and Anaïs Baudot. They co-supervise the PhD thesis of Nadine Ben Boina, supported by Centuri from 2021 to 2024. Nadine is developing a logical model for Dysferlinopathies, rare muscle diseases tied to *DYSF* gene mutations. In addition, Nadine is proposing an approach highlighting the links between different updating policies to refining Boolean models. We aim to expand this logical modeling framework into digital twins, drawing on the expertise of our supervisors in single-cell and omics data analysis, mathematical modeling, and rare genetic disease research.

Two to five references related to the project

- Ponce-de-Leon, Miguel, Arnau Montagud, Vincent Noël, Annika Meert, Gerard Pradas, Emmanuel Barillot, Laurence Calzone, and Alfonso Valencia. 2023. "PhysiBoSS 2.0: A Sustainable Integration of Stochastic Boolean and Agent-Based Modeling Frameworks." Npj Systems Biology and Applications 9 (1): 54.
- Varela PL, Ramos CV, Monteiro PT, Chaouiya C. "EpiLog: A software for the logical modeling of epithelial dynamics." F1000Res. 2019 Mar 11;7:1145.
- Argiro L, Chevalier C, Choquet C, Nandkishore N, Ghata A, Baudot A, et al. "Cardiopharyngeal Mesoderm specification into cardiac and skeletal muscle lineages in gastruloids." bioRxiv; 2023. [A collaborative effort to analyze time-series single cell dataset of an embryonic organoid]
- Hérault L, Poplineau M, Duprez E, and Remy E. "A novel Boolean network inference strategy to model early hematopoiesis aging." Computational and Structural Biotechnology Journal, 21:21--33, 2023. [We have developed a pipeline to infer a Boolean model from scRNAseq data, enabling us to understand the impact of aging on HSC differentiation.]

Two main publications from each PI over the last 5 years

- Cantini L, Zakeri P, Hernandez C, Naldi A, Thieffry D, Remy E, Baudot A. "Benchmarking joint multi-omics dimensionality reduction approaches for the study of cancer." Nature Communications. 2021 Dec.

An extensive benchmark of multi-omics data integration approaches, demonstrating our expertise in leverage data from different sources to better understand the molecular mechanisms of diseases.

- Baptista A, Gonzalez A, Baudot A. "Universal multilayer network exploration by random walk with restart." Commun Phys. 2022 Jul 1;5(1):1–9.

The implementation of an innovative network algorithm able to explore and extract information from heterogeneous network sources.

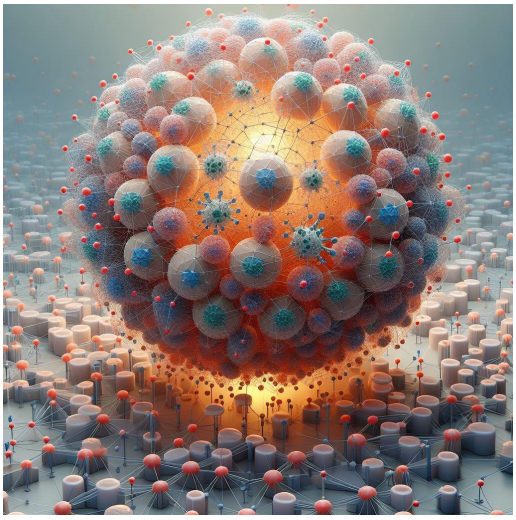
- Chapman S, Duprez E, and Remy E. "Logical modeling of myelofibrotic microenvironment predicts dysregulated progenitor stem cell crosstalk." BioSystems, 09 2023.

In the context of myelofibrosis, we have constructed an intercellular logical model focusing on the crosstalk between progenitor Haematopoietic Stem Cells and neighboring mesenchymal stem cells, that is perturbed during the disease progression.

- Remy E and Ruet P. "From multivalued to Boolean functions: preservation of soft nested canalization." Theoretical Computer Science, Volume 982 (To appear).

Theoretical study to characterize the class of canalizing multivalued functions, which are over-represented in genetic regulatory models.

Project's illustrating image



Generated with DALLE. 3

Prompt: I would like an abstract representation of a group of stacked cells organized in a spheroid, each cell having a small molecular network inside, and the cell being connected by edges to represent cell communication