THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Modeling human gut microbiota: from steady states to dynamic systems

Hao Luo

CHALMERS
UNIVERSITY OF TECHNOLOGY

Systems and Synthetic Biology
Department of Biology and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2022

Modeling human gut microbiota: from steady states to dynamic systems
Hao Luo

ISBN 978-91-7905-709-1

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5175
ISSN 0346-718X

Division of Systems and Synthetic Biology
Department of Biology and Biological Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover: Waves in the gut.
Sometimes turbulent, sometimes calm.

Printed by Chalmers Reproservice
Gothenburg, Sweden 2022

Modeling human gut microbiota: from steady states to dynamic systems

Hao Luo
Department of Biology and Biological Engineering
Chalmers University of Technology

Abstract

Human gut microbes are an essential part of human sub-microscopic systems and involved in many critical biological processes such as Type 2 diabetes (T2D) and osteoporosis. However, the underlying mechanisms are unclear. Several mathematical modeling approaches, such as genome-scale metabolic models (GEMs) and ordinary differential equation (ODE) based models, have been used to simulate the dynamics of human gut microbiota. This thesis aims to explore, simulate, and predict the behavior of gut microbial ecosystems and the relationships between gut microbes and humans by modeling.

The importance of the gut microbiome for bone metabolism and T2D has been demonstrated in mice and human cohorts. We first reconstructed a GEM for *Limosilactobacillus reuteri* ATCC PTA 6475, which is a probiotic that significantly reduces bone loss in older women with lower bone mineral density. To investigate the associations between T2D and the gut microbiota, GEMs for 827 gut microbial species and 1,779 community-level GEMs for T2D cohorts have also been constructed. With these GEMs, we investigated metabolic potentials such as short-chain fatty acids, amino acids, and vitamins that play vital roles in the host metabolism regulation. Furthermore, the integration of the models with machine learning method provides potential insights into the possible roles of gut microbiota in T2D.

Cybernetic models, which simulate metabolic rates by integrating the control of enzyme synthesis and enzyme activities, have been applied to explore the dynamic behaviors of small-size metabolic networks. However, only a few studies have applied cybernetic theory to the microbial community so far. The remaining part of this thesis focuses on the use of cybernetic models to explore human gut microbiota's interactions and population dynamics. Considering the high computing burden of the current cybernetic modeling approach for processing the full-size GEMs, we have developed a computing-efficient strategy for model reconstruction and simulation to reveal the metabolic dynamics of human gut microbiota.

In this thesis, we explore the human gut microbiota from single *L. reuteri* species to microbial gut communities, from simple steady state systems by GEMs to complex dynamic systems by cybernetic model.

Keywords: metabolic modeling, gut microbiota, genome-scale metabolic model, cybernetic model, dynamics.

List of Publications

This thesis is based on the work in the following publications and manuscript:

Paper I: Genome-scale Insights into the metabolic versatility of *Limosilactobacillus reuteri*
Hao Luo, Peishun Li, Hao Wang, Stefan Roos, Boyang Ji, Jens Nielsen. *BMC biotechnology* 21, 46 (2021).

Paper II: One-year supplementation with *Lactobacillus reuteri* ATCC PTA 6475 counteracts a degradation of gut microbiota in older women with low bone mineral density
Peishun Li*, Boyang Ji*, Hao Luo, Daniel Sundh, Mattias Lorentzon and Jens Nielsen. npj Biofilms and Microbiomes (2022), *In press*

Paper III: Modeling the metabolic dynamics at the genome-scale by optimized yield analysis
Hao Luo, Peishun Li, Boyang Ji, Jens Nielsen, *Manuscript submitted*

Paper IV: Metagenomic analysis of type 2 diabetes datasets identifies cross-cohort microbial and metabolic signatures
Peishun Li*, Hao Luo*, Boyang Ji, Jens Nielsen. *Manuscript*

Paper V: The metabolic network inference framework for shotgun metagenomics
Hao Luo*, Boyang Ji*, Jens Nielsen, *Draft Manuscript*

Additional papers and manuscripts not included in this thesis:

Gut microbiome and machine learning potential for personalized medicine
Peishun Li, Hao Luo, Boyang Ji, Jens Nielsen. *Manuscript submitted*

Bayesian genome scale modelling identifies thermal determinants of yeast metabolism
Gang Li, Yating Hu*, Jan Zrimec*, Hao Luo, Hao Wang, Aleksej Zelezniak, Boyang Ji, Jens Nielsen. *Nature Communications*. Vol. 12 (1) (2021)

Omics-based growth characterization and construction of metabolic network for *Halomonas bluephagenesis*
Jian-Wen Ye*, Lizhan Zhang*, Gang Li*, Helen Park, Hao Luo, Yina Lin, Fuqing Wu, Wuzhe Huang, Nigel S. Scrutton, Jens Nielsen, Guo-Qiang Chen. *Manuscript submitted*

Flux regulation through glycolysis and respiration is balanced by inositol pyrophosphates
Qi N, Li L, Ji X, Pereira R, Chen Y, Yin S, Li C, Wan X, Luo H, Zhang Y, Dong G, Zhang Y, Shi S, Larsson C, Chen Y, Tan T, Liu Z, Nielsen J. *Manuscript submitted*

* Contributed equally

Contribution summary

Paper I. Co-designed the study, reconstructed models, performed simulations, and wrote the manuscript.

Paper II. Assisted in data analysis and manuscript preparation.

Paper III. Co-designed the study, developed the algorithm, constructed models, performed simulations, and wrote the manuscript.

Paper IV. Constructed models, performed simulations, assisted in data analysis and manuscript preparation.

Paper V. Developed the toolbox, performed infant data analysis, wrote the manuscript.

Preface

This dissertation serves as partial fulfillment of the requirements to obtain the degree of Doctor of Philosophy at the Department of Biology and Biological Engineering at Chalmers University of Technology. The PhD research was carried out between September 2018 and October 2022 at the division of Systems and Synthetic Biology (Sysbio) under the supervision of Jens Nielsen. The project was co-supervised by Boyang Ji and Aleksej Zelezniak and examined by Verena Siewers. The project was mainly funded by the BioGaia company and the Novo Nordisk Foundation.

Hao Luo
October 2022

Contents

Abbreviations

Acronyms

GI	Gastrointestinal
T2D	Type 2 diabetes
GEM	Genome-scale metabolic model
ODE	Ordinary differential equation
SCFAs	Short chain fatty acids
RNA	Ribonucleic acid
GPR	Gene-protein-reaction
KEGG	Kyoto encyclopedia of genes and genomes
KO	KEGG orthology
FBA	Flux balance analysis
pFBA	Parsimonious FBA
EFM	Elementary flux mode
EFV	Elementary flux vector
HCM	Hybrid cybernetic models
L-HCM	Lumped hybrid cybernetic model
GR	Good response
PR	Poor response
BMD	Bone mineral density
vBMD	Volumetric bone mineral density
usCRP	Ultrasensitive c-reactive protein
MYA	Metabolic yield analysis
NGT	Normal glucose tolerance
Pre-D	Prediabetes
RF	Random forest
ML	Machine learning
LightGBM	Light gradient boosting machine
XGBoost	Extreme gradient boosting decision trees
AMN	Analyzer for Metabolic Networks
PCA	Principal component analysis
NMF	Nonnegative matrix factorization
BBH	Bidirectional best hit
GMM	Gut microbiota medium

Symbols

c	Biomass concentration
$\mathbf{e}$	The vector of enzyme level
$\mathbf{m}$	The vector of metabolites
$\mathbf{h}$	The vector of conversion factors required to express each metabolite concentration on a weight fraction
k	Reaction rate constant
K	Michaelis–Menten constant
p	Return on investment
r	Fully regulated specific reaction rate
$\mathbf{r}$	The vector of fully regulated specific reaction rates
$\mathbf{rM}$	The vector of fully regulated (uptake) fluxes through elementary modes
$\mathbf{S}$	Stoichiometric matrix of metabolites
$\mathbf{Sin}$	Stoichiometric matrix for intracellular metabolites
u	Cybernetic variable-controlling inductive syntheses of enzymes
v	Cybernetic variable-controlling enzyme activities; vertices of a convex hull
$\mathbf{x}$	The vector of metabolite concentrations (including both intracellular and extracellular metabolites); the vector of nonnegative fluxes obtained by decomposing reversible fluxes into irreversible pairs
α	Constitutive rate of enzyme synthesis
β	Rate constant for enzyme degradation
μ	Specific growth rate
σ	Parameter that scales the cost associated with resource investment
$\mathbf{diag(\cdot)}$	Operator forming a diagonalization from its argument
$\mathbf{Z}$	Elementary mode matrix

"All models are wrong, but some are useful."

-- George E. P. Box

1. Background

1.1. The human gut microbiota

1.1.1 Hello, human gut microbiota

Do you believe that half of the cells in your body do not belong to you?

In 1676, Antonie van Leeuwenhoek observed microorganisms from a microscope of his own design (Cocquyt et al., 2021). Since then, humans have been observing and exploring microorganisms (Lane, 2015). After years of study, it was realized that we humans are inseparably related to microorganisms and that our human body is the host of trillions of microorganisms (Bengmark, 1998; Thursby & Juge, 2017). These microbes colonize almost every nook and cranny of our body, such as the skin, intestine, mouth, and nasal cavity. Especially the gastrointestinal (GI) tract, which is one of the largest interfaces between the body and the external environment (about 250–400 m^2), is densely populated with a vast number of microbes and a wide variety of species (Luckey, 1972; Neish, 2009; Thursby & Juge, 2017). This collection of microbes is known as the 'human gut microbiota'. How many microbes are there in your body? Some studies estimate that there are approximately 10^{14} bacterial cells, and the ratio of microbial cells and human cells is close to 1:1 (Sender et al., 2016). This is why half of the cells in your body may not belong to you, they are microbes. One more exciting topic than cell counts is the study of gut microbiota species and composition. The human microbiota is a complex and dynamic population of microorganisms that consists of probably more than 2,000 different species of bacteria, viruses, fungi, and other microbes (Gill et al., 2006). Due to the advent of culture-independent approaches (Moore & Holdeman, 1974) such as high-throughput sequencing, i.e. 16S ribosomal RNA (rRNA) or shotgun metagenomic sequencing, methods (Mizrahi-Man et al., 2013; Poretsky et al., 2014), more and more species are identified (Gill et al., 2006). Some studies use metagenomic analysis classified them mainly into 12 different phyla, most belonging to Proteobacteria, Firmicutes, Actinobacteria and Bacteroidetes. A recent study identified 9,879,896 genes from 1,018 samples, 100 times more genes than that in the human genome (Hugon et al., 2015; J. Li et al., 2014). The composition and structure of gut microbiota have significant interindividual variations and multiple environmental factors throughout the life span can affect the gut microbiota, as shown in **Figure 1.1**.

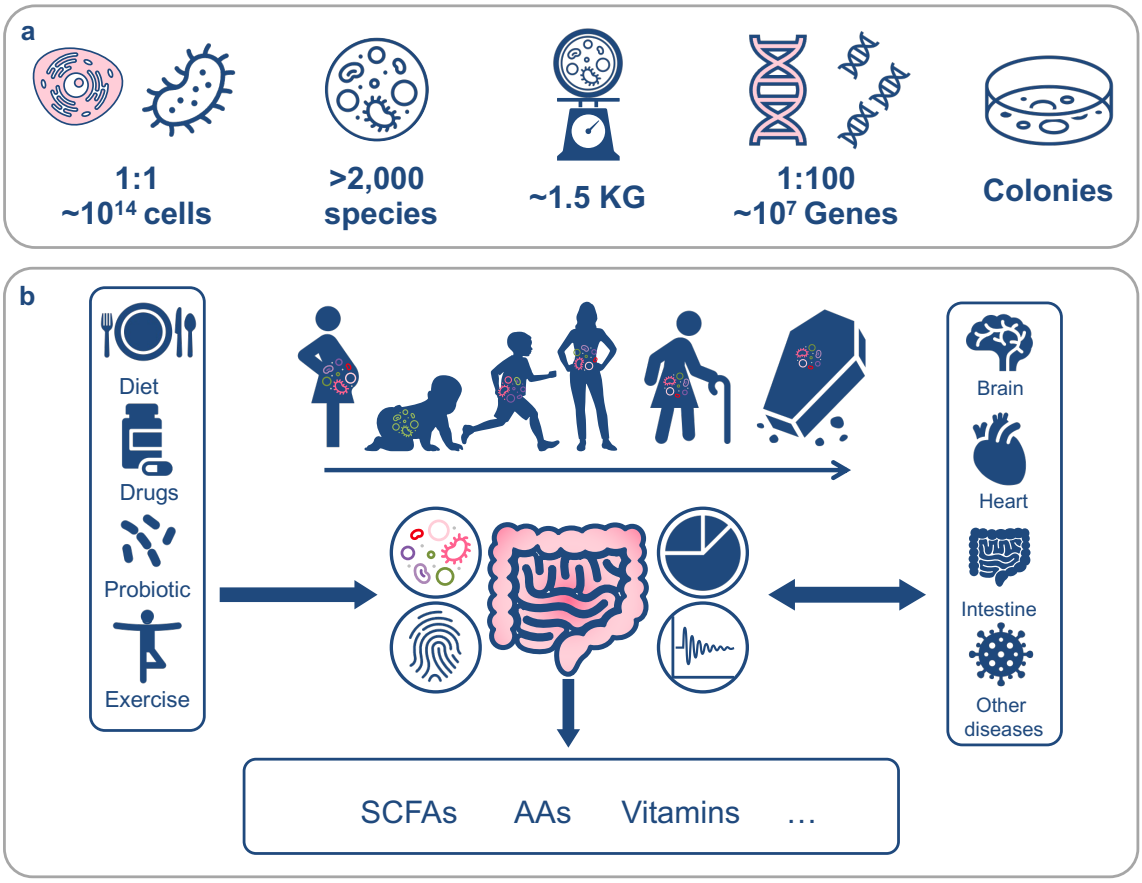


Figure 1.1| The human gut microbiota. a) There are approximately 10^{14} bacterial cells in our gastrointestinal (GI) tract, and the ratio of human cells and microbial cells is close to 1:1. There are over 2,000 species identified by metagenomic analysis. The average weight of adult human gut microbiota is approximately 1.5 kg. About 10^7 genes are identified, 100 times that in the human genome. Most species cannot be isolated in the laboratory. **b)** The microbiota is generally believed to begin at birth and stay with us for a lifetime. The original microbiota comes from our mother. Environmental factors and life events can shape the microbiota composition, such as daily diet, antibiotic or probiotic treatment, and illness. The microbiota offers many benefits to the host, such as producing short-chain fatty acids (SCFAs), branched-chain amino acids and vitamins. Microbiota composition is also associated with many diseases, such as type 2 diabetes (T2D) and osteoporosis.

