

# Approche hybride de modélisation explicable du métabolisme des écosystèmes microbiens

Hybrid approach for explainable metabolic modelling of microbial ecosystems'

Présenté par Maxime LECOMTE

November 21, 2023

## *Membres du jury*

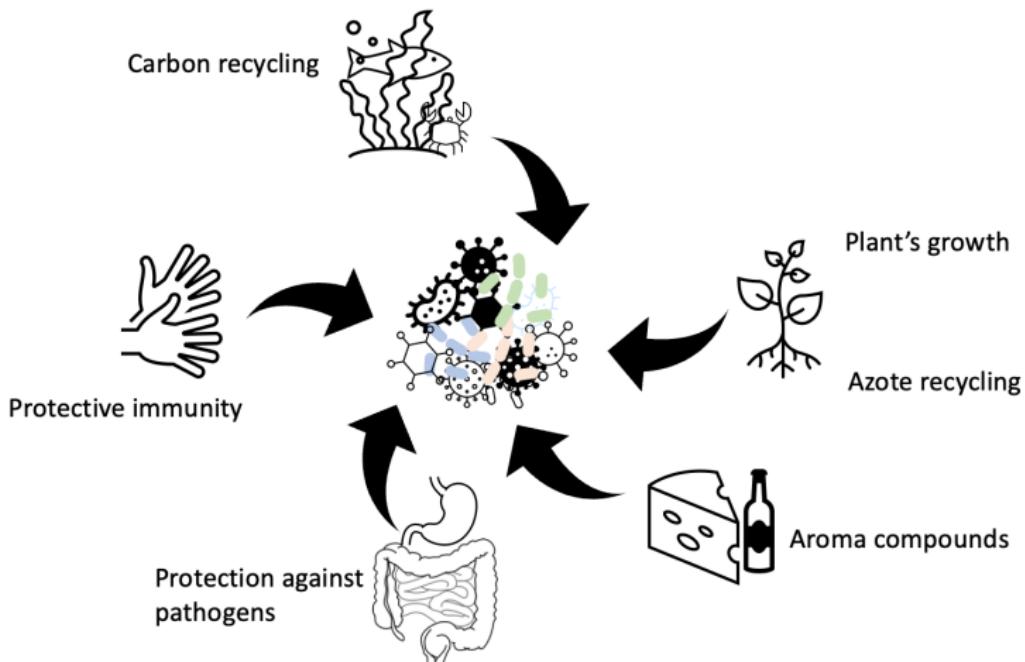
Président: SIMON Laurent

Rapportrices: BAROUKH Caroline  
COCAIGN-BOUSQUET Muriel

Examinateurs: COTTRET Ludovic  
LAROCHE Béatrice  
MARKOV Gabriel

Co-direction: David SHERMAN et Hélène FALENTIN  
Encadrement: Clémence FRIOUX

# Why the study of microorganisms is relevant ?



- High diversity of microorganisms
- Microorganisms roles specific to the environment (Royet and Plailly, 2004; Belkaid and Hand, 2014; Zhang et al., 2015; Hoorman, 2011; McSweeney and Sousa, 2000)

# Bacterial interaction are responsible of the observed roles

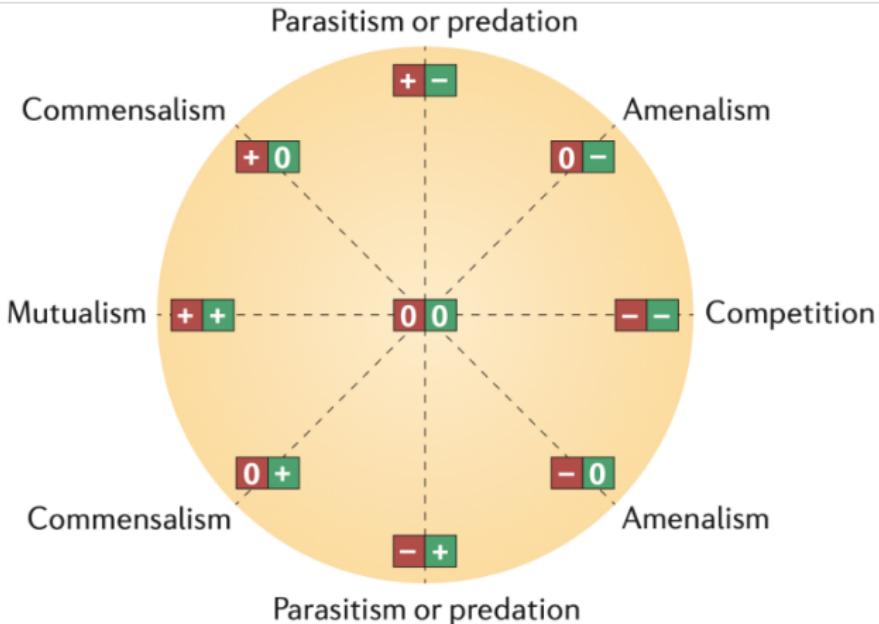


Figure 1: List of different types of bacterial interactions (Faust and Raes, 2012)

- Bacterial interactions are distinguishable within two species
- And within ecosystems composed of thousand of species ? → need of informatic

# How can we combine biological knowledge and infomatic program ?

Systems biology

## System biology

Associate an organism to a system and study the all system (Kitano, 2002)

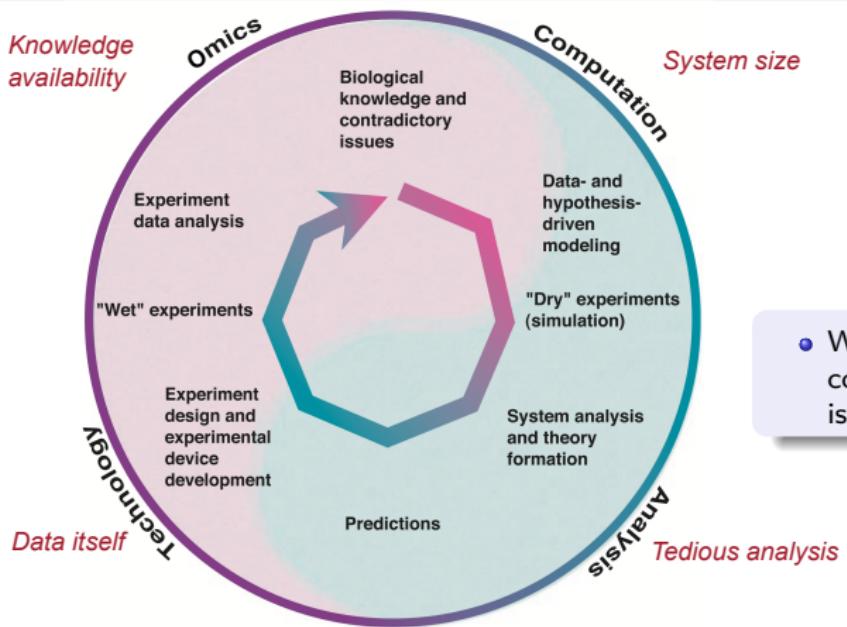
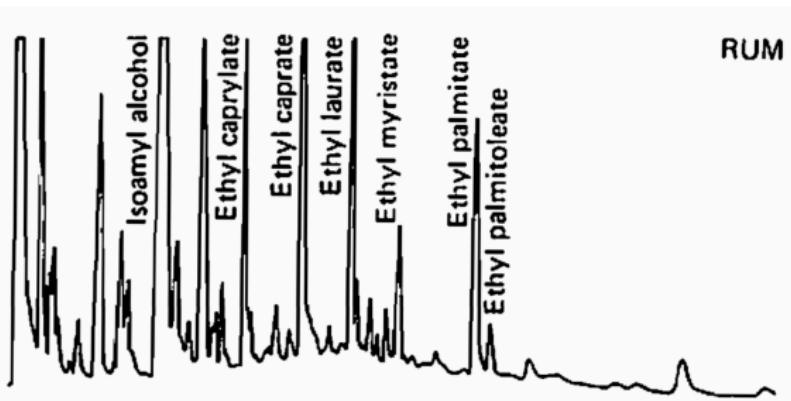


