University of Guadalajara

Center for University Studies in Exact Sciences and Engineering

Division of Technologies for Cyber-Human Integration

**Logo

Description automatically generated**

Precision Oncology for Breast Cancer Treatment

Graduate Protocol

Master of Science in Bioengineering and Intelligent Computing

Presented by

Eduardo Ruiz Robles

Thesis Director

Dr. Germán Andrés Preciat González

Codirector

Dra. Martha Patricia Gallegos Arreola

Tutor Tutor

Dr. Ernesto Borrayo Carbajal Dra. Ana Maria Puebla Pérez

Tutor

Dra.Patricia Martins Conde

Guadalajara, Jalisco, February 11, 2024

## **Abstract:**

## Breast cancer is the most common cancer among women worldwide, causing over 685,000 deaths and 2.3 million new cases in 2020, according to the World Health Organization. Its heterogeneity, reflected in different subtypes and unique tumor characteristics, requires personalized treatments to address this complexity.1

## This study aims to develop genome-scale metabolic models by integrating clinical and omic data, such as glucose levels, oxygen levels, and tumor microenvironment characteristics, using constraint-based modeling (COBRA Toolbox) and machine learning techniques. This interdisciplinary approach will enable the correlation of metabolic phenotypes with key clinical data to identify predictive biomarkers, specific metabolic subgroups, and personalized therapeutic responses. The goal is to overcome current limitations in diagnostic and therapeutic precision, providing more effective tools for managing breast cancer.

**Description**

In recent years, predictive models based on genomic data have been developed, though few achieve the necessary accuracy to be effectively implemented in clinical practice and public health. This limitation is partly due to the variability introduced by clinical data, such as blood biometrics, lipid profiles, electrolyte levels, electrocardiogram (ECG) records, and tumor microenvironment characteristics, which affect the precision of models relying solely on genomic data. Including multiple data sources such as transcriptomics or metabolomics improves the accuracy of metabolic models, suggesting that adding clinical data could also enhance precision.3

Breast cancer involves alterations across various biomolecular layers, with metabolism being one of the affected processes that influence both cellular behavior and the tumor microenvironment. Genome-scale metabolic models (GEMs) are detailed reconstructions of cellular metabolic networks, enabling the analysis of metabolism at the cellular level in different contexts. These models integrate metabolic reactions, mass balance, and associations between genes and proteins involved in metabolism. Cancer metabolism can be studied through the combination of GEMs and constraint-based modeling (COBRA), allowing for the evaluation of how genetic and environmental factors influence specific phenotypes. COBRA modeling represents the flux of metabolic reactions within a system, incorporating uptake and secretion rates to depict the system's phenotype. These tools can be tailored to specific cancer conditions, facilitating the identification of potential therapeutic targets in tumor metabolism.4

The primary aim of this project is to integrate patient-specific clinical data into genome-scale models using COBRA Toolbox constraint-based modeling. Additionally, this study seeks to correlate metabolic phenotypes with clinical data using statistical techniques (Pearson correlation) and machine learning methods, including Support Vector Machines (SVM) for their robustness in classifying high-dimensional data, neural networks for identifying non-linear relationships, and k-means clustering to group patients with similar metabolic characteristics. By correlating metabolic phenotypes with specific clinical features, this study aims to identify patterns useful for predicting responses to specific treatments and uncover key biomarkers, optimizing personalized treatment strategies for breast cancer.

**Hypothesis:**

There is a significant correlation between metabolic phenotypes derived from genome-scale metabolic models (GEMs) and clinical characteristics of patients, suggesting that cellular metabolic variability is related to the clinical information of each breast cancer patient.

**Goals and Objectives:**

**General Objective:**

To develop personalized genome-scale metabolic models (GEMs) that integrate relevant clinical data to analyze their relationship with cancer progression and/or prognosis in patients, using statistical and machine learning tools.

**Goal 1:** Integration of Clinical Data and Metabolic Modeling

Objective 1.1: Collect and preprocess clinical data from breast cancer patients for integration into the analysis.