It seems like the “army of microbes” has “taken over” our GI tract. In fact, the symbiotic relationship benefits both microbes and humans as long as the body is in a healthy state. Additionally, the microbiota provides many benefits, plays a critical role for the host, and has become an indispensable 'organ' for us (Bäckhed, 2011). For example, some physiological functions, such as harvesting energy, protecting against pathogens, and regulating host immunity (Gensollen et al., 2016). The status of the gut microbiota is also associated with many diseases. For example, in a state of dysbiosis of gut microbiota, the risk of intestinal disease increases (Schroeder & Bäckhed, 2016).

1.1.2 Where does the gut microbiota come from

When we are born, all we have come from our mothers, including our gut microbiota. A previous study reported that human genetics determines to some extent the composition of

the gut microbiota and that some microbial species from the phylum *Firmicutes* and the genus *Wolbachia* are heritable (Aagaard et al., 2014; Rodríguez et al., 2015). These species can be detected in the womb tissues of the mother. After birth, the GI tract is rapidly colonized, and the mode of delivery and feeding methods can affect the microbiota composition. It has been reported that babies delivered by natural birth have more *Lactobacillus* and *Prevotella* species (Aagaard et al., 2012; Avershina et al., 2014), and those delivered by cesarean section have more species of *Staphylococcus* and *Corynebacterium* (Jakobsson et al., 2014; Salminen et al., 2004). Compared to babies born by cesarean section, babies born naturally have a more similar microbiota to their mothers (Rodríguez et al., 2015). Human milk-fed infants have more abundances of *Bifidobacterium longum* and *Bacteroides* species. In the early stages of development, microbiota diversity is generally low and dominated by two phyla, *Actinobacteria* and *Proteobacteria* (Bäckhed, 2011; Rodríguez et al., 2015). During the first year of life, microbial diversity increases, and the microbiota composition tends to resemble the unique microbial spectrum of adults.

The composition of the intestinal microbiota is relatively stable during the adulthood, but it can still change over time (Rodríguez et al., 2015). Individuals older than 65 years have an increase in the abundance of *Bacillus*-like phylum and *Clostridium* class IV, and a more prevalent XIVa group compared to younger subjects (Biagi et al., 2010; Claesson et al., 2011). There is an important relationship between diversity and living arrangements, the microbiota of the centenarian population is significantly less diverse, and the microbiota of the elderly has a reduced ability to carry out metabolic processes such as SCFAs production and amylolysis (Biagi et al., 2013; Woodmansey et al., 2004). It has been hypothesized that the reduction in SCFAs may harbor inflammatory aging in the gut of the elderly processes (Biagi et al., 2013).

One of the main factors shaping the gut microbiota composition is diet. Some studies estimate that over 60 tons of food pass through the GI during a human lifetime (Bengmark, 1998). Different food contacts GI environment and influences community structure. The gut bacteria mainly rely on carbohydrates which have escaped the digestive enzymes and reached the colon (Donaldson et al., 2015). Other nutrients may promote the growth of different microbial species. For example, increased fiber availability promotes the diversity of gut microbes and the production of short-chain fatty acids (Woodmansey et al., 2004). Vegetarians have a different gut microbiota composition to process fiber (David et al., 2013; Zoetendal et al., 2012).

It has also been shown that antibiotic use can lead to dysbiosis of the microbiota. The direct exposure of gut microbes to antibiotics can kill large numbers of microbes (Jernberg et al., 2007; Maurice et al., 2013). The gut homeostasis can be disturbed to dysbiosis. Drug therapy may affect the gut microbiota and associate with human disease (Browne et al., 2016; Jernberg et al., 2007).

Gut microbiota is sensitive because it can be influenced by multiple environmental factors. The composition will be stable after a long period of environmental influence and adapt to changes to maintain balance (Gustafsson et al., 2012; Plovier et al., 2016; Tailford et al., 2015). Some studies show that a person's gut microbiota is more similar to members from the same family than to members from different families. Previous studies have shown that smoking may lead to an increase in *Prevotella* (Biedermann et al., 2013). And stress leads to altered gut microbial composition and reduced levels of *lactobacilli* (Tyakht et al., 2013).

1.1.3 Metabolic functions of the human microbiota

The gut microbiota is complementary to the human digestive system and provides many metabolic functions. One of the most critical functions is to digest food components that are otherwise indigestible. They can provide enzymes to help humans break down plant cellulose, such as complex polysaccharides and polyphenols. Gut microbes help humans absorb energy nutrients from food, and also synthesize some positive metabolites such as SCFAs, vitamins, and amino acids (Biagi et al., 2013; Louis et al., 2014; Macfarlane & Macfarlane, 2003; Morrison & Preston, 2016; Woodmansey et al., 2004).

Short-chain fatty acids (SCFAs): many species in gut microbiota express carbohydrate-active enzymes to ferment complex carbohydrates generating SCFAs. Include acetate, propionate, and butyrate and the proportion is close to 3:1:1 (Louis et al., 2014). SCFAs play vital roles in the regulation of host metabolism and are produced by different species from different metabolic pathways. Acetate is the most abundant SCFA and produced by most gut anaerobes; propionate is mainly produced by *Bacteroidetes* (Louis & Flint, 2017); butyrate is produced by *Firmicutes*. Propionate formation is mainly synthesized via the succinate or propanediol pathway and butyrate is synthesized from carbohydrates via glycolysis and acetoacetyl-CoA (Louis & Flint, 2017; Macfarlane & Macfarlane, 2003; Morrison & Preston, 2016; Musso et al., 2010).

Vitamins: The gut microbiota is necessary for vitamin synthesis (LeBlanc et al., 2013; Martens et al., 2002; Pompei et al., 2007). The host is incapable of synthesizing some vitamins like the majority of vitamin B and vitamin K. The vitamin B group is synthesized by lactic acid bacteria and different B vitamins can also be synthesized by different bacteria (Martens et al., 2002; Pompei et al., 2007). For example, folate is synthesized by certain species of bifidobacteria and cobalamin is synthesized by *L. reuteri* and propionibacteria (Palau-Rodriguez et al., 2015; Pompei et al., 2007).

Amino acids: The gut microbiota also contributes to producing amino acids from dietary nitrogen sources and carbohydrates (Pompei et al., 2007). It is a supplementary source of amino acids. Even though most amino acids for humans come from food, the gut microbiota produced amino acids are important for maintaining the community stable (Carbonero et al., 2012; Ridlon et al., 2014).

1.1.4 Limosilactobacillus reuteri

Limosilactobacillus reuteri, previously known as *Lactobacillus reuteri*, is a well-studied gut microbial species and has many applications (Su et al., 2012; Zheng et al., 2020):

As a **probiotic**, some specific strains of *L. reuteri* are widely used as probiotic supplements. Probiotics are "live microorganisms that, when administered in adequate amounts, confer a health benefit on the host" (Dore et al., 2019; Nelson et al., 2010). Many strains of *L. reuteri* are applied in disease treatment and food products. For example, some strains of *Lactobacillus* and *Bifidobacterium* affect host health and metabolism. Several studies have suggested that the health benefits of probiotics include positive effects on the immune system, prevention of infections, and treatment of antibiotic diarrhea (Britton et al., 2014; Nilsson et al., 2018; Savino et al., 2007). However, more details about health benefits remain unclear. Multiple factors shape the probiotic dynamics and their survival and mechanisms in the intestine. With the shown beneficial properties of *Lactobacillus/Limosilactobacillus* strains, *L. reuteri* proved to have positive effects on several diseases such as improving symptoms of infantile colic, reducing diarrhea in children, preventing bone loss in the elderly and promoting regulatory immune system development (Santos et al., 2009; Schepper et al., 2019).

As a **lactic acid bacterium**, some strains of *L. reuteri* have been applied to a large variety of food products and food supplements. The lactic acid bacterium is generally recognized as a safe microorganism such as the bacteria in yogurt (Alayande et al., 2020; Sauer et al., 2017; Vankerckhoven et al., 2008).

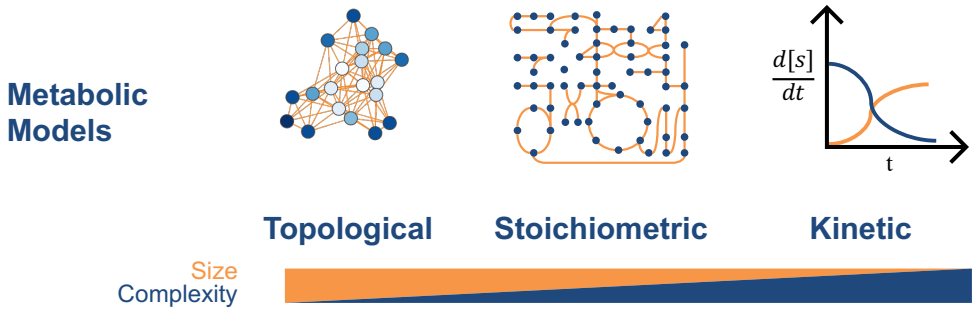
As a potential **cell factory**, *L. reuteri* produces several industrially important compounds such as 1,3-propanediol, folate, and reuterin. Some studies discussed the potential of *L. reuteri* as a cell factory and provided some metabolic engineering strategies (Bosma et al., 2017; Kristjansdottir et al., 2019).

Osteoporosis is a bone disease that reduces bone mineral density and decreases bone strength. This disease increases the risk of fracture in the elderly population (Britton et al., 2014; Nilsson et al., 2018). Due to the limited understanding of osteoporosis mechanisms, the treatment rate for patients is very low. The gut microbiota has been suggested to have a positive effect on bone metabolism, and gut microbiota supplements are a potentially effective intervention for the prevention and treatment of osteoporosis. Some studies showed that some specific probiotic strains of *Limosilactobacillus* can produce butyrate and thus promote bone formation in mice (Britton et al., 2014; Nilsson et al., 2018). Supplements of *L. reuteri* ATCC PTA 6475 chewable tablets could also prevent osteoporosis in a mouse model. The interest in studying *L. reuteri* has increased significantly in recent years. More and more studies have shown that *L. reuteri* or *L. reuteri,* which produced butyrate can improve bone mineral density loss (Mu et al., 2018).

1.1.5 Gut microbiota and type 2 diabetes

Type 2 diabetes (T2D) is characterized by hyperglycaemia and some metabolic disorders like dysfunction of insulin resistance and pancreatic beta-cell, abnormal glucose and fatty acid metabolisms and limited control of the postprandial glycemic level (Association, 2011; Tabák et al., 2012). Increasing metagenomic studies have suggested that diabetes is associated with altered gut microbiota composition and functional capacity. For example, a previous study indicated that discriminatory microbial markers of T2D were heterogeneous between cohorts and revealed an increase in the abundance of *Lactobacillus* species and a decrease in the number of *Clostridium* species in individuals with T2D (Karlsson et al., 2013). Furthermore, enrichment of *Escherichia coli* and an increased quantity of *Bacteroides* spp. also related to T2D. Some gut microbiotas affect glucose tolerance and T2D by producing detrimentally microbial metabolites such as branched-chain amino acid (BCAA) (Pedersen et al., 2016), imidazole propionate (Koh et al., 2018) and butyrate (J. Wang et al., 2012).

1.2. Steady-stay modeling and Genome-scale metabolic models (GEMs)



Example	Metabolic networks	GEMs	Cybernetic models
Type	Steady states	Steady states	Dynamic
Data structure	Edge & Nodes (& Weights)	Reactions (with GPR) & Metabolites (& Constraints)	Pathways (or Reactions) & kinetic parameters
Predictions	Pathways	Pathways & Fluxes	Concentrations (dynamic)
Papers	Paper V	Paper I & VI	Paper III

Figure 1.2| Different approaches of metabolic modeling. Different types of metabolic models have their advantages for different type of research questions. Topological metabolic networks require biochemical information to represent nodes and edges. Considering metabolites and enzymes as the nodes in a graph separately, the network could be 'metabolite-centric' or 'enzyme-centric'. Stoichiometric models like GEMs require gene-protein-reaction biochemical information of target organisms. Quantitative constraints such as substrate uptake rate can make quantitative flux predictions. Kinetic models like cybernetic models could predict metabolites abundance in a specific time span. Some kinetic models need to reduce the scope of the models to the level of pathways but also many kinetic parameters. These parameters are often used in specific contexts.