Figure 2: System biology modified from Kitano, 2002

# Metabolism as a starter pack for analysing bacterial interactions

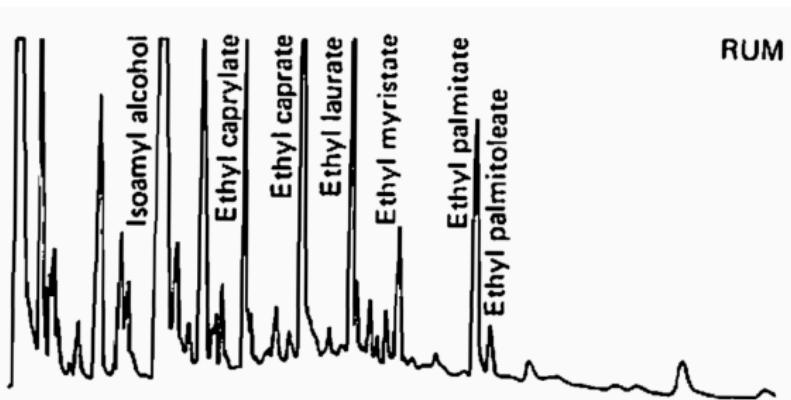
## Metabolism



**Figure 3:** Gas chromatograms of the major aroma compounds isolated from rum (from Suomalainen and Lehtonen, 1978)

# Metabolism as a starter pack for analysing bacterial interactions

## Metabolism



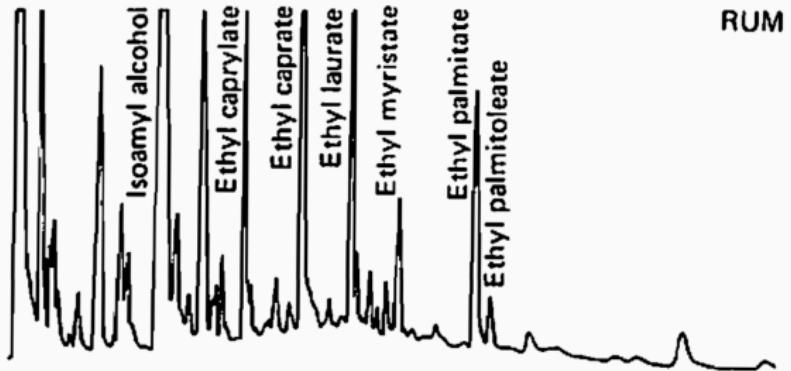
**Figure 3:** Gas chromatograms of the major aroma compounds isolated from rum (from Suomalainen and Lehtonen, 1978)

### What is metabolism ?

Set of all biochemical reactions occurring in the cell of an organism that permit the production of energy and metabolic goods. (Sánchez López de Nava A, 2023)

# What underlying mechanisms are responsible of the observed activity ?

Metabolism and Bacterial interactions



**Figure 3:** Gas chromatograms of the major aroma compounds isolated from rum (from Suomalainen and Lehtonen, 1978)

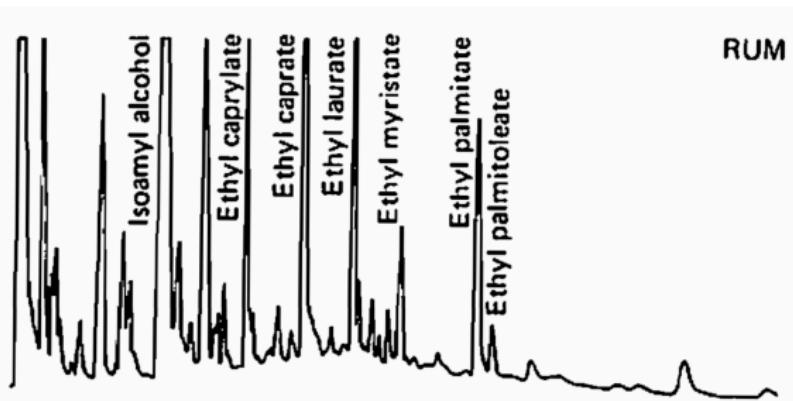
## What is metabolism ?

Set of all biochemical reactions occurring in the cell of an organism that permit the production of energy and metabolic goods. (Sánchez López de Nava A, 2023)



# What underlying mechanisms are responsible of the observed activity ?

Metabolism and Bacterial interactions



**Figure 3:** Gas chromatograms of the major aroma compounds isolated from rum (from Suomalainen and Lehtonen, 1978)

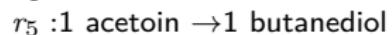
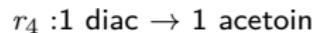
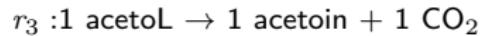
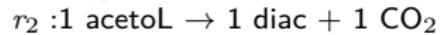
## What is metabolism ?

Set of all biochemical reactions occurring in the cell of an organism that permit the production of energy and metabolic goods. (Sánchez López de Nava A, 2023)

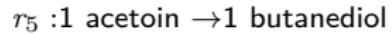
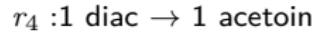
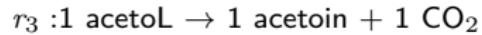
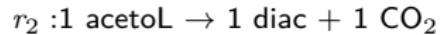
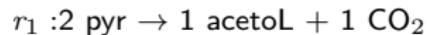


- Metabolism of an organism explain observable phenotype
- Is impacted by bacterial interactions

# How is the metabolism represented?



# How is the metabolism represented?



## Stoichiometry matrix

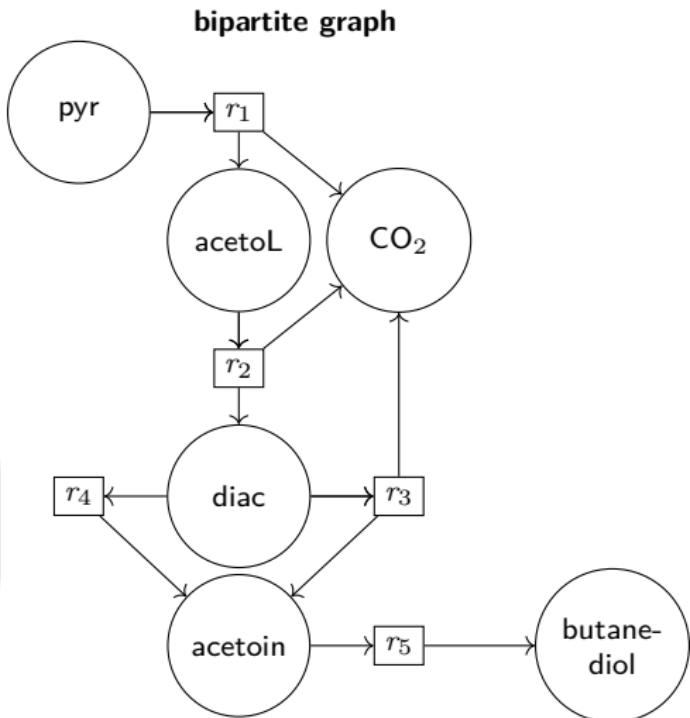
$$\begin{array}{c|ccccc} & r_1 & r_2 & r_3 & r_4 & r_5 \\ \text{pyr} & -2 & 0 & 0 & 0 & 0 \\ \text{acetoL} & 1 & -1 & -1 & 0 & 0 \\ \text{diac} & 0 & 1 & 0 & -1 & 0 \\ \text{CO}_2 & 1 & 1 & 1 & 0 & 0 \\ \text{acetoin} & 0 & 0 & 1 & 1 & -1 \\ \text{butanediol} & 0 & 0 & 0 & 0 & 1 \end{array}$$

