UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

DOTTORATO DI RICERCA IN Tecnologie Convergenti per i Sistemi Biomolecolari – XL CICLO

Research Topic ID: XL – 1.8

Proponent: Prof. Chiara Damiani

Project Title: Resolving metabolic heterogeneity in time and space via data integration

Scientific background and 'open issues'

Metabolism is inherently spatially and temporally heterogeneous. This heterogeneity arises from diverse cellular environments, cellular interactions, and temporal fluctuations in cellular processes. Characterizing spatio-temporal metabolic heterogeneity provides valuable insights into fundamental biological phenomena, including cellular differentiation, disease progression, and microbial interactions. Moreover, such characterization is essential for developing targeted therapeutic strategies, biomarker discovery, and biotechnological advancements.

In recent years, significant progress has been made in methodologies for characterizing spatio-temporal metabolism. Advanced imaging techniques enable the visualization and quantification of metabolites within cellular and tissue contexts. These techniques offer spatial resolution at the subcellular level, allowing researchers to probe metabolic activities within specific cellular compartments or microenvironments.

Furthermore, the integration of omics data, such as transcriptomics, proteomics, and metabolomics, provides a comprehensive understanding of metabolic regulation across spatial and temporal scales. Multi-omics approaches enable the correlation of metabolic phenotypes with molecular signatures, facilitating the identification of metabolic pathways underlying cellular functions and diseases.

Computational modelling plays a crucial role in elucidating spatiotemporal metabolic dynamics. Mathematical models, ranging from agent-based models to genome-scale metabolic models (GEMs), simulate metabolic processes within cellular environments, predicting metabolic fluxes, and metabolic responses to perturbations.

Emerging technologies, including single-cell and spatial transcriptomics, offer unprecedented opportunities to dissect cellular heterogeneity and spatial organization of metabolism at single-cell resolution. These cutting-edge approaches provide valuable data for constructing spatially resolved metabolic networks and uncovering novel metabolic interactions and regulatory mechanisms.

Objectives

The PhD candidate is expected to develop a computational framework that characterizes the metabolic program of cells within a mixed population, derived, for example, from a biopsy or from a 3D cell culture, at the single-cell and spatial resolution. The framework ideally

should also be able to characterize the evolution of the population in time after a genetic or environmental perturbation. The framework must be generalizable to different case studies. With the dual goal of validating the framework and obtaining new hypotheses to be tested in the wet-lab, the PhD candidate will apply the framework to: i) datasets produced within projects in which the proponent is involved, including embryo development datasets and cancer datasets; ii) public datasets.

Methodologies

Within the expanding field of genome-wide metabolic network modelling, many frameworks have been proposed to predict metabolic fluxes from gene expression data, ideally coupled with exo-metabolomics data [1,2]. Machine learning models, especially graph neural networks, are emerging either as alternative approaches to constraint-based modelling or as complementary tools to enhance its capabilities [3]

Chiara Damiani's research group has proposed a method to combine such predictions with information from intracellular metabolomics to discern transcriptional from metabolic regulation of metabolic fluxes [4]. The group has also proposed a framework to model metabolic interactions among cells, integrating single-cell data, but neglecting spatial phenomena [5]. The framework might be extended to take into account spatial interactions, leveraging spatial-transcriptomics datasets.

A challenge in spatial modelling consists in defining the source of nutrients and their diffusion. In this regard, some preliminary work combining Cellular Pots Models and Flux Balance Analysis has been put forward [6]. The visual representation of the simulation outcome of such complex models could be integrated with High Content Imaging.

Collaboration / Co-tutoring opportunities

The PhD candidate will collaborate with:

- Davide Maspero, who is Post-doc fellow at the Single Cell Genomics Team, CNAG – CRG in Barcelona and is an expert in spatial-transcriptomic data analysis methods,
- the Data and Computational Biology of the Department of Computer Science of the University of Milan Bicocca, which has experience in genomics data analysis,
- Martello'lab from the University of Padova with expertise in 3D human models of embryo development.
- Collaborations will be active also with the Institute of Bioimaging and Molecular Physiology of CNR.

Project's Sustainability & Mobility

Chiara Damiani runs the lab of Bioinformatics and Computational Biology at the Department of Biotechnology of Biosciences. The lab has long-standing experience in metabolic network modelling and omics data science. The lab develops user-friendly tools, such as Marea4Galaxy [7], to integrate omics data into metabolic networks

The paper from our lab presenting the single-cell Flux Balance Analysis framework [5], which serves as a precursor to this PhD project, has earned attention, accumulating up to 77 citations (source Google Scholar).

Chiara Damiani has a long-standing collaboration with the group of Prof. Marco Vanoni, PI of the UNIMIB unit of the PNRR infrastructure project Elixir NextGenerationIT and of the Sysbio-ISBE JRU and has full access to High Content facilities.

A suitable institution for mobility is the Single Cell Genomics Team at Centre Nacional d'Anàlisi Genòmica in Barcelona.

References

[1] Gopalakrishnan, S., Joshi, C. J., Valderrama-Gómez, M. Á., Icten, E., Rolandi, P., Johnson, W., ... & Lewis, N. E. (2023). Guidelines for extracting biologically relevant context-specific metabolic models using gene expression data. *Metabolic Engineering*, 75, 181-191.
[2] Machado, D., & Herrgård, M. (2014). Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS computational biology*, 10(4), e1003580.
[3] Alghamdi, N., Chang, W., Dang, P., Lu, X., Wan, C., Gampala, S., ... & Zhang, C. (2021). A graph neural network model to estimate cell-wise metabolic flux using single-cell RNA-seq data. *Genome research*, *31*(10), 1867-1884.
[4] Di Filippo, M., Pescini, D., Galuzzi, B. G., Bonanomi, M., Gaglio, D., Mangano, E., ... & Damiani, C. (2022). INTEGRATE: Model-based multi-omics data integration to characterize multi-level metabolic regulation. *PLoS computational biology*, *18*(2), e1009337.
[5] Damiani, C., Maspero, D., Di Filippo, M., Colombo, R., Pescini, D., Graudenzi, A., ... & Mauri, G. (2019). Integration of single-cell RNA-seq data into population models to characterize cancer metabolism. PLoS computational biology, 15(2), e1006733.
[6] Maspero, D., Damiani, C., Antoniotti, M., Graudenzi, A., Di Filippo, M., Vanoni, M., ... & Pescini, D. (2020). The influence of nutrients diffusion on a metabolism-driven model of a multi-cellular system. *Fundamenta Informaticae*, 171(1-4), 279-295.
[7] Damiani, C., Rovida, L., Maspero, D., Sala, I., Rosato, L., Di Filippo, M., ... & Mauri, G. (2020). MaREA4Galaxy: Metabolic reaction enrichment analysis and visualization of RNA-seq data within Galaxy. *Computational and structural biotechnology journal*, 18, 993-999.