Different types of metabolic models have been developed to understand and simulate biological metabolism (**Figure 1.2**). In systems biology, different models can be primarily categorized into three groups, topological, stoichiometric, and kinetic. Topological models require little data to create simple edges or links between nodes (Steuer, 2007). There are two types of networks, one is an enzyme-centric network, and the other is a metabolite-centric network. The former uses enzymes as nodes, and metabolites as edges, and the latter is the opposite (Aric A. Hagberg et al., 2008; Bauer & Thiele, 2018). Dynamic models are more complex than topological and stoichiometric models. They use enzyme capacity data as a parameter to simulate or predict dynamic systems under specific conditions. Genome-scale metabolic models (GEMs) are useful tools in metabolic engineering that could help us understand the metabolism and physiology of an organism (Fang et al., 2020). GEMs provide a platform to efficiently integrate genome sequences, experimental data, and other types of data (Y. Kim et al., 2020; Rana et al., 2020). In this thesis, we use both stoichiometric and kinetic models to study human gut microbiota metabolism.

1.2.1 A map of metabolic pathways: Genome-scale metabolic model

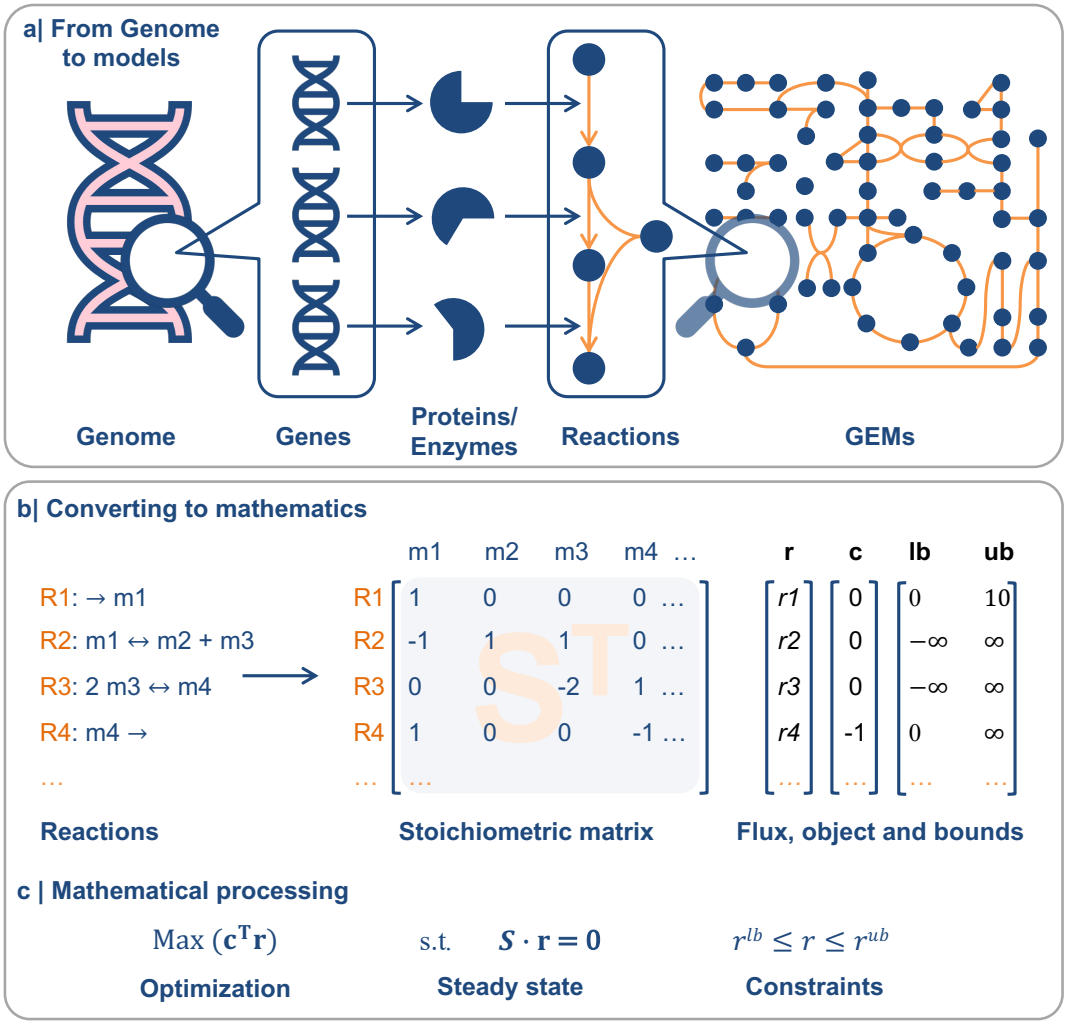


Figure 1.3| Genome-scale metabolic models. a) GEM reconstruction. Metabolic genes from the genome are annotated, and the information for encoding proteins or enzymes that catalyze metabolic reactions were extracted. All reactions in the GEM are connected by shared metabolites and make up a network. **b)** The stoichiometric matrix (**S**) represents all reactions and metabolites in a GEM. Each reaction is represented by assigning the stoichiometric coefficients of the participating metabolites, where negative values indicate consumption and positive values indicate production. The fluxes, objects, and bounds can be represented as vectors. **c)** The flux balance analysis (FBA) has been widely used to optimize objective functions and predict the fluxes (**r**). The unit in a GEM is in mmol per gram dry cell weight per hour (mmol/gdw/h). The optimization assumes a steady state and all fluxes are constrained by bounds vectors.

A genome-scale metabolic model (GEM) integrates genes, enzymes, metabolic reactions, and metabolites through gene-protein-reaction (GPR) associations (**Figure 1.3**) (Finley & Hatzimanikatis, 2021; Ye et al., 2022). The model serves as a platform for systematic metabolic analysis and can describe the complete metabolic transformation of cells. With the development of high-throughput and low-cost sequencing technologies, many GEMs have been reconstructed for different organisms, including *E. coli* (Monk et al., 2017), yeast (Lu et al., 2019), mouse and humans (H. Wang et al., 2021). These GEMs have many applications, such as industrial strain design in metabolic engineering, simulating species interactions, and analyzing coupled reaction sets (G. B. Kim et al., 2020; Proffitt et al., 2022).

GEM reconstruction. The first step of GEM reconstruction is genome annotation. Based on BLAST results and genes' functions, corresponding enzymes and metabolic reactions are collected from databases into a draft model. Some reactions do not have associated genes, such as natural exchange reactions and biomass synthesis reactions. These reactions need to be added on the basis of the experimental data. The draft model contains "gaps" or missing information due to our incomplete understanding of metabolism and genome. These gaps need to be added by algorithmics. In order to build a high-quality GEM, a lot of manual work is required. Many curations require experimental data and cannot be replaced by automatic tools. With the development of sequencing technology, more and more species of GEMs need to be constructed. Several automatic and semiautomatic tools, such as Carveme (Machado et al., 2018) and ModelSEED (Seaver et al., 2021), have recently been developed. These tools allow the construction of draft models based on the genome sequences, thus accelerating the reconstruction process.

1.2.2 Rate and yield

Rate and yield in production processes are both critical characteristics for evaluating industrial strains, and they focus on productivity and efficiency with different criteria (**Table 1.1**). Rate measures the speed of product formed and concerns per unit of time. Yield is a relative value to measure the efficiency of conversions and is more concerned with the amount of substrate consumed. For example, the biomass rate and yield indicate the amount of gram dry weight of biomass formed, but the rate is measured in units of time (hour), and the yield is measured in units of substrate consumed (g). The rate and yield are not independent of each other and usually there is a trade-off between them, such as the ATP rate and ATP yield in yeast respiration and fermentation. Respiration generates ATP with a high yield and low rate. Oppositely, fermentation has a higher ATP rate but with a lower yield.

Table 1.1| Comparison of rate and yield.

	Rate	Yield
Mathematical	r_p	r_p/r_s
Units	mmol/gDW/h; per unit of time	One (g/g or mmol/mmol); per amount of substrate consumed
Description	Productivity / Speed	Efficiency of conversions
Optimization problem	Linear program	Linear-fractional program
Optimization methods	FBA	opt-yield, FBA*, EFM/EFV+,

FBA*, fixed substrate uptake rate (fixed r_s)
EFM/EFV+, not optimization methods (strictly), return a set of pathways.

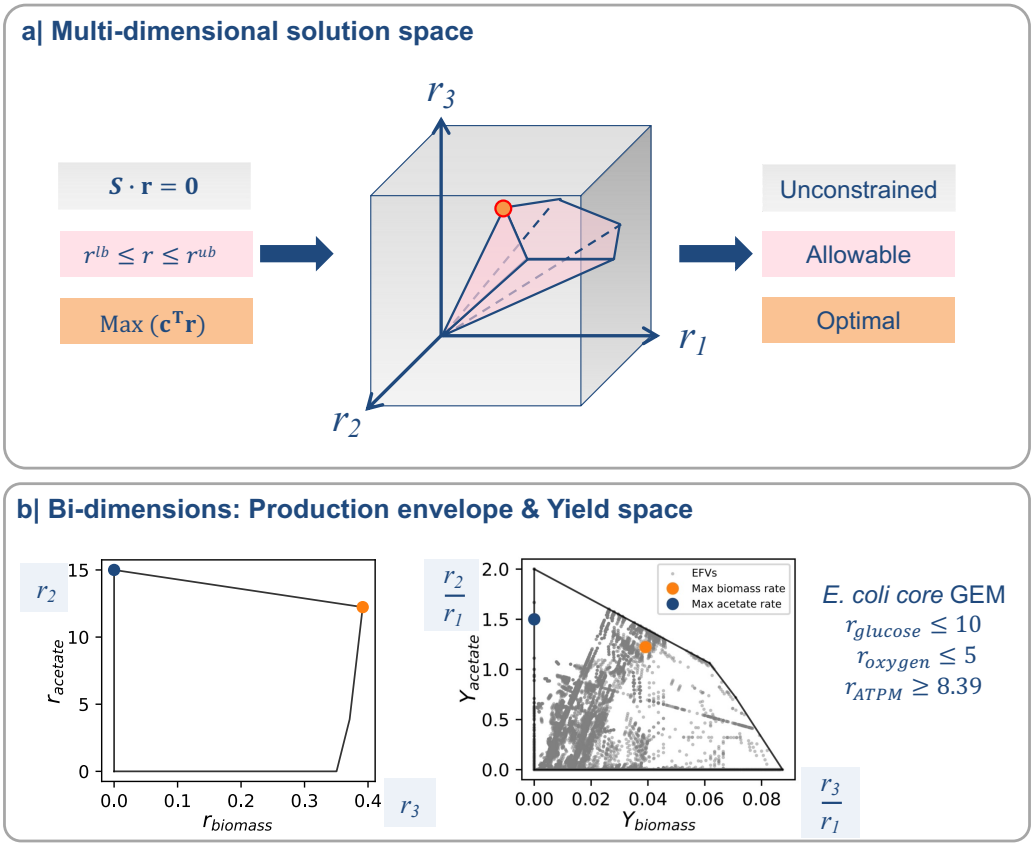


Figure 1.4| Solution space, production envelope, and yield space. a) Feasible solution space of the linear problem in a GEM. The solution space is infinite without constraints (grey background). After adding constraints for reactions in a GEM, the feasible solution space will also be constrained. If only linear constraints existed, the allowable solution space would be a convex polyhedron, indicated by the pink background. The optimization is searching for a specific solution in a feasible space. For example, FBA is optimizing the objective function indicated by orange. The returned maximum is unique, but the returned vector is not unique. For instance, multiple flux distributions (r vector) often result in the same optimal objects when maximizing growth. The distributions often contain loops and parsimonious FBA (pFBA) use a second optimization to remove loops. The pFBA returns a vector with the same objective values but with the lowest sum of fluxes. **b)** Production envelopes and yield space of acetate-biomass in the core *E. coli* GEM and the locations of optimal acetate and biomass rate pathways. The solution space can be reduced to a two-dimensional space. For example, keep the r_2 and r_3 and get the production envelope of r_2 and r_3. Letting r_2 and r_3 divide by the substrate flux r_1, then get yield space.

The feasible solution space of reaction flux vectors is a bounded polyhedron in geometry **Figure 1.4a**, and both rate and yield optimization could be considered as searching for the optimal pathways in it. If we focus on the variability of two target products, the solution space could be present as 2-dimensional production envelopes or yield space. The production envelope and the yield space could help us to understand the rate and yield relationships between the two products. The difference between the rate and the yield solution space can be highlighted. The yield is a property of a pathway and not as flexible as fluxes in our toy model. In **Figure 1.4b**, we map the optimal acetate and biomass rate pathways to the yield space based on the flux distributions, and it is clear that the optimal rate pathway cannot represent the optimal yield pathway. The rate-optimal pathway is not unique (underdetermined system), and the yield might not be unique either.

Rate- and yield- optimizations for genome-scale metabolic models (GEMs) are helpful methods to help us design strains according to the metabolic flux distributions. Flux-balance analysis (FBA) is a fundamental method for optimizing target reactions or predicting flux distributions in a steady state. The objective function of FBA usually is the growth rate or product rate, and the optimization can be solved by linear programming. FBA is clearly maximizing rates in mathematics, and it is also used to solve yield problems by fixing the substrate uptake rate in some studies. For example, by fixing substrate uptake rates as experimentally observed values or normalizing substrate uptake as one, optimizing the target rate is equal to optimizing yield. And there are also other methods to find optimal yield by EFM/EFV or linear fractional programming.