# How is the metabolism represented?

$r_1 : 2 \text{ pyr} \rightarrow 1 \text{ acetoL} + 1 \text{ CO}_2$   
 $r_2 : 1 \text{ acetoL} \rightarrow 1 \text{ diac} + 1 \text{ CO}_2$   
 $r_3 : 1 \text{ acetoL} \rightarrow 1 \text{ acetoin} + 1 \text{ CO}_2$   
 $r_4 : 1 \text{ diac} \rightarrow 1 \text{ acetoin}$   
 $r_5 : 1 \text{ acetoin} \rightarrow 1 \text{ butanediol}$

## Stoichiometry matrix

$$\begin{array}{c}
 & r_1 & r_2 & r_3 & r_4 & r_5 \\
 \text{pyr} & \left( \begin{array}{ccccc} -2 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right) \\
 \text{acetoL} \\
 \text{diac} \\
 \text{CO}_2 \\
 \text{acetoin} \\
 \text{butanediol}
 \end{array}$$



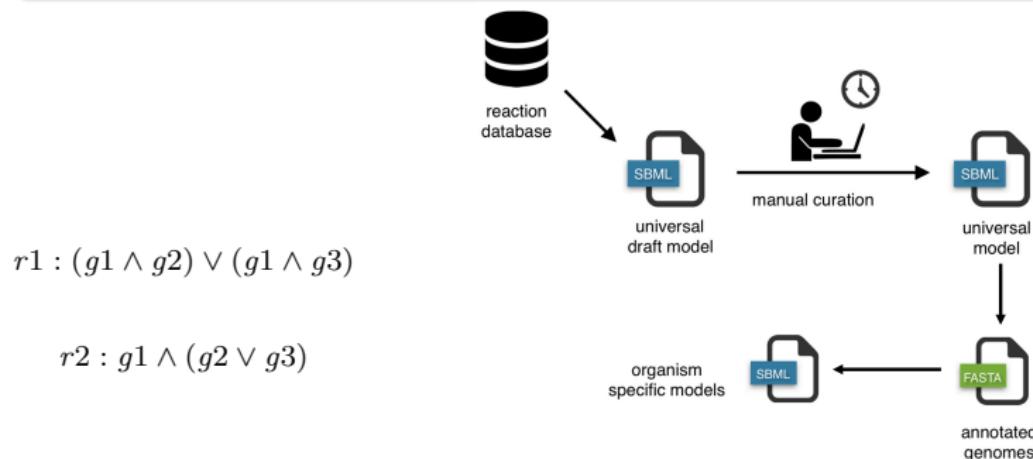
**Stoichiometry matrix** is commonly used for quantitative analysis instead of **graph**, more focused on topology analysis

# How is the metabolism reconstructed?

Genome-scale metabolic network (GEMs) reconstruction

## Genome-scale metabolic network (GSMNs)

Contain metabolic reactions predicted from the entire genomic content through gene-protein-reaction (GPR) relationships (Thiele and Palsson, 2010)



**Figure 4:** Top down genome-scale metabolic network reconstruction approach (modified from Machado et al., 2018)

- For bacteria: average of 1500 reactions, 1000 genes, 800 metabolites
- Informatic can help to resolve combinatorial problem

# Reasoning-based metabolic analysis

Definition

## Reasoning-based

Allow us to infer qualitative models from logical rules based on biological knowledge

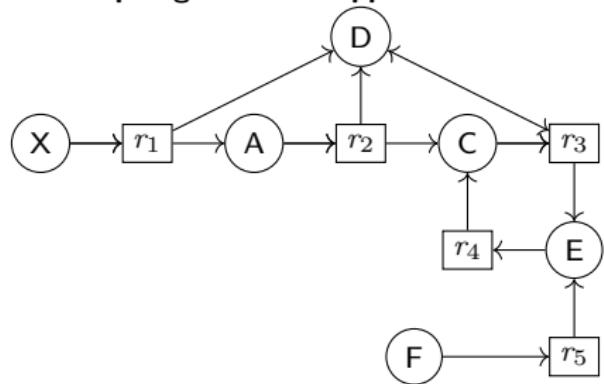
# Reasoning-based metabolic analysis

Definition

## Reasoning-based

Allow us to infer qualitative models from logical rules based on biological knowledge

### topological-based approaches



How to compute metabolic capability ?

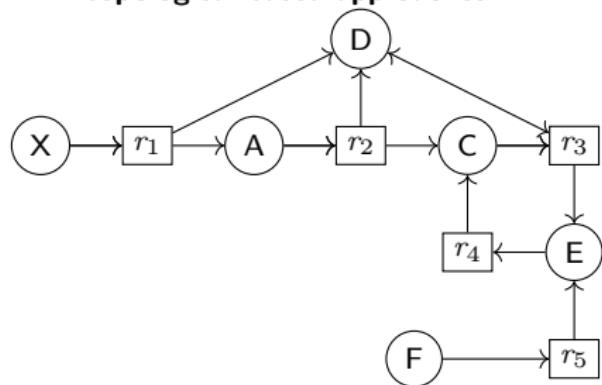
# Reasoning-based metabolic analysis

## Definition

### Reasoning-based

Allow us to infer qualitative models from logical rules based on biological knowledge

### topological-based approaches



- Producibility is initiated by the presence of nutrients,
- The products of a reactions are producible if all reactants of this reaction are themselves producible

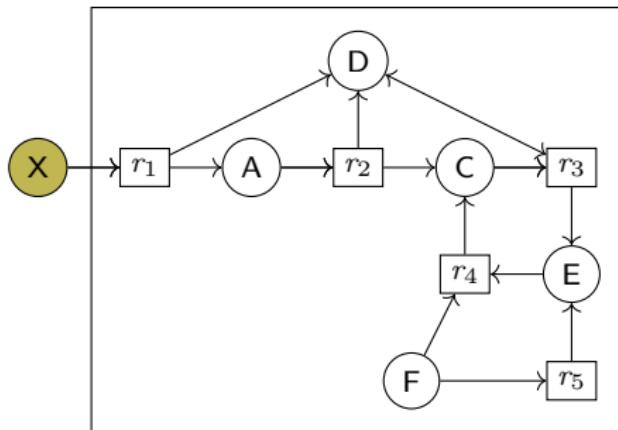
The scope, i.e. the metabolic capacity, a network is reached in 2 logical rules (Ebenhöh, Handorf, and Heinrich, 2004)

# Reasoning-based metabolic analysis

## Reasoning-based

Allow us to infer qualitative models from logical rules based on biological knowledge

## topological-based approaches



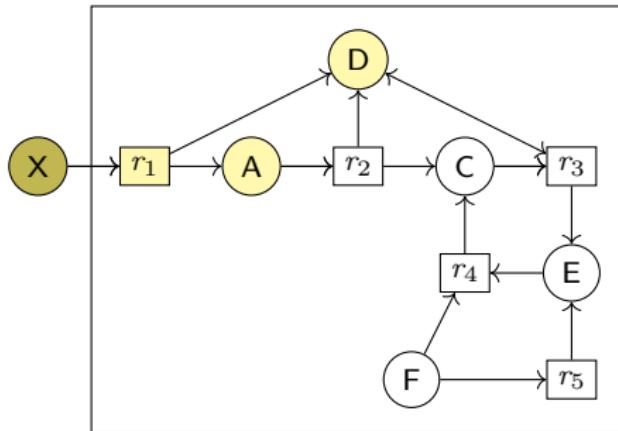
X is a seed

# Reasoning-based metabolic analysis

## Reasoning-based

Allow us to infer qualitative models from logical rules based on biological knowledge

