Establishing a Baseline for a Healthy Gut Microbiome

Artur Gomes and Óscar Dias
University of Minho
2024/2025

Index

01

Contextualization and Motivation

1.1 Gut Microbiome in Health and Disease

Gut Microbiome

- Diverse and active metabolism
- Physiological contribution
- Maintain homeostasis
- Protects against pathogens
- Modulates inflammatory response

- Composition influenced by many factors

Dysbiosis

Disruption of microbial balance

Associated with a wide range of diseases

Leads to impaired metabolic and immune functions

Host Health

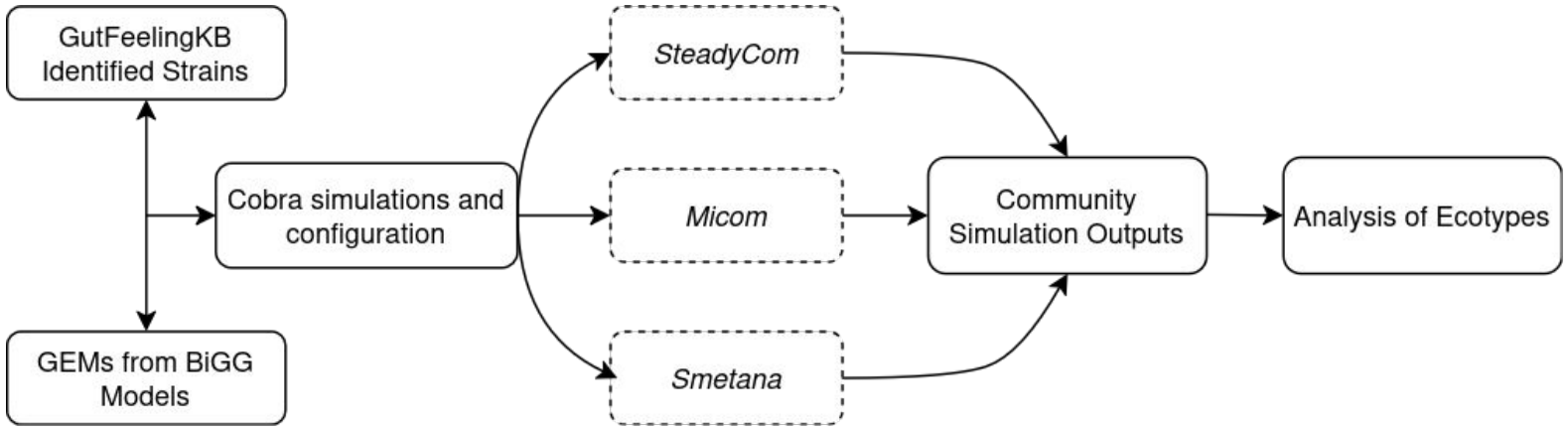
- Influenced by microbiome's

Interventions

Probiotics, prebiotics, dietary modulation

Antibiotic or FMT (Fecal microbiota transplantation)

1.2 Aims and Plan



- ➔ Use GutFeelingKB to identify healthy/dysbiotic strains
- ➔ Retrieve models from BiGG and test with Cobra
- ➔ Simulate communities using SteadyCom, Micom, and Smetana

02

Materials and Methods

2.1 GFKB and Bigg Database

GutFeelingKB

Curated database identifying bacterial strains associated with **healthy** and **dysbiotic** gut microbiomes.

Used to select **representative microbial species** for community modeling.

BiGG Models

Database of high-quality **genome-scale metabolic models (GEMs)**

Provides pre-constructed GEMs for many gut bacteria

Ensures standardized and curated metabolic reconstructions

Gut Models

These assembled models are used in **community simulations**

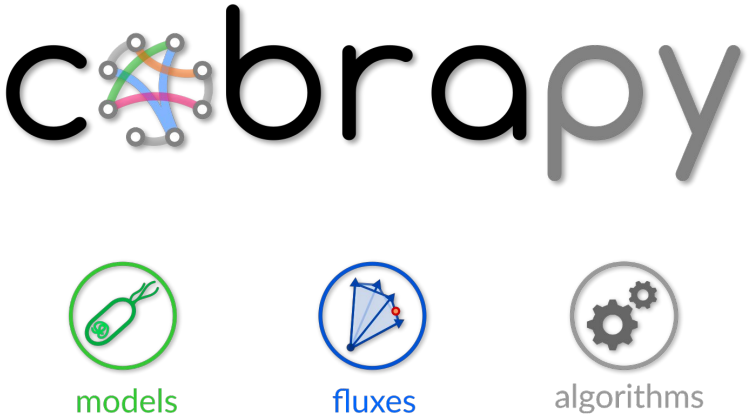
2.2 Constraint-Based Reconstruction and Analysis

COBRA (*Constraint-Based Reconstruction and Analysis*) is a framework for analyzing the **metabolic capabilities** of organisms using **GEMs (genome-scale metabolic models**.