Elementary flux mode and vector (EFM/EFV): EFMs are a set of topologically feasible, non-decomposable pathways underlying the steady state in models. This set of pathways fully characterizes the available metabolic space of a GEM. EFMs are usually used as a structural concept to identify interested targets robustness of models and a non-negative linear superposition of EFMs could express any feasible results without constraints. EFV is similar to EFM but is used to deal with inhomogeneous constraints and can be considered as a catalogue of FBA pathways (Klamt et al., 2017). With the existing inhomogeneous constraints, EFV could identify disserved pathways under complex conditions. Both EFM and EFV are used to find optimal yield or yield space and they performed very well for medium size models. The mathematical problem is similar to the enumeration of a polyhedron's corners and edges. Some tools like efmtool (Terzer & Stelling, 2008) that is based on double description method (DDM) approaches require a huge amount of random-access memory and resources. Currently, both EFMs and EFVs limited to small-size or medium-size networks. Some other approaches like the lexicographic reverse search (lrs) based motheds have made some progress but still no many applications. MATLAB toolbox efmtool and Metatool (von Kamp & Schuster, 2006), Python packages efmlrs (Buchner & Zanghellini, 2021) (recommended, compatible with cobrapy) and efmtool (python) support EFM and EFV calculations.

1.2.3 Metabolic modelling of microbial communities

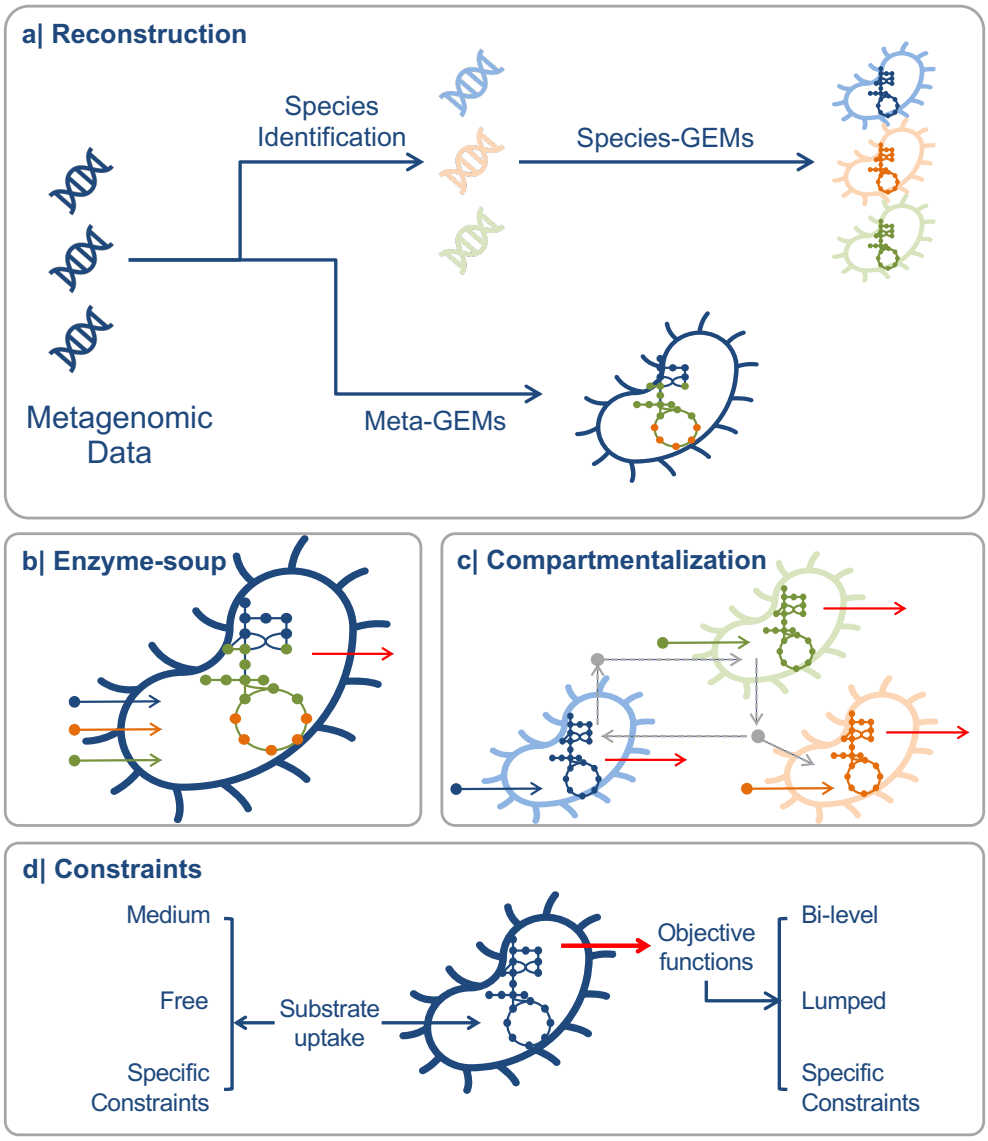


Figure 1.5| Genome-scale metabolic models for microbial communities. a) Metagenomic data-based GEMs reconstruction approaches. From the metagenomic data, species can usually be identified by 16S rRNA genes or other marker genes. Based on taxonomic profiles, the specific whole genome sequences could be downloaded from public databases such as NCBI. The following step is typical GEM reconstruction. Depending on a large number of species, some automatic and semi-automatic tools are generally used. An alternative approach is constructing an enzyme-soup model directly from the metagenomic data. **b)** Enzyme-soup model. Assuming that all reactions and metabolites are present in one 'super-cell-pool'. **c)** Compartmentalization. Compartments separate reactions and metabolites from different species. **d)** Constraints for substrate and objective functions. The specific constraints usually according to the taxonomic abundance profiles. The red arrows represent the objective function. In enzyme-soup and compartmentalization methods, one lumped objective function is optimized. Some bi-level methods have inner and outer objective functions.

With the increasing availability of high-throughput omics data, an ever increasing number of GEMs have been reconstructed by automatic tools. It is now possible to apply GEMs to systems of multispecies and communities. Communities GEMs would offer insight into metabolic machinery and interactions between species. Several approaches have been

developed for different concerns. Depending on research questions, some frequently used methods for modeling microbial communities are the enzyme-soup model, compartmentalization, pan/core-models, meta-GEMs, separation of species-level and community-level objective functions, etc.

Enzyme-soup approach is the simplest method to evaluate the potential metabolic capacity of a community. **Meta-genome** models are built by the enzyme-soup approach, as shown in **Figure 1.5a** and **1.5b**. The model is built by annotating all the genetic material in an environmental sample. The model contains many individual organisms' genomes and reactions. Model reactions are compiled into a single stoichiometric matrix and usually have one objective function for the community. This approach focuses on environment-community interactions; it does not provide species-species interactions and cross-species metabolic exchanges (Abubucker et al., 2012; Henry et al., 2016) .

Compartmentalized models are developed to predict cross-species metabolic exchanges and species-species interactions. Each species model is compartmentalized and maintains all metabolic capacity. This approach usually creates a new environmental compartment to let species share external metabolites and medium. The S matrices are combined into a "meta-stoichiometric matrix". The objective function could be bi-level, species objects, and an umped community object. The relative abundance of microbes in the community is a vital evaluation indicator for metagenomic data and is usually used as a constraint. The constraints are used to limit substrate distributions of the community members or formulate a weighted linear combination of the biomass reactions (Shoaie et al., 2013; Shoaie & Nielsen, 2014; Stolyar et al., 2007) .

Pan/core models are used to analyze the genome characteristics and phenotypic diversity among different strains of a species. It is a collection of GEMs for single microorganisms and can be further developed into enzyme-soup models or compartmentalized models (Y. Kim et al., 2020; G. Li et al., 2019; Lu et al., 2021).

1.3. Dynamic modeling and Cybernetic models

Cybernetic modeling is a dynamic modeling approach that incorporates control strategies for optimal investment of intracellular resources. The hypothesis is based on the fact that microorganisms can regulate the synthesis and activity of enzymes as a survival mechanism for cells when encountering environmental perturbations. Different from regular kinetic modeling, this framework introduced two key regulatory variables (*u* and *v*) that were derived based upon two laws (the matching law and proportional law) to regulate the "key" enzymes that were responsible for the uptake of different substrates. The models can describe not only the dynamic behavior of microbes in mixed carbon substrate environments but also the multi-species coculture relationship between biomass and products among multi-species.

1.3.1 Mitigating "traffic congestion" in metabolism: cybernetic Approach

In 1942, Monod observed the phenomenon of "diauxic" growth, in which the bacteria prefer to recruit glucose as carbon sources rather than xylose (Monod, 1942). When all glucose had been nearly consumed, the bacteria began to utilize other carbon sources, such as xylose. In this context, many researchers have tried to explain these phenomena rationally with scientific laws linking cause and effect. As one of the outstanding hypotheses, Monod thinks that the enzymes responsible for glucose uptake should be synthesized before those for xylose (Monod, 1978). Based on this clue, Doraiswami Ramkrishna and his group were led to postulate that the organism must make frugal use of its resources for enzyme synthesis and tried to describe "diauxic" growth using a mathematical framework (Ramkrishna, 1983). They believe that the regulatory phenomena cooperated with a cybernetic system that could regulate the resources for enzyme synthesis to conform to their hypothesis. Here the word "cybernetic" is derived from the Greek "χυβερτησ" which means steersman approach. The term was defined as the science of "control and communication in the animal and machine" by Norbert Wiener (Wiener, 1965). Within this framework, microorganisms were considered to have an optimal control strategy toward their survival goal, i.e., the maximization of growth under different substrates environments. The simple explanation of this theory is that if we consider that there are "key" enzymes responsible for the uptake of different carbon sources, such as glucose and xylose, the "diauxic" growth phenomenon can be considered as a wise resource investment for the synthesis of different "key" enzymes through their optimal regulation strategy (Monod, 1942). Based on the assumption that the microorganisms will drive metabolism through the optimal investment of their resources, control of enzyme synthesis and activity as survival mechanistic, the development focuses on dynamic modeling and has excellent achievement, especially under environments with multiple complementary substrates (Narang et al., 1997; Ramakrishna et al., 1996). The idea of regulatory processes is also unique from other kinetic modeling or other dynamic modeling. In the cybernetic model framework, a microorganism cell could be considered a combination of a series of machines (Ramkrishna & Song, 2018). The components of such adaptive machinery within the cell can be referred to **Figure 1.6**.

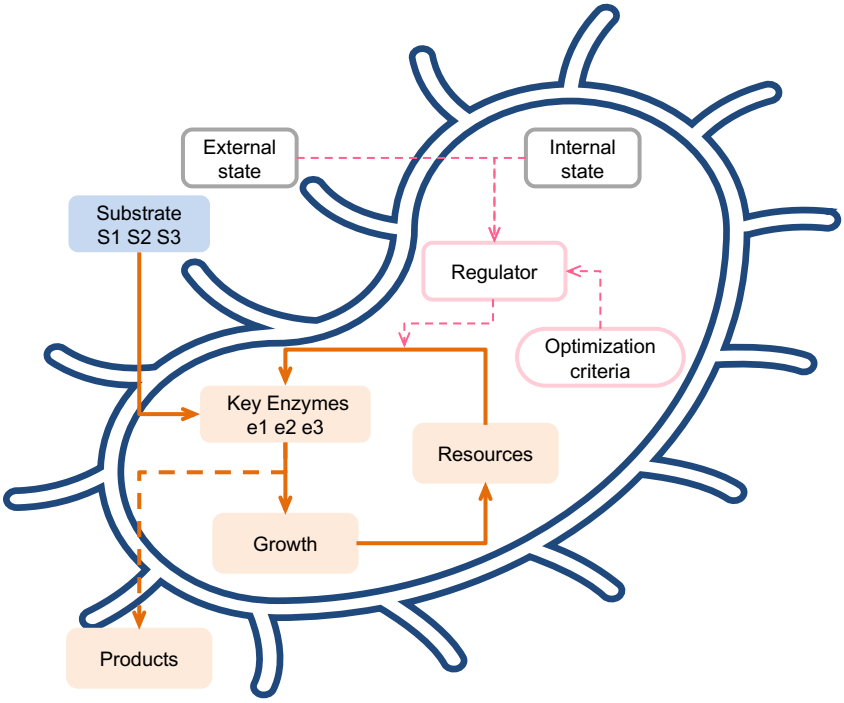


Figure 1.6| The cell in a cybernetic view. The cell of a microorganism could be viewed as a combination of machines. Adaptive machinery (external and internal states adaptive), permanent machinery (permanent reactions for growth), and regulators (control the resources). The bold orange lines are recourses flow, from the substrate to growth or production. The dashed lines in pink are adaptive or regulatory processes. The key enzymes respond for different substrate uptake, and the key enzymes also require synthesis reactions that are regulated by the resources and regulators. The regulator can adapt to external and internal states, as well as to optimization criteria.

1.3.2 The development of cybernetic modeling

The motivation of the cybernetic modeling approach is aimed to describe the "diauxic" growth phenomenon of bacteria that prefer to utilize glucose instead of xylose (Monod, 1978, 1942). After a lot of combinations of substrate experiment observations, some researchers believe that the cell response is under navigation toward a survival goal. After three decades of development, the following researchers believe that cybernetics refers to “the art of steering a system toward a goal” (Ramkrishna & Song, 2018). Based on the assumption that microorganisms will drive their metabolic behavior through the optimal investment of cellular resources, i.e., control of enzyme synthesis and activity, modeling development focusing on dynamic modeling has achieved some significant progress, especially in modeling complex environments with multiple complementary substrates (Song et al., 2011; Song & Ramkrishna, 2013).