### topological-based approaches



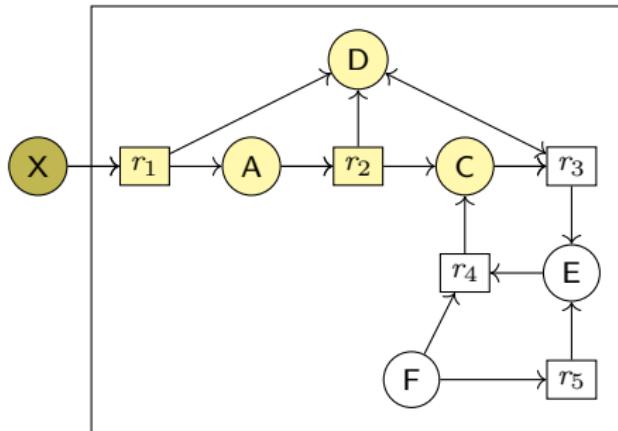
X is available;  $r_1$  is activated and A,D are producible

# Reasoning-based metabolic analysis

## Reasoning-based

Allow us to infer qualitative models from logical rules based on biological knowledge

### topological-based approaches



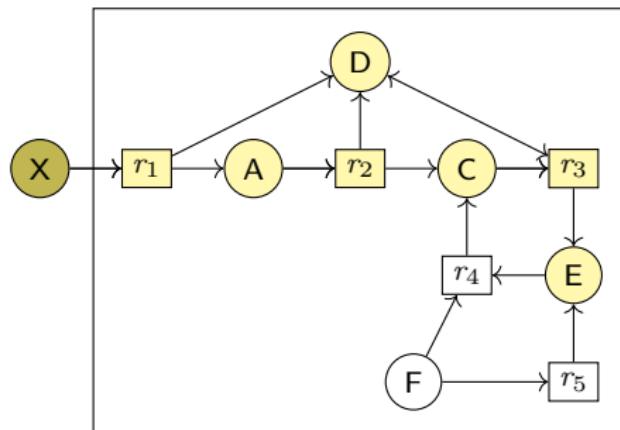
A,D are available;  $r_2$  is activated and C is producible

# Reasoning-based metabolic analysis

## Reasoning-based

Allow us to infer qualitative models from logical rules based on biological knowledge

## topological-based approaches



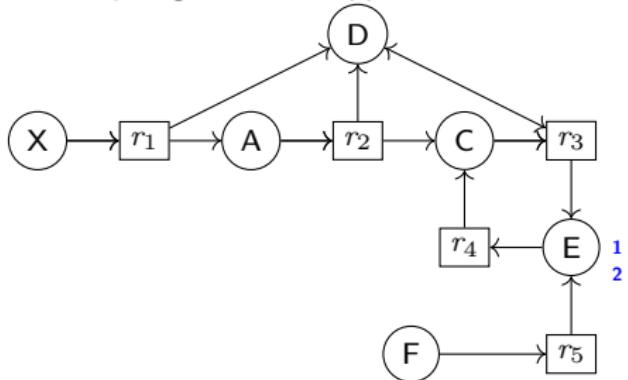
- The potential metabolic capability and topology dependant

# Logical rules implementation

## Reasoning-based

Allow us to infer qualitative models from logical rules based on biological knowledge

### topological-based approaches



- Producibility is initiated by the presence of nutrients,
- The products of a reactions are producible if all reactants of this reaction are themselves producible

```
scope(M) :- seed(M).  
scope(M) :- bacteria(B), product(M,R,B),  
reaction(R,B), scope(M2) : reactant(M2,R,  
B).
```

Logical rules are embedded using Answer Set Programming (Lifschitz2008)

# Why using Answer Set Programming

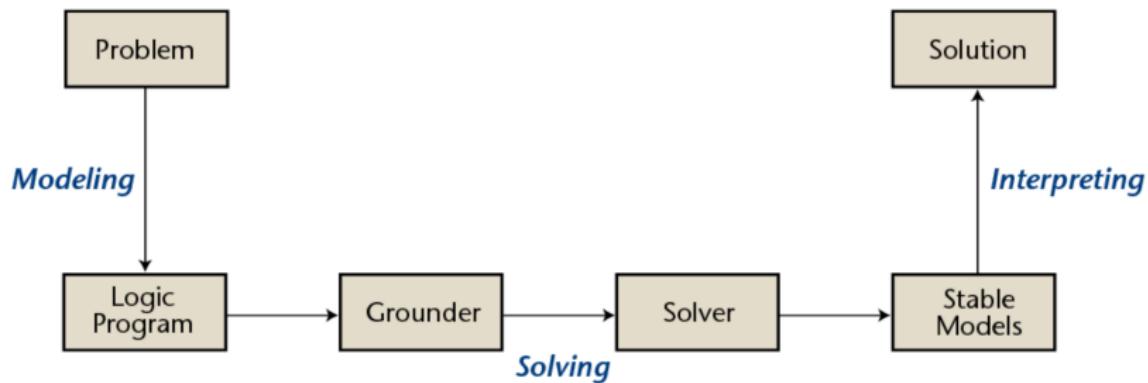


Figure 5: The workflow of Answer set programming (ASP) (Kaufmann et al., 2016)

- Close assumption
- Explainable model
- Solve combinatorial problems

# Numerical metabolic model of the metabolism

definition

## Metabolic model

From a GEM, a model metabolic has the capacity to simulate and to predict on the metabolic content

# Numerical metabolic model of the metabolism

Flux Balance Analysis (Orth, Thiele, and Palsson, 2010)

## Metabolic model

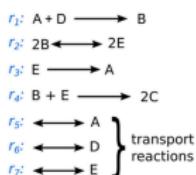
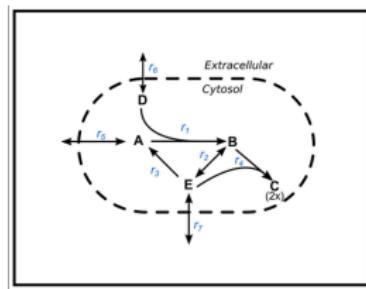
From a GEM, a model metabolic has the capacity to simulate and to predict on the metabolic content

## Constraint-based approaches

maximiser/minimiser  $f_{obj}$

$$\text{tel que } (S.v)_{int} = 0$$

$$\text{et } v_{i_{min}} \leq v_i \leq v_{i_{max}}$$



	$r_1$	$r_2$	$r_3$	$r_4$	$r_5$	$r_6$	$r_7$
A	-1	0	1	0	1	0	0
B	1	-2	0	-1	0	0	0
C	0	0	0	2	0	0	0
D	-1	0	0	0	0	1	0
E	0	2	-1	-1	0	0	1

(Stoichiometric values)

$$\bar{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} \quad (\text{Metabolic flux values})$$

Figure 7: Stoichiometry matrix representation and the flux vector  $v$

Figure 6: Example of metabolic network

# Numerical metabolic model of the metabolism

Flux Balance Analysis (Orth, Thiele, and Palsson, 2010)

## Metabolic model

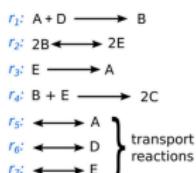
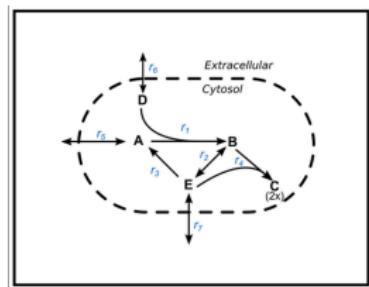
From a GEM, a model metabolic has the capacity to simulate and to predict on the metabolic content