It models metabolism as a **network of reactions** constrained by:

- Mass balance
- Thermodynamics
- Nutrient availability

One of the most common approach in COBRA is **Flux Balance Analysis (FBA)**, which estimates the **optimal distribution of metabolic fluxes** through a metabolic network that maximizes or minimizes a specified objective function.



2.3 Smetana, Micom, SteadyCom

Smetana

Models microbial communities by inferring **metabolic interactions** through **metabolite exchange**, without using growth rates or biomass optimization. It identifies **cross-feeding** and **dependencies**, making it ideal for analyzing metabolic complementarity.

MICOM

Models microbial communities using known **species abundances**, without assuming equal growth. It balances individual and community growth using a **cooperative trade-off strategy**, and estimates species-specific flux distributions under **steady-state conditions**.

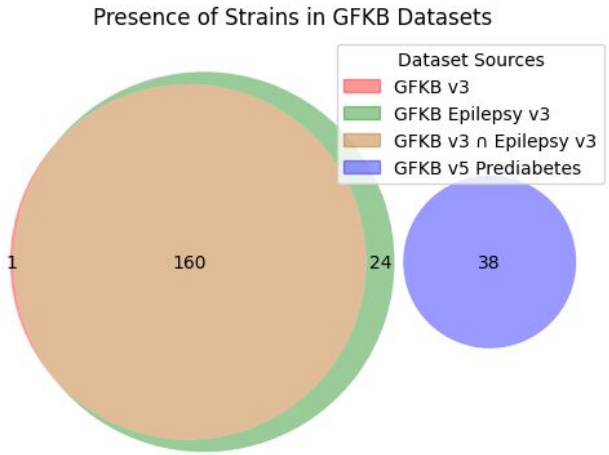
SteadyCom

Models stable microbial communities by assuming all species **grow at the same rate** and infers the **species abundances** that make this possible. It efficiently predicts **community composition** and **fluxes** under steady-state conditions.

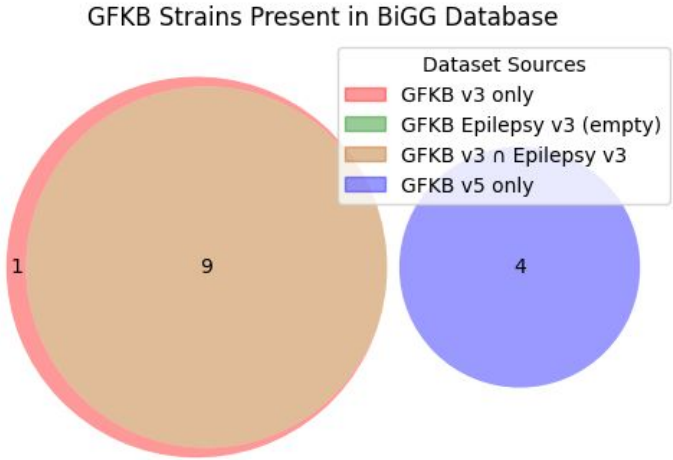
03

Results

3.1 Cross Referencing (GFKB and BiGG)



Number of microbial strains found in different updates of the **GutFeelingKB** database — including **v3**, **epilepsy v3 subset**, and the **prediabetes v5**.



Shows only the strains from each version that also have **genome-scale metabolic models available in the BiGG database**. For this study, we chose 13 of this 14 models to have a more close distribution to the original one.

3.1.1 Chosen Models from BiGG

UP name	Present in GFKB v3	Present in GFKB v5	BiGG ID
Escherichia coli BL21-DE3	Y	N	iEC1356_BL21DE3
Escherichia coli SMS-3-5	Y	N	iEcSMS35_1347
Escherichia coli UTI89	Y	N	iUTI89_1310
Escherichia coli O17:K52:H18	Y	N	iECUMN_1333
Escherichia coli O25b:H4-ST131	Y	N	iECSF_1327
Escherichia coli O6:H1	Y	N	iC_1306
Escherichia coli O78:H11	Y	N	iETEC_1333
Escherichia coli O83:H1	Y	N	iNRG857_1313
Escherichia coli UMNK88	Y	N	iUMNK88_1353
Peptoclostridium difficile	Y	N	iCN900
Escherichia coli SE11	N	Y	iECSE_1348
Klebsiella pneumoniae	N	Y	iYL1228
Shigella dysenteriae serotype 1	N	Y	iSDY_1059

Escherichia coli O25b:H4-ST131

A multidrug-resistant, pathogenic *E. coli* lineage linked to gut dysbiosis and extra-intestinal infections.