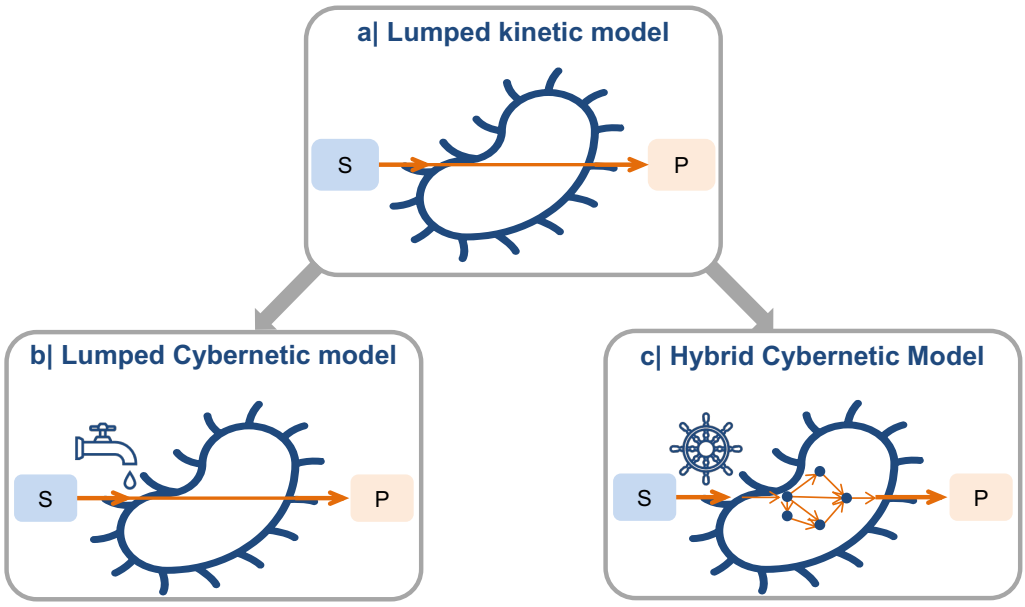


Figure 1.7| Schematic description of the traditional lumped kinetic model, lumped hybrid cybernetic models (L-HCM) and hybrid cybernetic models (HCM). The traditional lumped kinetic model has no regulations. L-HCM incorporated enzyme regulation and lumped all pathways from the same substrate to production into one pathway. The model is helpful for complex media that contain more than three carbon sources. The HCM retains more representative pathways from EFMs, which means there will be more key enzymes and more parameters than L-HCM.

Regulatory processes make it distinct from other kinetic modeling or other dynamic modeling. Ramkrishna developed a mathematical framework in 1982 that fixed the number of resources used to synthesize enzymes from different carbon sources (Ramkrishna, 1983). The theory via integrating regulation of the resource investment has described many diauxic behaviors for several substrate combinations. After years of improvement, Straight (1991) tried to apply the theory to metabolic networks to describe metabolic performance with complementary substrates (Straight & Ramkrishna, 1991). Ramakrishna (1996) refined Straight's work and formulated a cybernetic model using a simple network (Narang et al., 1997). And Varner published some papers to apply cybernetic modeling on more extensive networks and Young (2005) further extended to large networks by developing elementary modes method, called hybrid cybernetic models (HCM), with a fresh approach to the derivation of cybernetic laws (Young, 2005). With the expansion of metabolite networks and the knowledge of molecular details, this framework has some derivative methods for processing large metabolic networks, such as lumped hybrid cybernetic model (Song & Ramkrishna, 2010). The models can accurately simulate and describe microorganisms' dynamic behavior and even predict product formations. In 2008, Kim differentiated the intracellular and the cellular variables, which made get variables essayer and possible for genome-scale networks (J. Kim et al., 2008).

Cybernetic models adopted some advantages of lumped kinetic model, which sampled the internal cell process, and extended some regulation machineries. Two outstanding types of cybernetic models are lumped hybrid cybernetic model (L-HCM) (Song & Ramkrishna, 2010)

and hybrid cybernetic models (HCM) (J. Kim et al., 2008). Complex models introducing enzymatic parameters for each reaction are not presented here because they are not feasible within current computational resources and knowledge. The difference between the lumped kinetic model, L-HCM and HCM is shown in **Figure 1.7**.

1.3.3 Cybernetic variables and cybernetic laws

Cybernetic variables: As we have mentioned above, cybernetic models could regulate enzyme synthesis and enzyme activities, but how to regulate? As shown in **Figure 1.8**, the cybernetic model introduced two cybernetic variables, where *u* controls the enzyme synthesis and *v* controls enzyme activities.

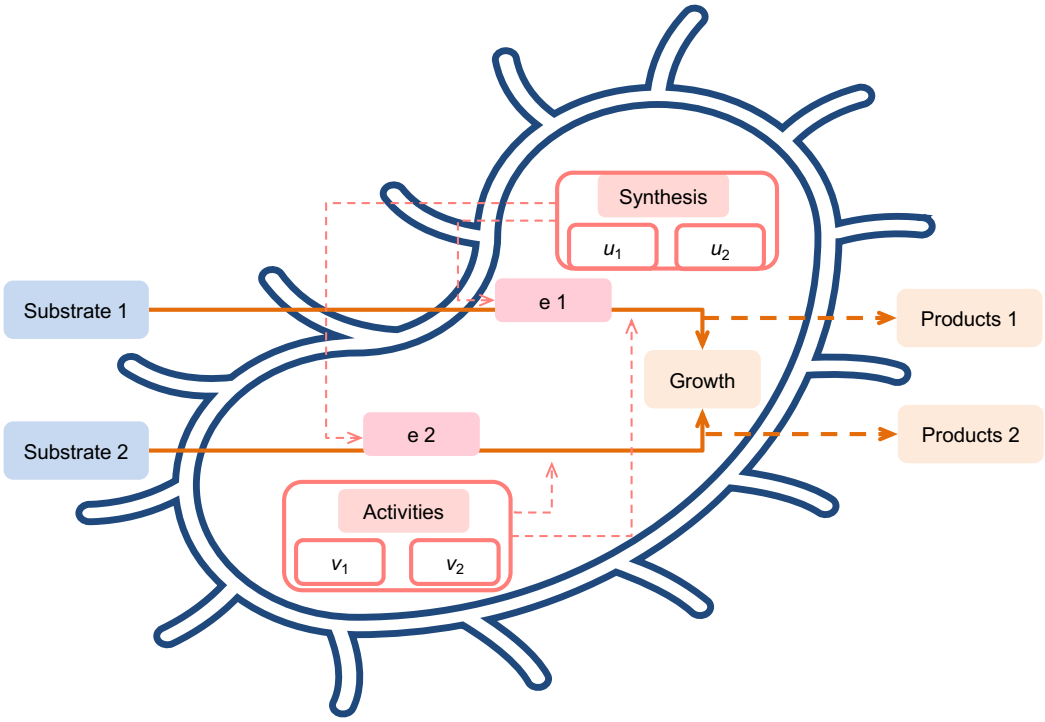


Figure 1.8| The regulation machinery via cybernetic variables. The cybernetic variable represents the percentage of resources allocated. The u_i controls the synthesis of key enzymes and v controls the activities of enzymes.

The u_i is the fractional allocation of resources for synthesizing e_i, and v_i is the fractional allocation of resources for the activity of e_i. Therefore, the actual rate of enzyme synthesis of enzyme will be regulated by u_i:

$$r_{Ei}\, u_i\,, (0 \leq\ u_i \leq 1) \& \left(\sum\ u_i = 1\right)$$

And as defined, the u_i is a fraction, should less than or equal to 1, and sum to 1. Similar with v_i, the actual rate of activity of enzyme:

$$r_i\, v_i\,, (0 \leq\ v_i \leq 1)$$

The v_i should less than or equal to 1, but not necessarily sum to 1. Because the resources for enzyme synthesis process are competitive, but enzyme activity is not strictly competition. In other words, all parallel enzyme activities can be maximized, but synthesis cannot be. For example, the total growth rate:

$$\sum r_i \, v_i \ , (0 \leq \ v_i \leq 1)$$

Here all the v_i could be 1, which means all the enzymes for growth are maximally activated. But not all the u_i can be 1, because the sum of resources for the synthesis is 1, which are competing with each other for cellular resources.

The u_i incorporates regulatory action of repression and induction and the v_i incorporates inhibition and activation. In the next two sections, the mathematic desecrations of $u_i and \ v_i$ will be introduced, as can be derived based on two cybernetic laws, i.e., the matching law and the proportional law.

Matching Law: Mathematical expressions should be based on real-world phenomenon. Before introducing the cybernetic laws, we will introduce how the cybernetic variables regulate the recourses. At first the variables u_i controlled how much resource could be used to synthesis specific enzyme.

That means the investment of u_i is the resource, and the return is the amount of enzyme e_i. According to the law of diminishing marginal utility and limited resources,

$$\text{Max } Total \ enzymes \ = \sum e_i(r_i)$$
$$\text{s.t.}: Total \ Resources \ R = \sum r_i$$

Furthermore, the p_i as return-on-investment for resource allocated to the i_{th} enzyme, which is linear with the amount of enzyme. The resource investment is linear with the amount of enzyme and return obtained:

$$\frac{e_1}{r_1} = \frac{e_2}{r_2} = \frac{e_3}{r_3} = \frac{e_n}{r_n} \ \& \ \frac{p_1}{r_1} = \frac{p_2}{r_2} = \frac{p_3}{r_3} = \frac{p_i}{r_i}$$

Therefore, the mathematical definition of u_i, the Matching Law:

$$u_i = \frac{r_i}{\sum_j r_j} = \frac{p_i}{\sum_j p_j} \ (0 \leq \ u_i \leq 1) \& \ (\sum \ u_i = 1)$$

Proportional Law: Compared with Matching Law, Proportional Law is more complex, the postulates should be clear here.

Postulate 1: The activity of an enzyme that supports the fastest partially regulated reaction is promoted to the utmost. In other words, if the i th reaction has the fastest partially controlled rate, v_i is set equal to unity.

Postulate 2: The activity of an enzyme supporting a reaction, which has a partially controlled rate lower than the maximum, is proportional to its rate.

Based on the two postulates, v is proportional to the fastest reaction rate:

$$v_i = \lambda\, r_i$$

This proportionality combined with constraints determine the bound on λ:

$$0 \le v_i \le 1 \Rightarrow 0 \le \lambda \le 1 \; or \; \lambda \le \frac{1}{\max\,(r_j)}$$

The actual rate of reaction as we mentioned:

$$\sum\nolimits_j r_i\, v_i\,, (0 \le\ v_i \le 1)$$

Combine the three equations:

$$\sum\nolimits_j r_i\, v_i \;\Rightarrow\; \lambda \sum\nolimits_j r_i^2 \le\ \frac{1}{\max\,(r_j)} \sum\nolimits_j r_j^2\,, (0 \le\ v_i \le 1)$$

The maximum reaction rate is when:

$$\lambda = \frac{1}{\max\,(r_j)}$$

And then the v_i math definition by Proportional Law:

$$v_i = \frac{r_i}{\max\,(r_j)}$$

The proportional law is concerned with allosteric and covalent regulatory controls that modulate relative enzyme activities. Enzyme activities are properties of the enzyme itself; the regulation is not strict competition as enzyme synthesis. And the investment in signals (s_i) is NOT linear with the return on activity (a_i). Here, the relative activity of an enzyme means activity relative to its maximum value. And the regulation seems like signal regulates, not recourses itself. So, the sum of v_i could more than 1.

1.3.4 Rate equation and mass balances of cybernetic models

Basic Model formulation: To introduce the cybernetic model, the assumed variables **x** should contain metabolites, enzymes and growth. The vector **x** describes the state of the metabolic system at any instant, which contains vector **m,** the concentration vector of metabolites, c is biomass concentration, **e** is the vector of enzyme concentration, **x** vector y was given by: Where the vector of metabolites **m** is further decomposed into extracellular ($\mathbf{m_{ex}}$) and intracellular metabolites ($\mathbf{m_{in}}$).

$$\mathbf{x} = \begin{bmatrix} \mathbf{m} \\ \mathbf{e} \\ c \end{bmatrix}, \mathbf{m} = \begin{bmatrix} \mathbf{m_{ex}} \\ \mathbf{m_{in}} \end{bmatrix}$$

For most network models like GEMs, the c could be considered as the biomass, and the enzyme synthesis reactions network is considered to be excluded from the network. The key enzyme could be a combination of enzymes or a pseudo enzyme. Since the system needs to consider two-parts mass balance: metabolites balance and enzyme balance.