## Constraint-based approaches

maximiser/minimiser  $f_{obj}$

$$\text{tel que } (S.v)_{int} = 0$$

$$\text{et } v_{i_{min}} \leq v_i \leq v_{i_{max}}$$



### Subject to

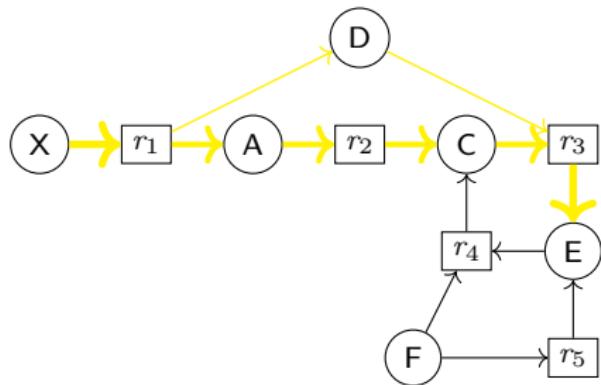
$$S\vec{v} = \vec{0} = \begin{cases} \frac{dA}{dt} = -v_1 + v_3 + v_5 & 0 \leq v_1 < \infty \\ \frac{dB}{dt} = v_1 - 2v_2 - v_4 & -\infty < v_2 < \infty \\ \frac{dC}{dt} = 2v_4 & 0 \leq v_3 < \infty \\ \frac{dD}{dt} = -v_1 + v_6 & 0 \leq v_4 \leq \infty \\ \frac{dE}{dt} = 2v_2 - v_3 - v_4 + v_7 & -\infty < v_5 < \infty \\ & 0 \leq v_6 \leq \infty \\ & 0 \leq v_7 \leq \infty \end{cases}$$

(Steady state system) (Reaction bounds)

Figure 7: Linear programming problem.

Figure 6: Example of metabolic network

# Flux application

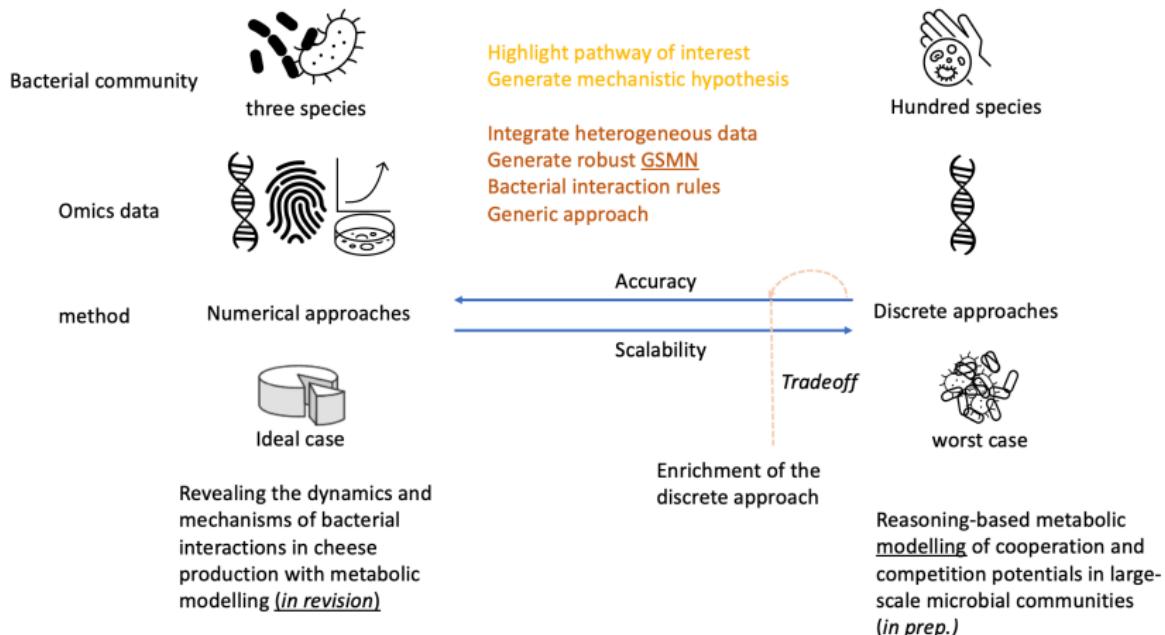


- Topology and stoichiometry of metabolic goods
- Can explain metabolic observations through reaction fluxes
- difficult to apply in larger scale

# Contributions and objective

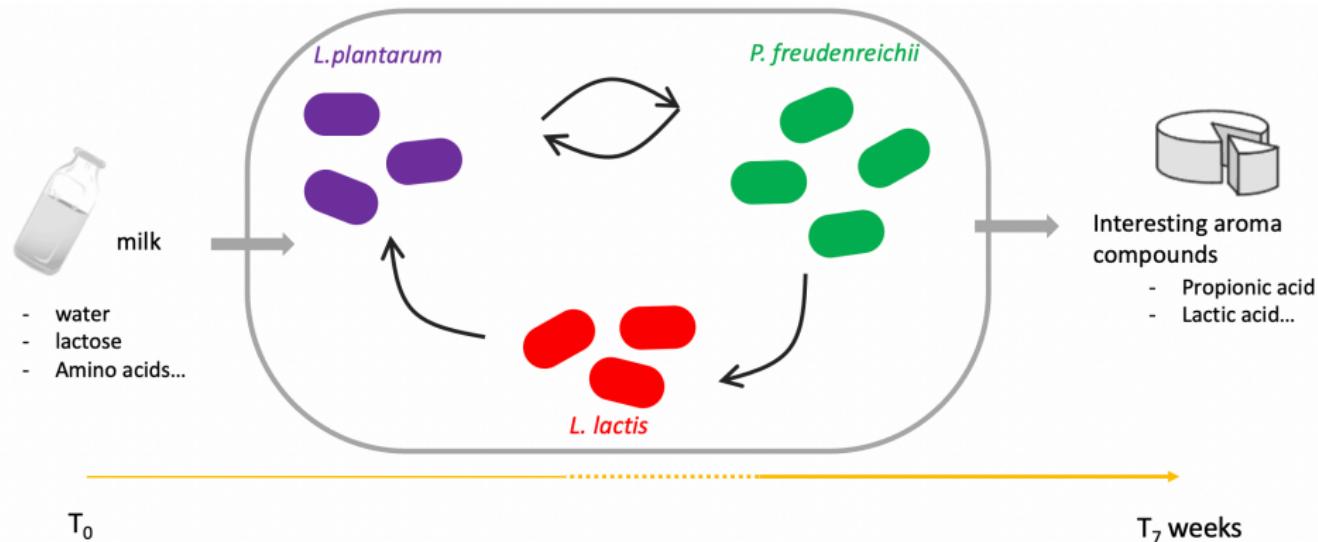
## Objective

Contribute to analyzing metabolic interactions of bacterial communities associated to two use cases: controlled and uncontrolled environment



**Figure 8:** Contributions in my thesis. In yellow and brown-red are respectively biological and methodological contributions

# Biological context: cheese bacterial fermentation



Heterogenous data are necessary for analysing bacterial fermentation

# Multi-omics strategy



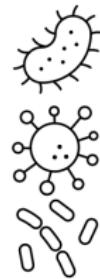
Annotated genomes



Genes expression  
(metatranscriptomics)

Acétate-HPLC-F1	Acétate-HPLC-F3
0,01	0,01
0,04	0,05
0,44	0,36
0,92	0,81
1,05	0,97
2,00	1,77
2,59	2,52