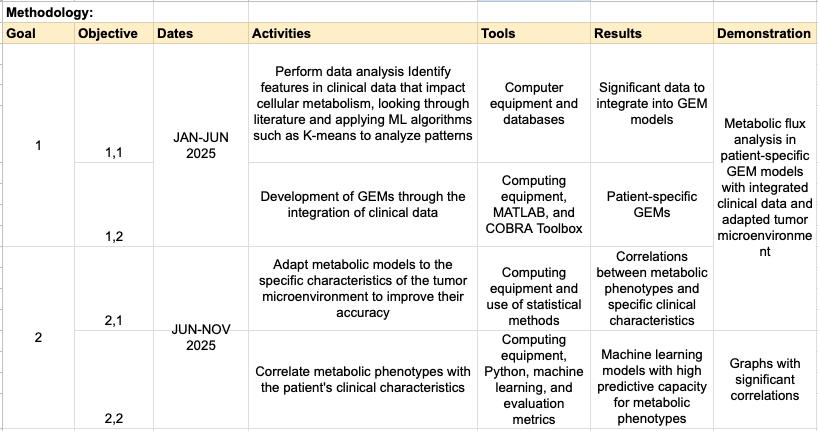
Objective 1.2: Design and construct personalized genome-scale metabolic models (GEMs) incorporating clinical information associated with cancer progression and/or prognosis.

**Goal 2:** Advanced Analysis and Adaptation to the Tumor Microenvironment

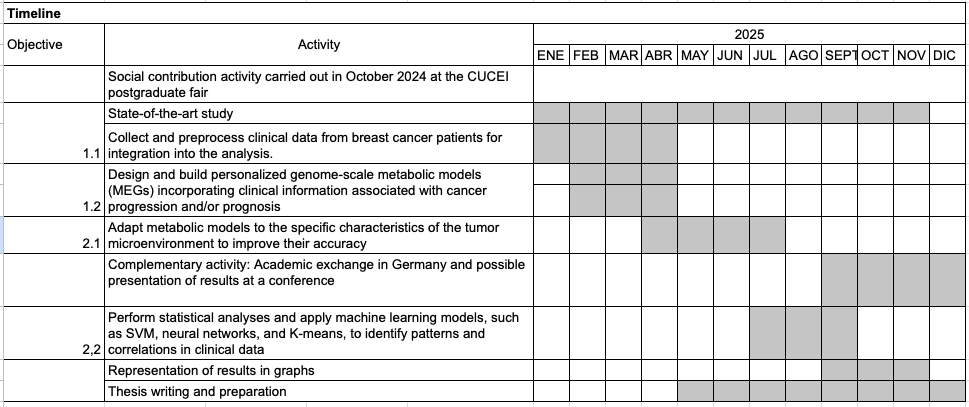
**Objective 2.1:** Adapt metabolic models to the specific characteristics of the tumor microenvironment to improve their accuracy.

**Objective 2.2:** Perform statistical analyses and apply machine learning models, such as SVM, neural networks, and K-means, to identify patterns and correlations in clinical data.