Mass balance: After defining the variable of vector **x**, all the dx/dt could be described. Mass balances for extracellular intracellular metabolites $\mathbf{m_{in}}$. Sometimes the intracellular metabolites are under a pseudo-steady state, the concentration will not change,

$$\frac{d\mathbf{m_{in}}}{dt} = \mathbf{S_{in}} \cdot \mathbf{r} - \mu \cdot \mathbf{m_{in}} = \mathbf{0} \tag{1.1}$$

For extracellular metabolites concentration ($\mathbf{m_{ex}}$):

$$\frac{d\mathbf{m_{ex}}}{dt} = \mathbf{S_{ex}} \cdot \mathbf{Z} \cdot \mathbf{diag}(\mathbf{v}) \cdot \mathbf{r} \cdot c \tag{1.2}$$

In equation (1.3), the $\mathbf{m_{ex}}$ (shape: n_{mets}) is a vector of extracellular metabolites concentrations; the $\mathbf{S_{ex}}$ (shape: n_{mets} * n_{rxns}) is the stoichiometric matrix of GEMs extracellular parts; the Z is the (shape: n_{rxns} * n_{paths}) stoichiometric matrix of pathways and extracellular metabolites, the pathways could be EFMs or EFVs or opt yield FBA pathways; the $\mathbf{diag}(\mathbf{v})$ is a diagonalization **v** (shape: n_{paths}), **v** is the vector of the cybernetic control variables that regulate the enzyme activities that will defend by equation (1.11); the **r** (shape: n_{paths}) is the vector of exchange fluxes from corresponding pathways; c is the biomass concentration per unit volume. Here the results of $\mathbf{S_{ex}} \cdot \mathbf{Z}$ could be normalized with substrates. All the extracellular metabolites concentration related to the pathway yield from $\mathbf{S_{ex}} \cdot \mathbf{Z}$, rate **r** is regulated by cybernetic control variables (**v**). All metabolites' concentrations are treated by per c relating to unit biomass. For some detailed small systems (three - five reactions), the **Z** could be ignored.

Enzyme balance: The rate of enzyme synthesis is related to the available resources and substrate concentration. Employing a Michaelis–Menten type of kinetic expression for the inductive rate of enzyme synthesis, an enzyme balance within the cell becomes feasible, as shown below (victor of **e** or variable *e*).

$$\frac{d\mathbf{e}}{dt} = \boldsymbol{\alpha} + \mathbf{diag}(\mathbf{u}) \cdot \mathbf{r_E} - (\mathbf{diag}(\boldsymbol{\beta}) + \mu) \cdot \mathbf{e} \tag{1.3}$$

$$\frac{de}{dt} = \alpha + \frac{k_e s}{(K+s)} u - (\beta + \mu) \cdot e$$

In the equation (1.3), represent the enzyme ($\mathbf{e}$) (n_{paths}), mass fraction of the enzyme in the biomass. $\boldsymbol{\alpha}$ (n_{paths}) denotes the vector of the constitutive rate of enzyme synthesis; while $\mathbf{r_E}$ represents the maximum enzyme synthesis rate which inductive the synthesis of the enzyme occurs without limitation of the resources. The $\mathbf{diag}(\mathbf{u})$ is a diagonalization $\mathbf{u}$ (n_{paths}), $\mathbf{u}$ is a vector of the cybernetic control variables that regulate the enzyme synthesis, which represents the limitation of resources, that will be defined by equation (1.10); $\boldsymbol{\beta}$ (n_{paths}), the vector is degradation rate, $\mu \cdot \mathbf{e}$ is the dilution rate by growth, In the equation (1.4), μ is the specific growth rate.

Kinetic and rate describing: Rates of pathways and enzyme synthesis and growth could be defended by the fowling equations:

$$\frac{dc}{dt} = \mu \cdot c \tag{1.4}$$

$$r_i = k_{maxi} \cdot e_i^{rel} \cdot \frac{s_j}{(K_{ij} + s_j)} \tag{1.5}$$

$$r_{Ei} = k_{ei} \cdot \frac{s_j}{(K_{ij} + s_j)} \tag{1.6}$$

$$\mu = \mathbf{h} \cdot \mathbf{r} \quad or \quad \mu = \mathbf{h} \cdot \mathbf{S_{in}} \cdot \mathbf{r} \tag{1.7}$$

Equations (5) and (6) are the Michaelis–Menten type of kinetic expression of rates of pathways maximized rate and enzyme synthesis. Here the *i* is the index of pathways, j is the index of substrates, different pathways could consume different substrates; where the k_{maxi} is the reaction rate constant (g/gDW/h); $K_{ij} and\ k_{ei}$ are the Michaelis–Menten constant (g/L); s_j denotes the substrate concentration; $\mathbf{h}$ is the matrix to find the biomass-related pathways and sum the biomass produce rate. The e_i^{rel} is the relative enzyme, which could be described by the following equations:

$$e_i^{rel} \equiv \frac{e_i}{e_i^{max}} \tag{1.8}$$

$$e_i^{max} = \frac{\alpha_i + k_{ei}}{\beta_i + \mu_i^{max}} \tag{1.9}$$

These two equations are under the condition of maximized enzyme level, which are related to enzyme balance described in equation (1.3). When left-hand side to zero and u = 1, the maximum enzyme level can be established as equation (1.9). Finally, formulation of matching law and proportional law to specify the cybernetic control variables as follows

$$u_i = \frac{p_i}{\sum_J p_j} \tag{1.10}$$

$$v_i = \frac{p_i}{\max(p_j)} \tag{1.11}$$

where the variable v_i and u_i are bounded between 0 and 1, signifies regulation 1 of enzyme activity and synthesis through allosteric control. Here the p is the return on investment:

$$p_i = f_{carboni} r_i \tag{1.12}$$

Here $f_{carboni}$ is the number of carbon atoms of substrate through the *i* pathways. For now, all the formulation of the cybernetic model is completed.

1.4. Aim and Scope

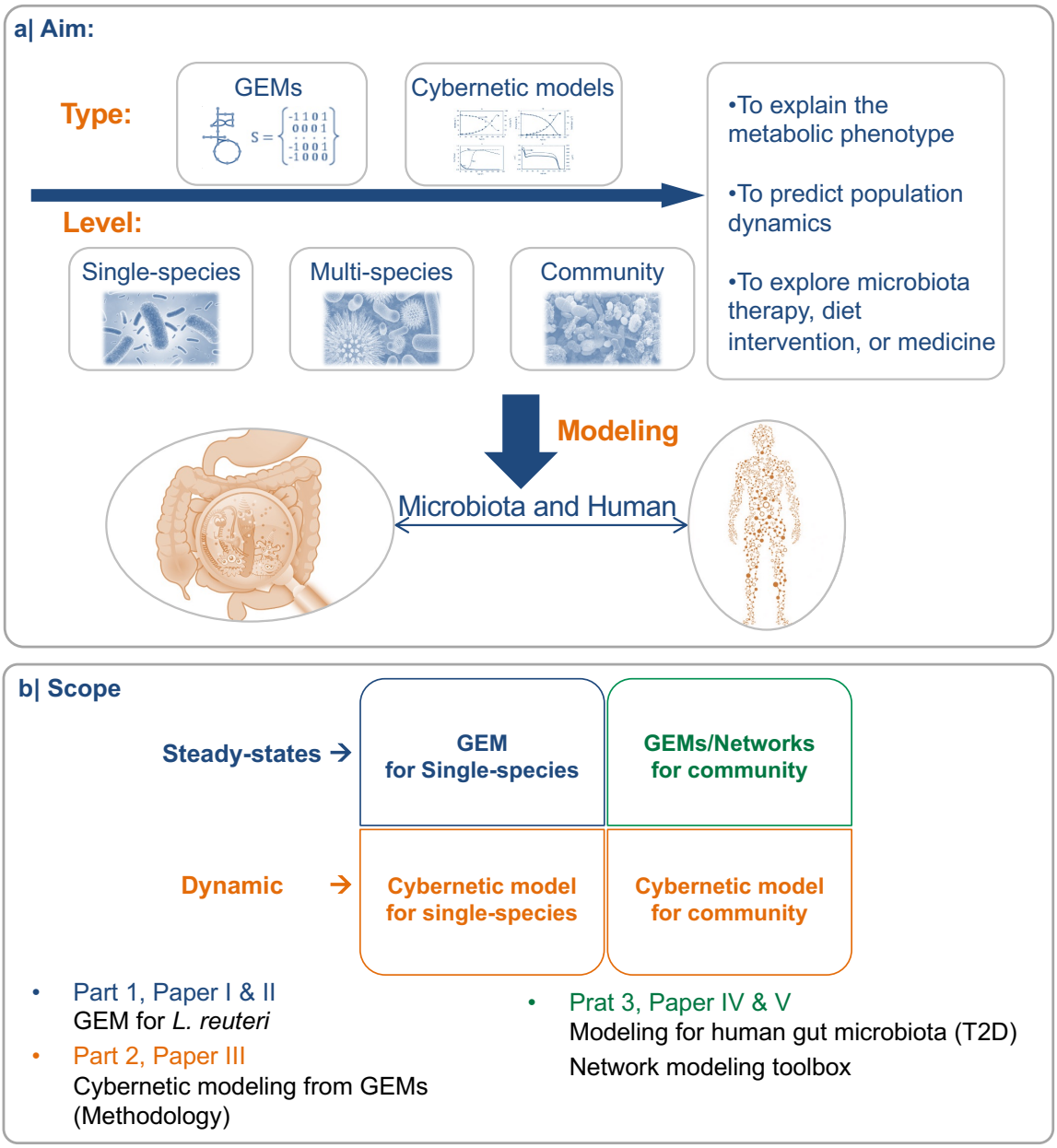


Figure 1.9| The aim and scope of this thesis.

The aim of this thesis is to explore, simulate, and predict the gut microbial ecosystem and the relationship between gut microbes and humans, as shown in **Figure 1.9a**. Due to the complexity of the gut microbial system, it is a challenge to describe the behaviors of the gut microbes by one type of model. We have divided our studies into two aspects: from the aspect of modeling methods (as mentioned in the introduction section), steady-states and dynamic-based modeling will be engaged; from the aspect of modeling subjects, single species to complex community will be addressed.

The scope matrix of this study is shown in the **Figure 1.9b**, modeling human gut microbiota from steady states to dynamics systems, from single species to community systems.

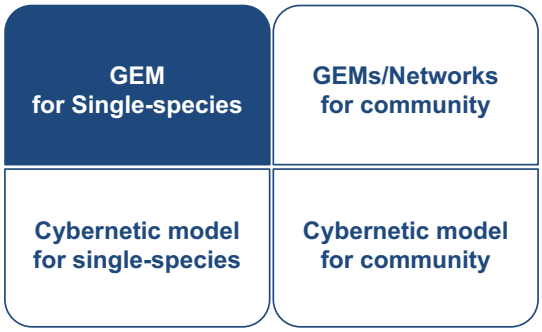
There are three parts in this thesis:

Part1, Steady states modeling for single species by *L. reuteri* GEMs (**Paper I & II**): in this part, we reconstruct GEMs for *L. reuteri ATCC PTA 6475* and 35 other *L. reuteri* strains from three types of hosts. We systematically investigate the metabolic features of *L. reuteri ATCC PTA 6475* and the metabolic versatility of *L. reuteri*. After that, we explore the mechanisms underlying the effect of *L. reuteri* ATCC PTA 6475 on bone metabolism and identify factors important for a good response to the probiotic.

Part2, Exploration of dynamic modeling by a cybernetic model (**Paper III**): in this part, we develop a hybrid cybernetic model strategy that can be applied to genome-scale metabolic models, we also illustrate the strategy by both single species and communities.

Part3, Steady states modeling for microbiota communities (**Paper IV & V**): in this part, we reconstructed communities GEMs for T2D cohorts and investigated T2D-related gut microbial signatures by machine learning. We also developed a python package for fast reconstruction of metabolic networks and analysis of the structure and properties of the networks.

2. Part I (Paper I & II): GEMs for *L. reuteri* and the effect on human metabolism



Related Papers:

Paper I: Genome-scale insights into the metabolic versatility of *Limosilactobacillus reuteri*

Paper II: One-year supplementation with *Lactobacillus reuteri ATCC PTA 6475* counteracts a degradation of gut microbiota in older women with low bone mineral density

Limosilactobacillus reuteri ATCC PTA 6475 (earlier known as *Lactobacillus reuteri*) is a perfect candidate for our study because it is a well-studied strain with both laboratory and clinical data. It has been identified as a probiotic and a member of the human gut microbiota. Some studies have confirmed the beneficial effects of orally administered *L. reuteri ATCC PTA 6475*, such as preventing bone loss. It is widely used in the market as a dietary supplement. Its health benefits may be due, in part, to the production of beneficial metabolites. To study its metabolic capacities, *L. reuteri* ATCC PTA 6475 GEM has been constructed. Considering the strain-specific effects and genetic diversity of *L. reuteri* strains, we were interested to study the metabolic versatility of these strains. Finally, we present alterations in the gut microbiota of order women with good or poor responses to orally administered *L. reuteri ATCC PTA 6475.*