Metabolomics data



Growth and pH  
data in pure  
cultures

Dynamic and numerical model of the metabolism can integrate all the data

# Model fitting strategy

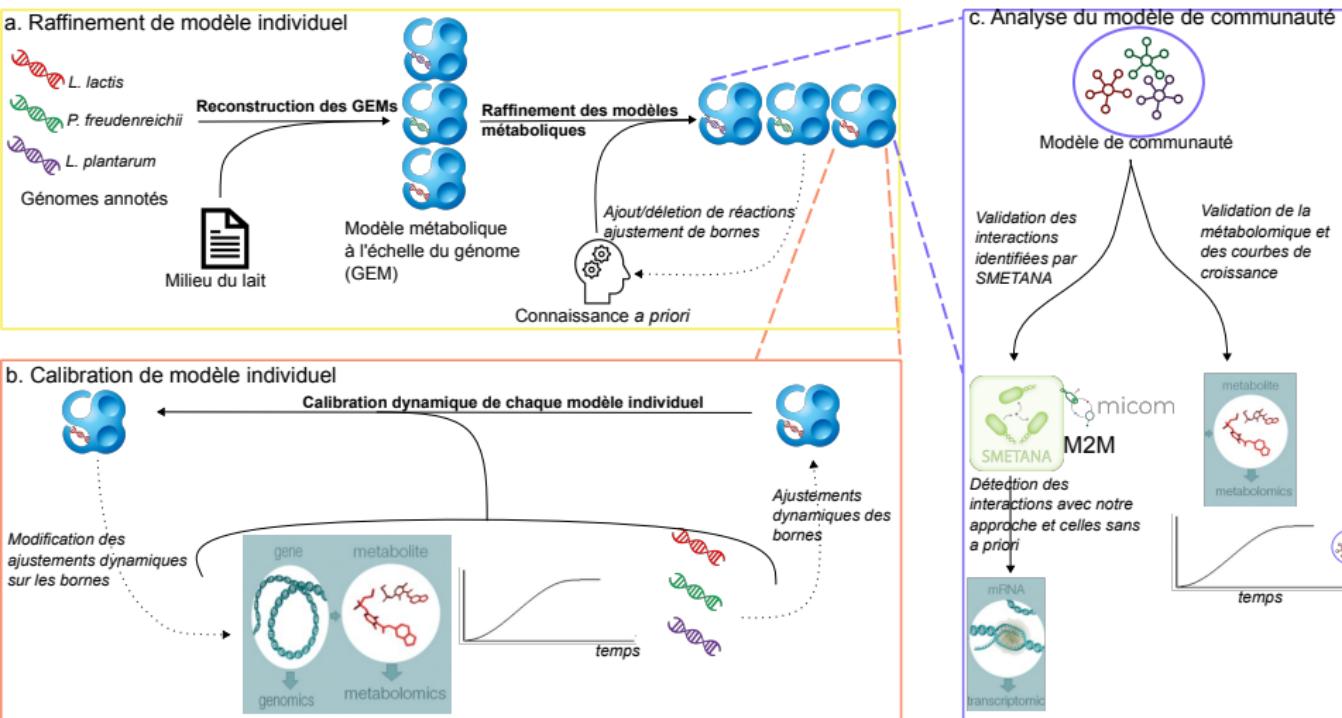


Figure 9: Numerical workflow for analysing cheese bacterial fermentation

# Refinement of metabolic networks - 1

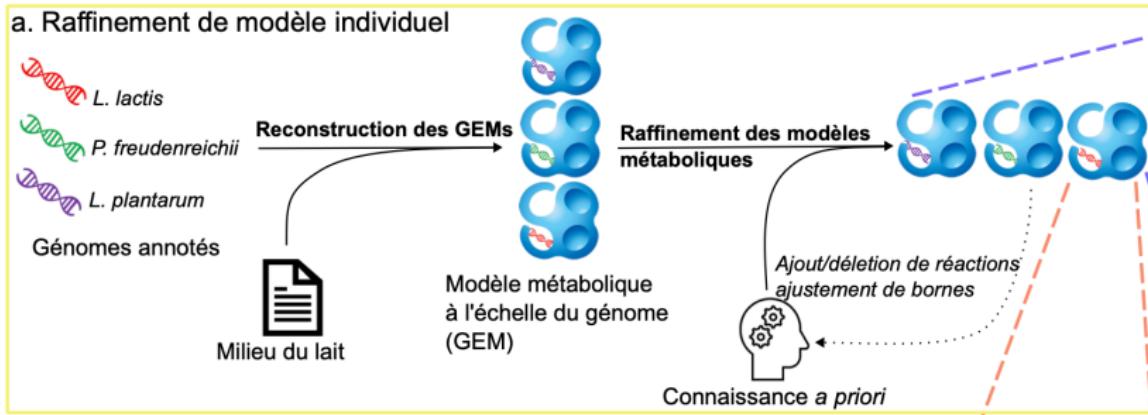
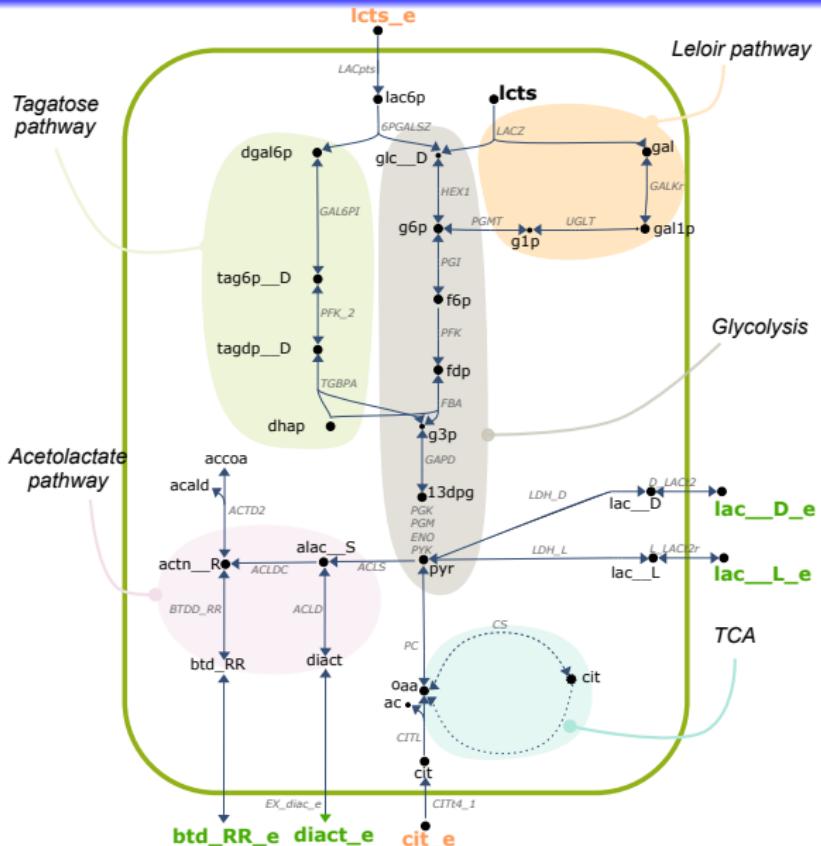


Figure 10: First part of the numerical workflow

- Adding/removing reactions
- Qualitative check of existing pathway and metabolic goods
- Tedious analysis