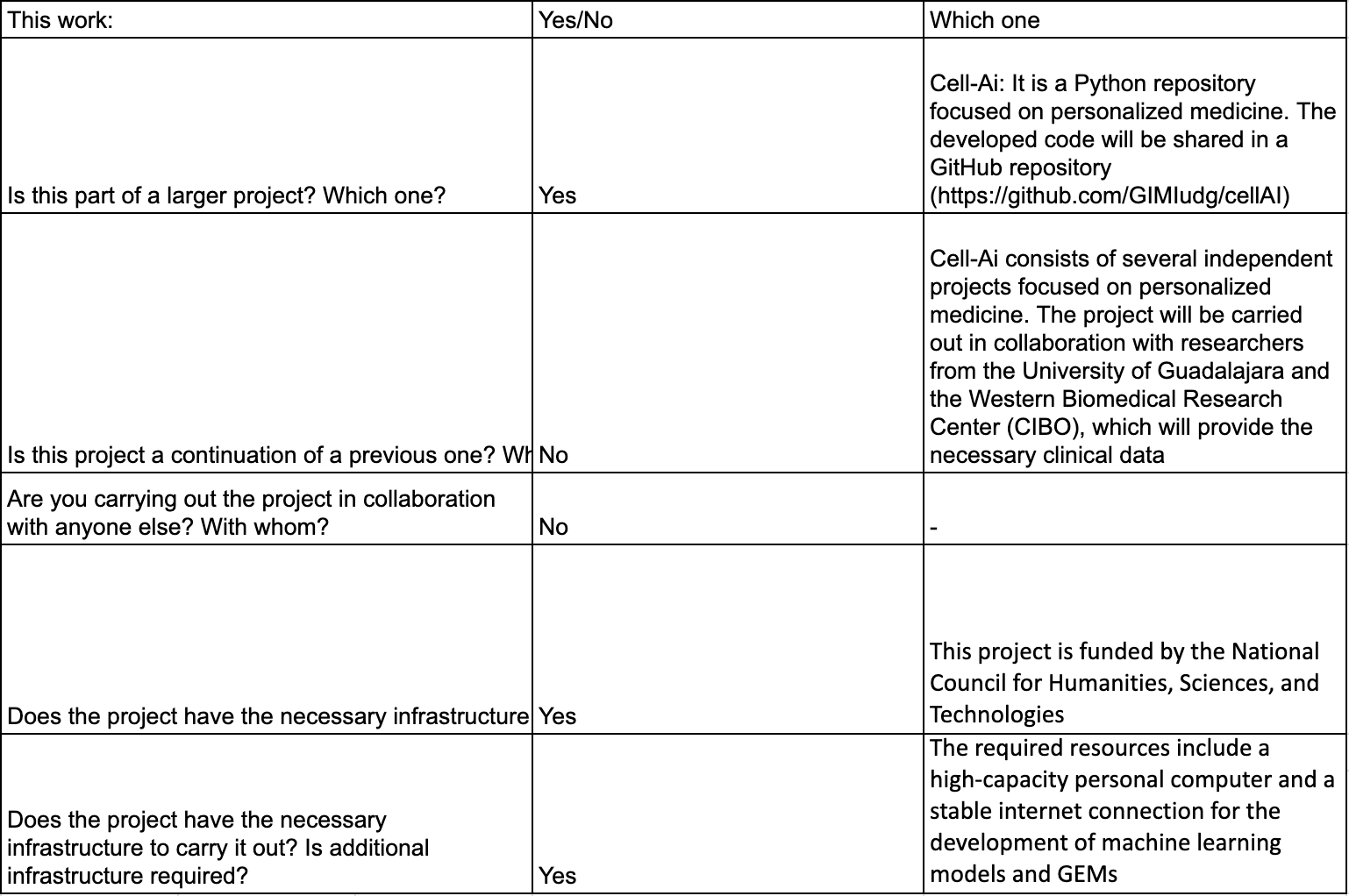
**Methodology:**



**Timeline:**

****

**Feasibility and follow-up**



**Expected Results and Impact:**

The expected results are: 1) identify specific patterns and correlations between metabolic phenotypes and clinical characteristics of patients with a predictive accuracy of at least 85% in therapeutic responses; 2) develop GEMs for different types of patients by integrating relevant clinical data; and 3) correlate metabolic phenotypes with clinical characteristics of patients to predict responses to specific treatments or disease prognosis. These results would contribute to a deeper understanding of the metabolic pathways that affect breast cancer progression and resistance to certain treatments. Furthermore, they would allow the identification of new metabolic biomarkers and improve the predictive capability for therapy responses, facilitating more effective and personalized treatments. A significant challenge is the lack of uniformity in clinical data, which includes both qualitative and quantitative variables, complicating their integration into metabolic models. To address this problem, the literature will be reviewed to identify clinical data, and ML algorithms such as K-means will be applied to detect patterns in the data that may be associated with cancer metabolism, with the goal of incorporating this information into GEMs

**References:**

1. 1. Papalexis, P., Georgakopoulou, V. E., Drossos, P. V., Thymara, E., Nonni, A., Lazaris, A. C., et al. (2024). Precision medicine in breast cancer (Review). Molecular and Clinical Oncology, 21, Article 78. https://doi.org/10.3892/mco.2024.2776
2. 2. López de Maturana, E., Alonso, L., Alarcón, P., Martín-Antoniano, I. A., Pineda, S., Piorno, L., et al. (2019). Challenges in the integration of omics and non-omics data. Genes, 10(3), Article 238. https://doi.org/10.3390/genes10030238
3. Preciat, G., Lucumi Moreno, E., Wegrzyn, A. B., Willacey, C. C. W., Modamio, J., Monteiro, F. L., et al. (2021). Mechanistic model-driven exometabolomic characterisation of human dopaminergic neuronal metabolism. bioRxiv. https://doi.org/10.1101/2021.06.30.450562
4. 3. Ng, R. H., Lee, J. W., Baloni, P., & Diener, C. (2022). Constraint-based reconstruction and analyses of metabolic models: Open-source Python tools and applications to cancer. Frontiers in Oncology, 12, Article 914594. <https://doi.org/10.3389/fonc.2022.914594>