2.1. Reconstruction and characteristics of *L. reuteri* ATCC PTA 6475 GEM

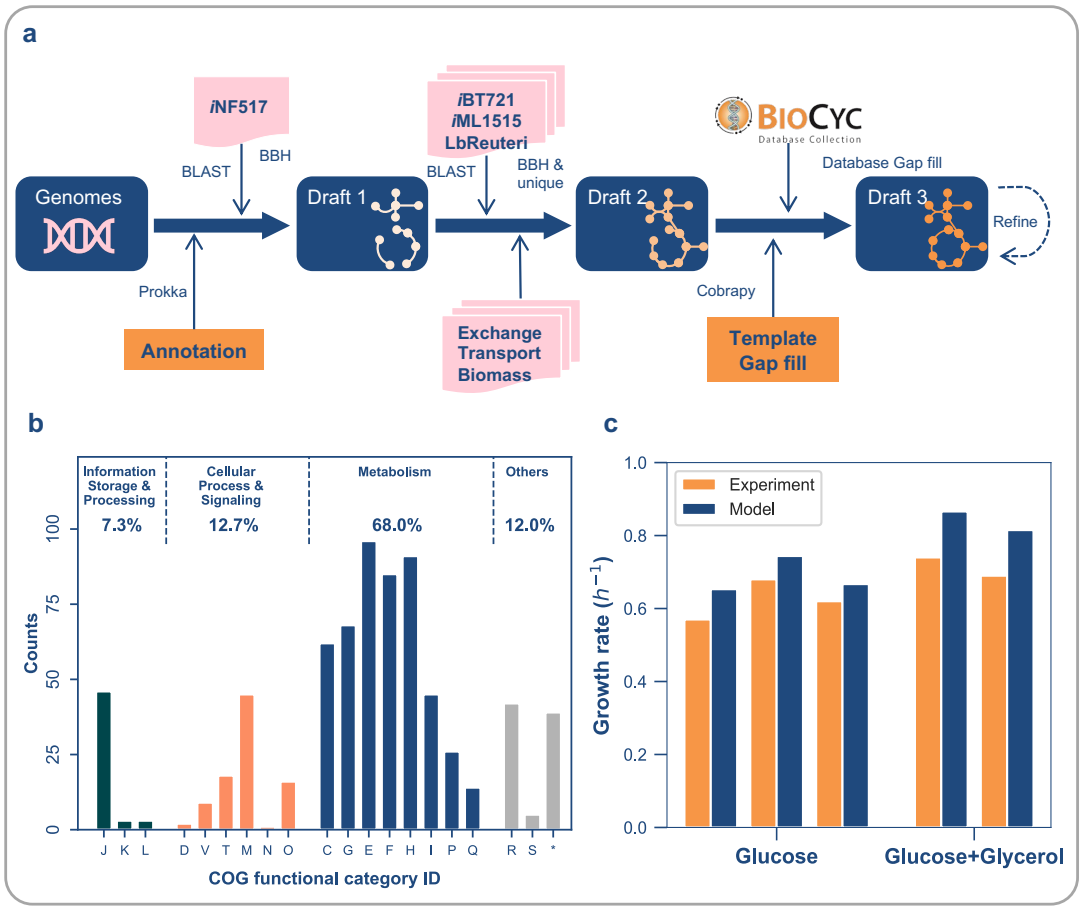


Figure 2.1 | **Reconstruction and characteristics of *L. reuteri ATCC PTA 6475* GEM. a)** Templated-based GEM reconstruction pipeline. To generate the draft models, the *i*NF517 was employed as the main template GEM. Ortholog genes and reactions and extracted based on bidirectional best hits (BBHs). More reactions were added to the draft model from LbReuteri, *i*ML1515 and *i*BT721 also based on BBHs. The exchange and transport reactions were added from the templates according to the transporter annotations and corresponding medium composition. The gap-filling was performed by COBRApy and the *i*NF517 used as template model. During the simulation and validation, GEM was also manually curated based on experimental data. **b)** GEM genes COG functional distribution. C, energy production and conversion; G, carbohydrate transport and metabolism; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; P, inorganic ion transport and metabolism; Q, secondary metabolites biosynthesis, transport and catabolism; J, translation, ribosomal structure and biogenesis; K, transcription; L, replication, recombination and repair; D, cell cycle control, cell division, chromosome partitioning; V, defense mechanisms; T, signal transduction mechanisms; M, cell wall/membrane/envelope biogenesis; N, cell motility; O, posttranslational modification, protein turnover, chaperones; R, general function prediction only; S, function unknown. *, no COG categories. **c)** Growth rate simulation and comparison with experimental data. The experimental data for each dataset is shown in orange, and the prediction is shown in dark blue.

The genome sequences of *L. reuteri* ATCC PTA 6475 are sequenced by BioGaia, and the annotation yielded 2,019 protein-encoding genes. On the basis of the sequence, GEM reconstruction applied a template-based approach, as shown in **Figure 2.1a**. The initial draft model is reconstructed on the basis of the template model *i*NF517. The other three template models were also applied to integrate extra biochemical reactions into the draft model and

generated draft 2. After incorporating exchange reactions and transport reactions to enable nutrient uptake and by-product secretion, the gap-filling was performed for growth. Many manual curations were performed to remove potential errors in reactions or metabolites. Altogether, the final GEM of *L. reuteri* ATCC PTA 6475 includes 869 reactions and 713 metabolites with intracellular and extracellular components. 30.8 % of the genome and a total of 623 genes were associated with the reactions, and 584 of them were associated with COG categories (**Figure 2.1b**). The experimental growth rates we collected are 0.751 ± 0.03 h^{-1} with glycerol supplementation and 0.623 ± 0.04 h^{-1} without glycerol, both are close to the values predicted by the GEM.

2.2. Comparison with other lactic acid bacteria GEMs

Table 2.1| Model characteristics of *i*HL622 and comparison with template GEMs

Model	*i*HL622	*i*NF517	LbReuteri	*i*BT721	*i*ML1515
Organism	*L. reuteri* ATCC PTA 6475	*L.lactis* MG1363	*L. reuteri* JCM 1112	*L. plantarum* WCFS1	*E. coli* MG1655
Genes	2,019	2,339	1,943	3,063	4,243
Included	622(31%)	516(22%)	530(27%)	724(24%)	1,516(36%)
Reactions	869	754	714	778	2,712
Common with iHL622	869	483	531	392	509
With GPR[a]	709(82%)	541(72%)	606(85%)	528(68%)	2,266(86%)
Internal	644	530	507	538	1,548
Transport	122	119	123	127	833
Exchange	103	105	84	113	331
Metabolites	713	650	660	662	1,877
Unique	605	545	561	549	1,071
Biomass consistency	1.00	0.83	-[b]	-[b]	1.00
MEMOTE Score	**80%**	60%	57%	38%	68%

a Gene-Protein-Reaction Associations
b Not applicable

After GEM reconstruction, we compared our GEM with other published other lactic acid bacteria GEMs (and template models). As shown in **Table 2.1** and **Figure 2.2a**, there are 392 to 531 common reactions and 155 unique reactions in *i*HL622 (our GEM). In addition, *i*HL622 contained 31% more genes than the other three lactic acid bacteria templates models and up to 82% reactions in *i*HL622 associated with enzymes and genes. To evaluate model quality, MEMOTE was used for quality control. MEMOTE reports that *i*HL622 has the highest quality scores compared to other GEMs. In addition, *i*HL622 was used to predict the growth capability of *L. reuteri* ATCC PTA 6475 using amino acid as nitrogen sources (**Figure 2.2c**). Previous studies have shown that *L. reuteri* strains have the capacities to synthesize lactate, acetate, ethanol (Spinler et al., 2008), histamine, folate (Santos et al., 2008), cobalamin (Santos et al., 2007, 2008, 2009), 1-propanol (Siebert & Wendisch, 2015; Sriramulu et al., 2008; Walther & François, 2016) and 1,3-propanediol (Mishra et al., 2012), which may be related to the

probiotic effects of *L. reuteri*. As shown in **Figure 2.2b**, we explored the *i*HL622's ability to synthesize these products and production of all these metabolites can be correctly predicted by *i*HL622.

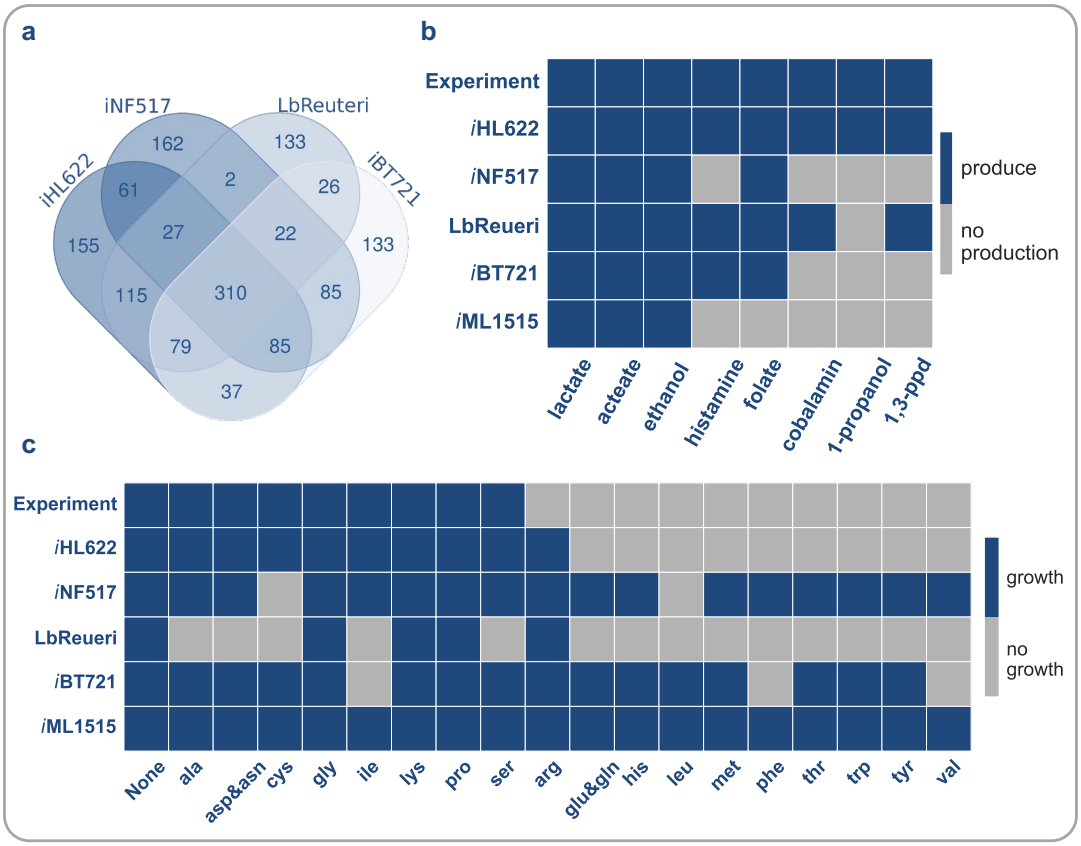


Figure 2.2| **Comparison with template GEMs and evaluation with experimental data. a**) The venn diagram of common and unique reactions in the four lactic acid bacterium models. *i*HL622 is the GEM of *L. reuteri* ATCC PTA 6475 in this study, *i*NF517, LbReuteri and *i*BT721 are the GEMs of *L. lactis* MG1363, *L. reuteri* JCM 1112 and *L. plantarum* WCFS1 separately. **b**) The predictions of representative metabolites. Eight products of lactate, acetate, ethanol, cobalamin, histamine, 1-propanol, folate and 1,3-propanediol) were simulated. The first row shows the experimental data of *L. reuteri*, and the remaining rows show the model results. The produce is shown in dark blue, and no productions are shown in grey. **c)** Growth capability under amino acid omitted medium. The first row shows the experimental data of *L. reuteri*, and the remaining rows show the model results. Growth is shown in blue, and no growth is shown in grey.

2.3. Characteristics of core- and pan-GEMs of 35 *L. reuteri* strains from different hosts

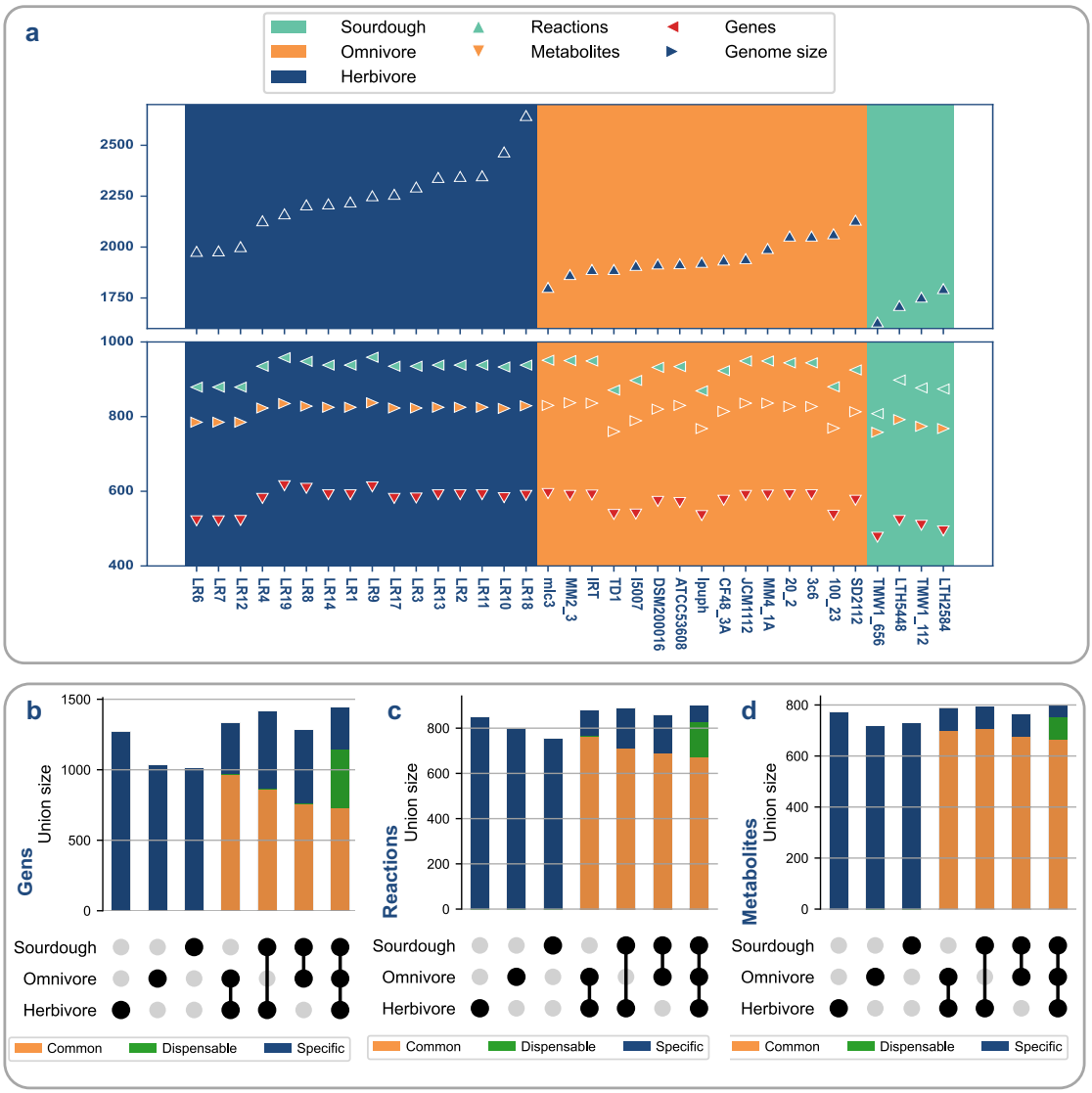


Figure 2.3| GEMs characteristics comparison of 35 *L. reuteri* strains. a) GEMs characteristics and Genome size. These 35 *L. reuteri* strains are isolated from three groups of hosts: herbivore, omnivore, and sourdough. Strains were sorted in descending order of genome size in each group. **b,c,d)** Upset plot of genes, reactions and metabolites between three groups. The total height of the bar indicates the union size of the corresponding group in the horizontal coordinate. In the final bar plotted, specific is considered as only appearing in one group, common is considered as appearing in all groups and appearing in two (between one and all) groups is considered dispensable. The sizes of common, dispensable and specific from all combinations were plotted.