# Refinement of metabolic networks - 2

*L. lactis* case

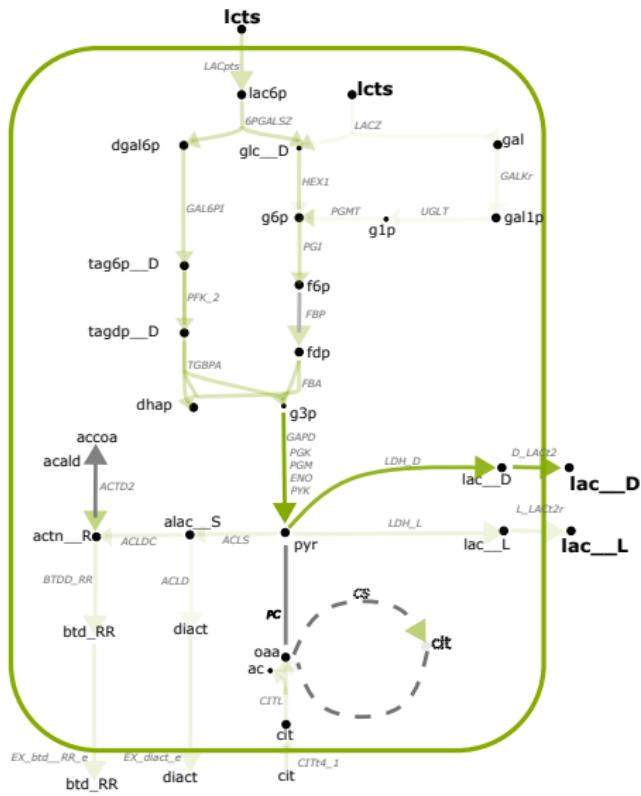


## Objectif

- Production of butanediol
- activation of acetolactate pathway

# Refinement of metabolic networks - 2

*L. lactis* case



## Objectif

- Production of butanediol
- activation of acetolactate pathway

## Refinement

- Acétoine-dehydrogenase was blocked (ACTD2)
- Modification of bounds: acétolactate decarboxylase (ACLDC) and acétolactate synthase (ACLS)

# Dynamic individual calibration

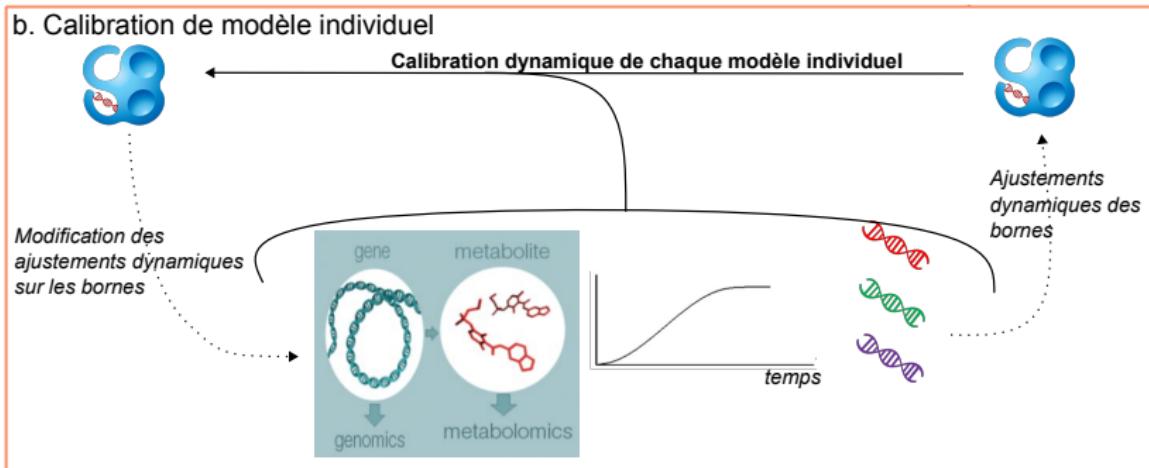


Figure 11: Second part of the numerical workflow

- Finding optimal parameters for explaining **individual** biological observations
- Quantitative check of metabolic goods and biomass density
- Tedious analysis

# Dynamic model of the metabolism

Quorum sensing and dynamic of molecules

$$\mathcal{R}_i = \lambda_i \left(1 - \frac{b_i}{\beta_i}\right), \text{ for } i \in \mathcal{B} \quad (1)$$

$$c_{min,i,j}^{ex} = \max\left(-\frac{m_{lcts_e}}{\Delta t * \sum_{i \in \mathcal{M}(lcts_e)} b_i}, -\mu_{max,lcts} * 10^{(-k_{lac} * \phi_{undiss})} - \mu_{min,lcts}\right) \quad (2)$$

# Dynamic model of the metabolism

## Dynamic regulation of pH

$$\phi_{undiss}(m_{lac\_L\_e}, m_{lac\_D\_e}) = \frac{m_{lac\_L\_e} + m_{lac\_D\_e}}{1 + 10^{c_1 * (m_{lac\_L\_e} + m_{lac\_D\_e}) + c_2}}. \quad (1)$$

# Dynamic model of the metabolism

$$F_j = \sum_{i \in \mathcal{B}} \mu_{i,j} \left( (c_{min,i}^{ex}, c_{max,i}^{ex})(b^n, m^n) \right) b_i ,$$

$$b_i^{n+1} = b_i^n + \Delta t * F_{b_i}$$

$$m_j^{n+1} = \begin{cases} m_i^n + \Delta t * F_j & \text{si } F_j > 0 \text{ (cas explicit)} \\ m_j^n / (1 - \Delta t * F_j / m_j^n) & \text{sinon (cas implicite)} \end{cases}$$

# Individual metabolic model calibration

## Goal

Make individual GSMN accurate for inferring mechanistic bacterial behavior

## LAB

$$J(b_i, pH | \theta_i, b_{i,exp}, pH_{exp}) = \left\| \frac{\log_{10}(b_i) - \log_{10}(b_{i,exp})}{\sigma_{log,i,exp}} \right\|^2 + \alpha \left\| \frac{pH - pH_{exp}}{\sigma_{pH,exp}} \right\|^2 \quad (1)$$

### *P. freudenreichii*

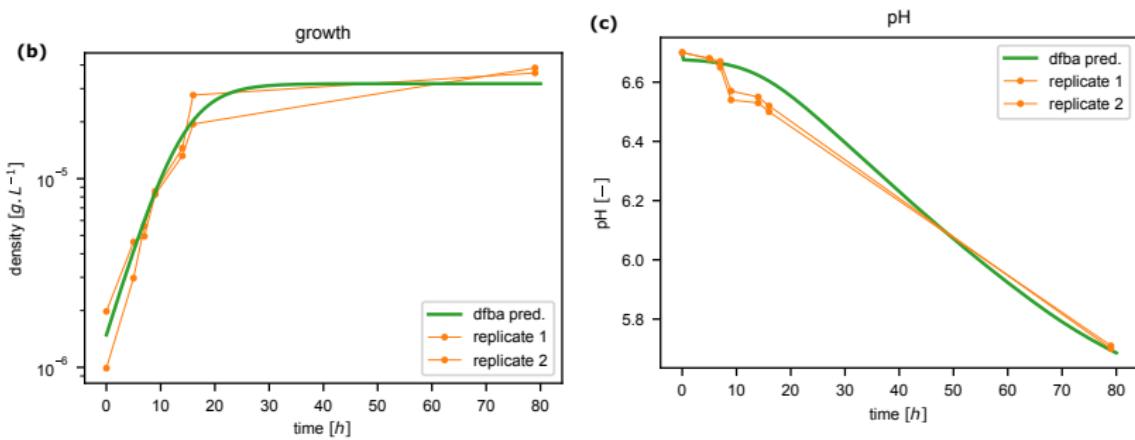
$$J(b, m | \theta_i, b_{exp}, m_{exp}) = \left\| \frac{\log_{10}(b) - \log_{10}(b_{exp})}{\sigma_{log,b,exp}} \right\|^2 + \alpha \left\| \frac{m - m_{exp}}{\sigma_{m,exp}} \right\|^2 \quad (2)$$