Clostridium difficile

Known for its role in antibiotic-associated colitis and dysbiosis — a key marker of a disrupted gut microbiome.

Klebsiella pneumoniae

An opportunistic pathogen often linked to gut dysbiosis, inflammation, and antimicrobial resistance.

Shigella dysenteriae serotype 1

A highly virulent enteric pathogen responsible for severe inflammatory diarrhea and gut epithelial damage.

3.2 Models Individual Simulations

1st → **individual simulations for each metabolic model** retrieved from the BiGG database.

For each strain, I generated a **model summary** and calculated its **minimum growth-supporting medium** using COBRApy. These simulations were essential for:

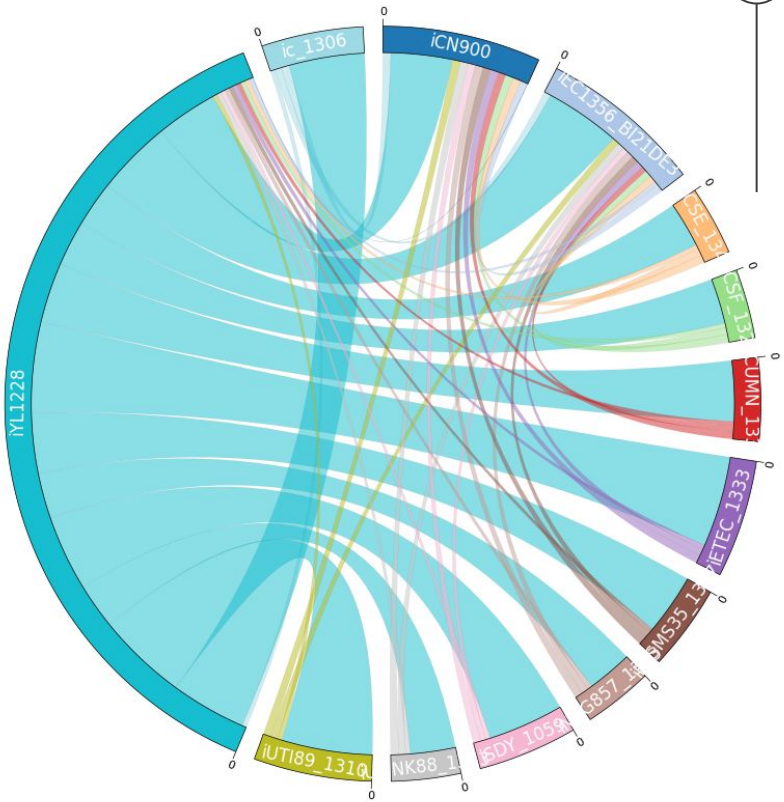
- ➢ Validating model functionality
- ➢ Identifying blocked reactions
- ➢ Preparing for community-level simulations under shared constraints

During this step, we discovered that the model **iCN900** (Clostridium difficile) had a predefined environment that did not allow biomass production under typical minimal conditions. This required special attention, as it could distort community simulation results if not adjusted.

3.3 Community Simulations – Smetana

Interpretation:

- **Each segment** on the circle represents a microbial strain (e.g., *ML1228*, *iCN900*, *ETEC 1333*)
- **Arcs (connections)** connecting the strains represent **cross-feeding interactions**, where a metabolite is transferred from a **donor** to a **receiver**.
- The **thickness** of each connection reflects the **SMETANA score**, which quantifies the strength or likelihood of the metabolic interaction.
- The **color of the arc** corresponds to the **donor strain**, making it easier to trace who supplies metabolites.



Key observation:

- *Klebsiella pneumoniae* (e.g., *iYL1228*) dominates the interaction network, acting as a **major donor**, supporting many others.
- Strains with **fewer outgoing links** might be more dependent or less metabolically flexible in the simulated environment.

3.3 Community Simulations – SteadyCom

The **abundances of individual strains** were estimated using **SteadyCom** under a shared, defined **minimal medium**, constructed by merging the individual minimal media computed with COBRApy.