## Data

- Metabolomics
- Growth and pH in pure cultures

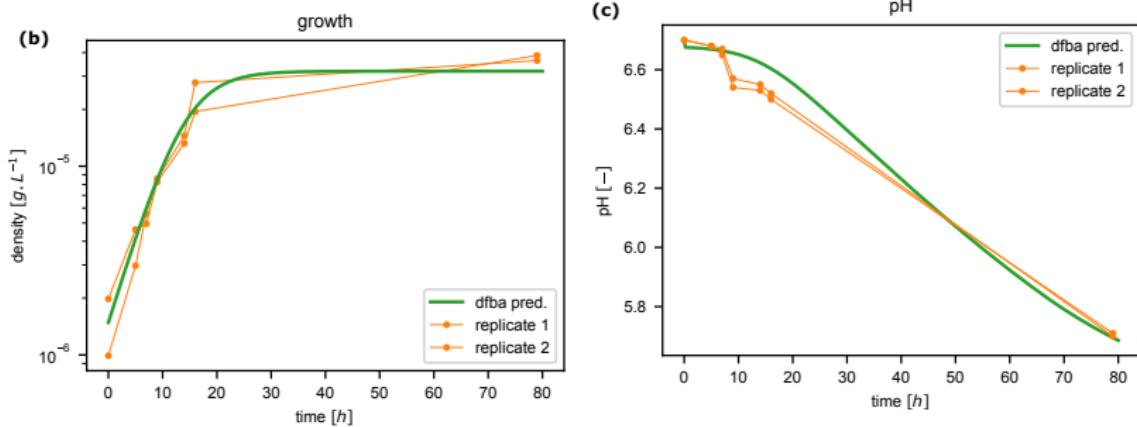
# Model validation

*L. plantarum*



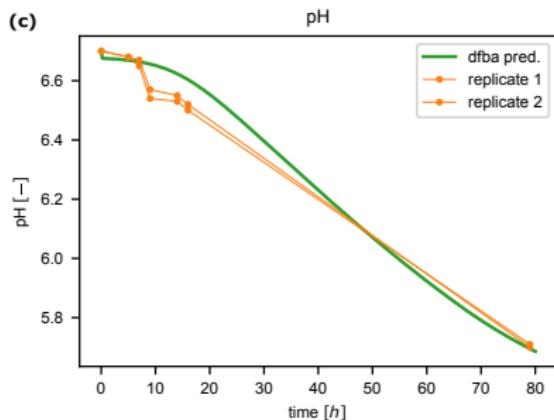
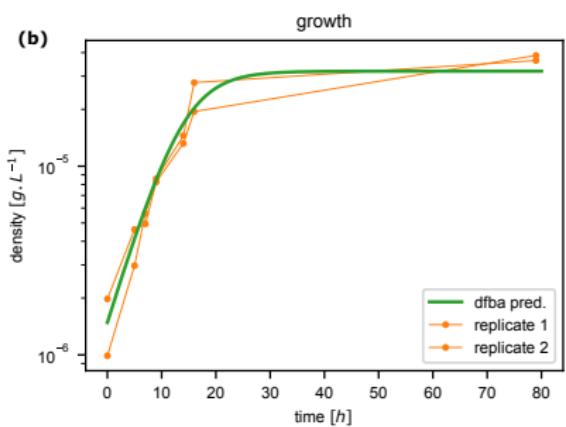
# Model validation

*L. lactis*



# Model validation

*P. freudenreichii*



Motivation  
○○○○

Background  
○○○○○○○○○○

Numerical model  
○○○○○○○○○○

References

-  Belcour, Arnaud et al. (2020). "Metage2metabo, microbiota-scale metabolic complementarity for the identification of key species". In: *eLife*. ISSN: 2050084X. DOI: 10.7554/eLife.61968.
-  Belkaid, Yasmine and Timothy W. Hand (2014). "Role of the Microbiota in Immunity and Inflammation". In: *Cell* 157.1, pp. 121–141. ISSN: 0092-8674. DOI: <https://doi.org/10.1016/j.cell.2014.03.011>. URL: <https://www.sciencedirect.com/science/article/pii/S0092867414003456>.
-  Ebenhöh, Oliver, Thomas Handorf, and Reinhart Heinrich (2004). "Structural analysis of expanding metabolic networks.". In: *Genome informatics. International Conference on Genome Informatics* 15.1, pp. 35–45. ISSN: 09199454.
-  Faust, Karoline and Jeroen Raes (2012). "Microbial interactions: From networks to models". In: *Nature Reviews Microbiology* 10.8, pp. 538–550. ISSN: 17401526. DOI: 10.1038/nrmicro2832. URL: <http://dx.doi.org/10.1038/nrmicro2832>.
-  Frioux, Clémence et al. (2018). "Scalable and exhaustive screening of metabolic functions carried out by microbial consortia". In: *Bioinformatics* 34.17, pp. i934–i943. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/bty588. URL: <https://doi.org/10.1093/bioinformatics/bty588>.
-  Hoorman, James J (2011). "The Role of Soil Bacteria". In: *The Ohio State University Extension*, pp. 1–4.
-  Kaufmann, Benjamin et al. (2016). "Grounding and Solving in Answer Set Programming". In: *AI Mag.* 37, pp. 25–32. URL: <https://api.semanticscholar.org/CorpusID:30782765>.

-  Kitano, Hiroaki (2002). "Systems biology: A brief overview". In: *Science* 295.5560, pp. 1662–1664. ISSN: 00368075. DOI: [10.1126/science.1069492](https://doi.org/10.1126/science.1069492).
-  Machado, Daniel et al. (2018). "Fast automated reconstruction of genome-scale metabolic models for microbial species and communities". In: *Nucleic Acids Research*. ISSN: 13624962. DOI: [10.1093/nar/gky537](https://doi.org/10.1093/nar/gky537).
-  McSweeney, Paul L.H. and Maria José Sousa (2000). "Biochemical pathways for the production of flavour compounds in cheeses during ripening: A review". In: *Lait* 80.3, pp. 293–324. ISSN: 00237302. DOI: [10.1051/lait:2000127](https://doi.org/10.1051/lait:2000127).
-  Orth, Jeffrey D., Ines Thiele, and Bernhard O. Palsson (2010). *What is flux balance analysis?* DOI: [10.1038/nbt.1614](https://doi.org/10.1038/nbt.1614).
-  Royet, Jean-Pierre and Jane Plailly (Oct. 2004). "Lateralization of Olfactory Processes". In: *Chemical Senses* 29.8, pp. 731–745. ISSN: 0379-864X. DOI: [10.1093/chemse/bjh067](https://doi.org/10.1093/chemse/bjh067). eprint: <https://academic.oup.com/chemse/article-pdf/29/8/731/930410/bjh067.pdf>. URL: <https://doi.org/10.1093/chemse/bjh067>.
-  Suomalainen, By Heikki and Matti Lehtonen (1978). "JJu It". In: 85, pp. 149–156.
-  Sánchez López de Nava A, Raja A (2023). *Physiology, Metabolism*. URL: [www.ncbi.nlm.nih.gov/books/NBK546690/](http://www.ncbi.nlm.nih.gov/books/NBK546690/).
-  Thiele, Ines and Bernhard O Palsson (2010). "A Protocol for Generating a High-Quality Genome-Scale Metabolic Reconstruction". In: *Nature Protocols* 5.1, pp. 93–121. ISSN: 1754-2189. DOI: [10.1038/nprot.2009.203](https://doi.org/10.1038/nprot.2009.203).



Zhang, Yu Jie et al. (2015). "Impacts of gut bacteria on human health and diseases". In: *International Journal of Molecular Sciences* 16.4, pp. 7493–7519.  
ISSN: 14220067. DOI: 10.3390/ijms16047493.