Strain ID	Abundance
iYL1228	0.6238
iSDY_1059	0.1549
iCN900	0.0126
iECSE_1348	0.0
iUMNK88_1353	0.0022
iETEC_1333	0.1673
iECSF_1327	0.0
iEcSMS35_1347	0.0012
iEC1356_BL21DE3	0.0
iECUMN_1333	0.0365
iNRG857_1313	0.0013
iUTI89_1310	0.0
ic_1306	0.0001
Community growth	3.0254

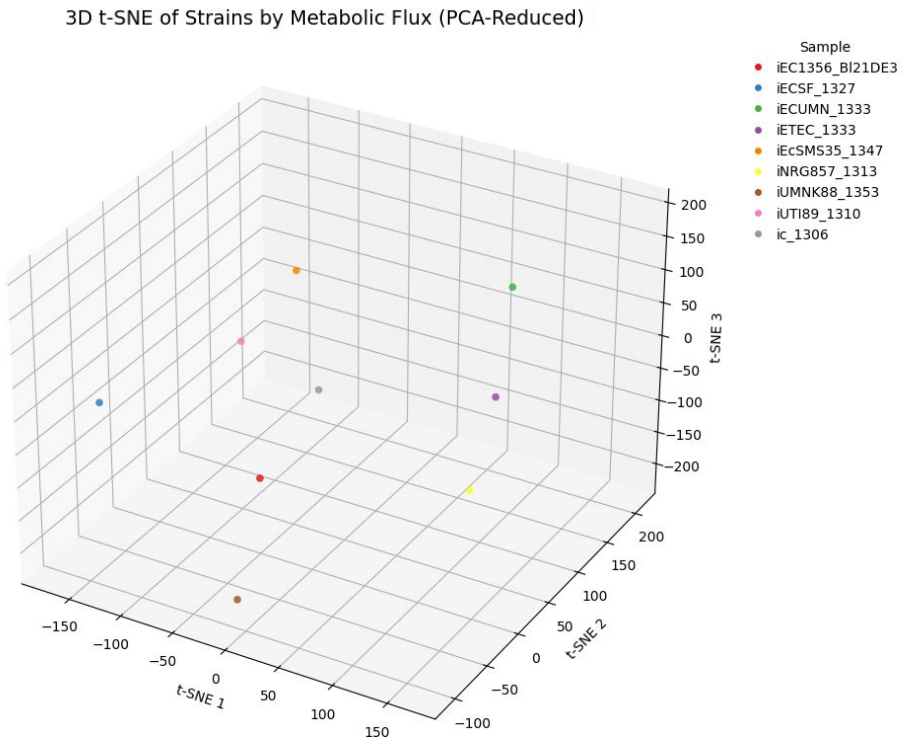
In this constrained environment, most strains were assigned **very low or zero abundance**, indicating that they **do not significantly contribute** to the stable community under the given nutrient conditions.

A few strains — notably **iYL1228** (*Klebsiella pneumoniae*), **iETEC_1333** (*Escherichia coli* O78:H11), and **iSDY_1059** (*Shigella dysenteriae*) — were assigned **higher relative abundances**, suggesting that these strains are **less metabolically efficient** and require **more biomass** to support the same growth rate as the others.

3.3 Community Simulations – MICOM

The plot presents a **3D t-SNE projection** of metabolic flux profiles for each strain simulated using **MICOM**.

Each point represents a strain's **flux distribution across all internal reactions**, reduced to three dimensions for visualization. The data was generated from simulations using the **cooperative trade-off strategy**, which balances individual growth with community benefit.



The clear separation between points suggests that different strains adopt **distinct metabolic strategies** within the same environment. This highlights **functional diversity** and potential **differentiation** among members of the synthetic gut community.

04

Future Perspectives

4. Future Perspectives

Simulate with a More Gut-like Medium

- Replace the merged minimal medium with a realistic **gut medium**
- Assess how nutrient richness impacts strain interactions and community stability

Compare Healthy vs. Dysbiotic Communities

- Build separate communities from **commensal (non-pathogenic)** and **pathogenic** strains
- Use simulations to identify **metabolic markers** or network structures that distinguish healthy and dysbiotic states
- Evaluate metabolic robustness and inter-strain dependencies across state

Integrate SteadyCom Abundances into MICOM

Use predicted **relative abundances from SteadyCom** as input abundances in MICOM

- Assess how flux predictions change with SteadyCom abundance distribution
- Evaluate whether MICOM reproduces community-level behavior seen in SteadyCom