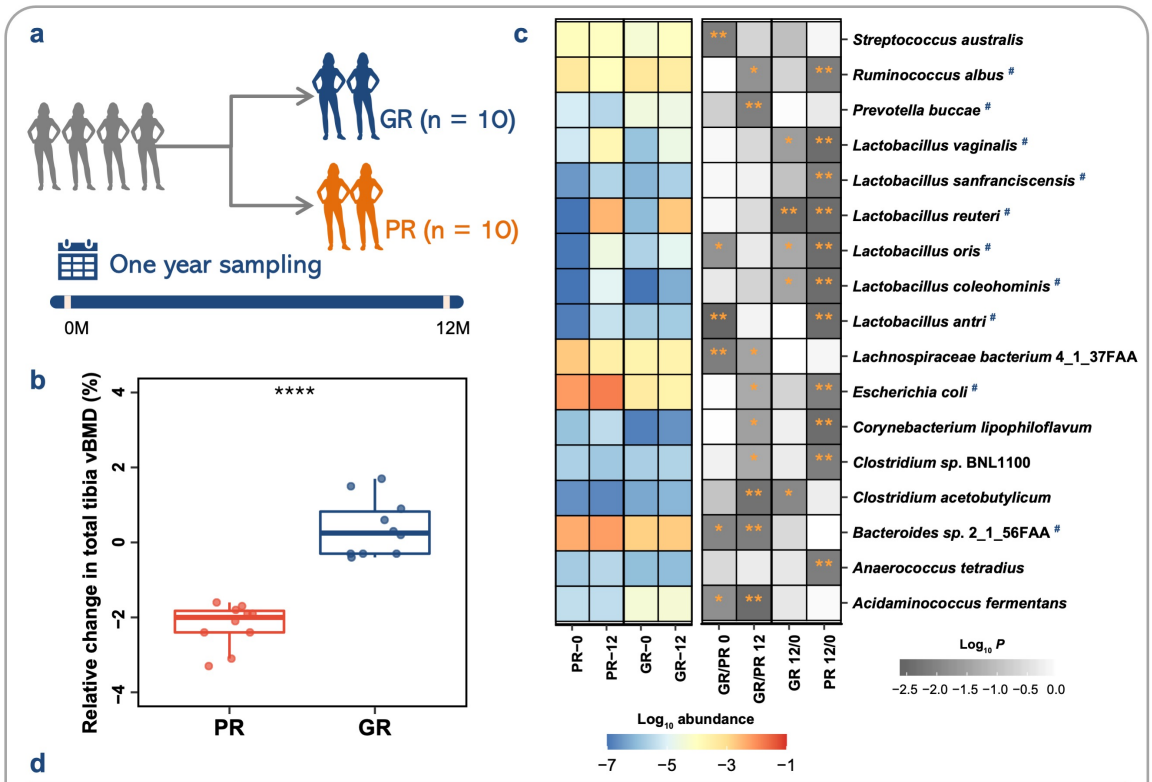
Previous studies show that *L. reuteri* strains have metabolic diversity at the genome level and *L. reuteri* species from different ecological origins are closely associated with their living environment and genomic diversity (Oh et al., 2010; Yu et al., 2018). Considering the strain-specific effects and genetic diversity of *L. reuteri* strains, we were interested to study the metabolic versatility of these strains. Here we analyzed metabolism diversity of *L. reuteri* by GEMs. The genome sequences of 35 *L. reuteri* strains used for GEMs reconstruction were

collected from NCBI. Based on their corresponding host, these 35 strains can be classified into three distinct groups: herbivore, omnivore, and sourdough. The distribution is 16, 15, and four strains into the three groups respectively. The genome size and GEM characteristics are shown in **Figure 2.3a**. The genome size is 2058.3 ± 222.9 CDS, GEMs reactions counts is 919.8 ± 35.0 and metabolites counts is 811.0 ± 25.7. Reactions are linked with 567.1 ± 35.6 encoding genes in GEMs. The results show that the GEMs size is weakly correlated with genome size, the genome size is sorted in descending order while none of the model characteristics correspond to this order (**Figure 2.3a**). We further compared the differences in reactions and metabolites between the groups. The comparison shows that only appearing in one group is considered specific, appearing in all groups is considered common and appearing in two (between one and all) groups is considered dispensable. Previous comparative genomic analysis shown in **Figure 2.3b** and model comparison shown **Figure 2.3c** and **2.3d**. There are host-specific genes in different groups. In our GEMs, there are 7.8% specific reactions and 5.5% specific metabolites correspondingly 74.8% common reactions and 83.7% common metabolites. The percentage of specific genes is more than specific model reactions and metabolites, while the common percentage is opposite, low correlation suggests that many of the differences in the genome are not inherited to GEMs.

2.4. The effects of *L. reuteri* ATCC PTA 6475 intake on older women with low bone mineral density

After comparing different *L. reuteri* strains GEMs, we studied the effects of *L. reuteri* ATCC PTA 6475 intake on older women with low bone mineral density. In this study, 20 elderly women were selected to receive a one-year supplement of *L. reuteri* ATCC PTA 647 (**Figure 2.4**). ten women with a good response (GR group) and ten women with a poor response (PR group). Serum and fecal samples were collected from older women at baseline and 12 months. After one year, the results show that *L. reuteri* ATCC PTA 6475 supplementation can potentially prevent a deterioration of the gut microbiota and inflammatory status in elderly women with low bone mineral density. As shown in **Figure 2.4d,** probiotic intake prevents bone loss in the GR group: the relative change in the total volumetric bone mineral density (vBMD) of the tibia showed an increase in the GR group (0.39 ± 0.77) compared to the PR group (-2.22 ± 0.58; $P < 0.001$ by the t-test) and a decrease in the level of total volumetric BMD of the tibia was observed in the PR group at 12 months ($P < 0.05$). The ultrasensitive c-reactive protein (usCRP) showed a significantly reduced level in the GR group at 12 months ($P < 0.05$). Furthermore, three species are shown to be differential at baseline, including *Streptococcus australis*, *Lactobacillus antri*, and the *Lachnospiraceae bacterium* 4_1_37FAA. Four species were identified to be differential at 12 months, including *Prevotella buccae, Clostridium acetobutylicum, Bacteroides sp.* 2_1_56FAA, and *Acidaminococcus fermentans*. The differences in the endogenous baseline microbiota might be important for a good response to the probiotic intake. The results suggest that *L. reuteri* ATCC PTA 6475 supplementation has the potential to prevent a deterioration of the gut microbiota and

inflammatory status in elderly women with low bone mineral density, which might have beneficial effects on bone metabolism.

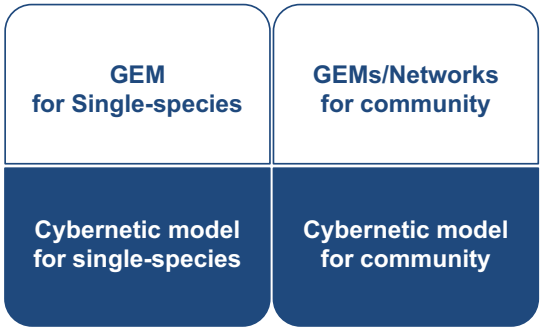


d

Characteristics	Baseline		12 months		P^a	P^b
	GR (n = 10)	PR (n = 10)	GR (n = 10)	PR (n = 10)		
Total tibia vBMD (mg/cm^3)	247±39.2	231±44.9	248±39.8	**226±44.2#**	0.42	0.26
usCRP (mg/L)	2.14 (1.53-3.68)	0.98 (0.8-2.47)	**1.57 (1.13-1.90)***	1.36 (0.67-3.19)	0.25	0.91
Total fat mass	28.9 (27.1-32.0)	20.8 (19.3-24.5)	28.2 (24.3-31.7)	20.0 (18.5-23.7)	**0.04**	**0.04**

Figure 2.4| The experimental design and alterations in the gut microbial composition. **a)** Women with a good response (GR group, n=10) and with a poor response (PR group, n=10) were selected. Serum samples and fecal samples were collected from the older women at baseline and 12 months. **b)** Relative change in total tibial volume BMD in the GR and PR groups after one year of treatment. The box plot shows the upper quartile, median, and lower quartile. **c)** The heat map on the left shows a logarithmic transformation of the mean abundance of the different species in the GR and PR groups at baseline and at 12 months. The gray color on the right heatmap indicates the P value of the comparative analysis; '*' indicates P < 0.05; '**' indicates P < 0.01. **d)** Comparison of characteristics of the GR and PR groups at baseline and 12 months. Note: Mean ± SD. Nonnormally distributed variables are presented as medians with an interquartile range. The t-test or Wilcoxon test were used as appropriate. '*' and '#' indicates a significant difference (P < 0.05) between baseline and 12 months in the GR and PR groups, respectively. P a and P b values are from comparisons between the GR and PR groups at baseline and 12 months, respectively. The significant differences (P < 0.05) are highlighted in bold. vBMD: volumetric bone mineral density; usCRP: ultrasensitive c-reactive protein.

3. Part II (Paper III): Cybernetic modeling for microbial communities



Related Papers:

Paper III: Modeling the metabolic dynamics at the genome-scale by optimized yield analysis

To simulate dynamic systems, we performed the cybernetic model approach, which integrates enzyme synthesis and activity regulation. Cybernetic models have been widely applied in bioreaction engineering. The advantage of cybernetic models is that they do not require a lot of experimental data, which is important for gut microbiota modeling. However, we also found some limitations when we tried the traditional hybrid cybernetic modeling (HCM) strategy. In this section, we explain these limitations and present new HCM strategy to simulate metabolic dynamics at the genome-scale.

3.1. Hybrid cybernetic modeling strategy and its limitations

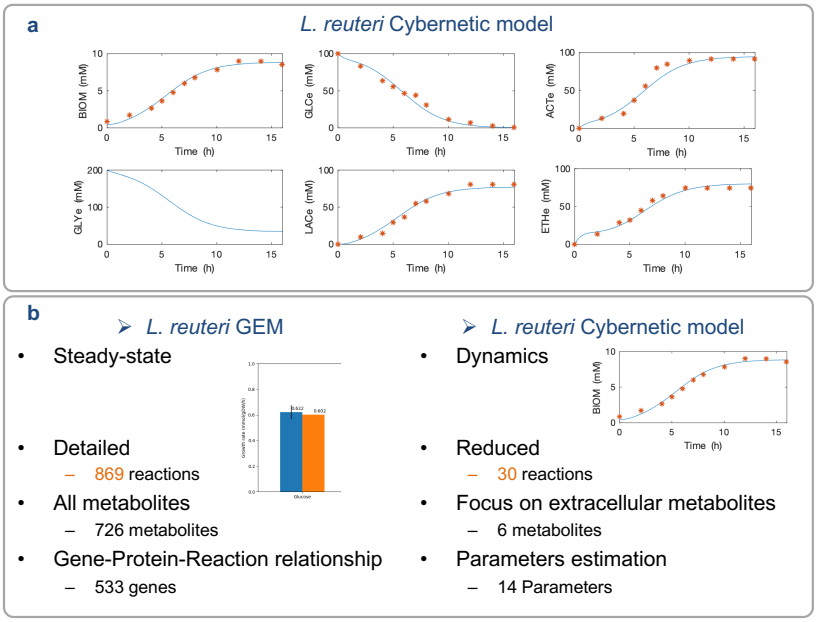


Figure 2.5| The overview of *L. reuteri* cybernetic model and compare with GEMs. **a)** The simulation of *L. reuteri* growth and metabolism with cybernetic modeling approach, the orange dots indicate experimental data, and the blue lines indicate cybernetic model simulation. **b)** Comparison of *L. reuteri* GEM and *L. reuteri* cybernetic model.

Based on the model constructed in the previous section, we constructed a cybernetic model and it performs well (**Figure 2.5a**). As shown in **Figure 2.5b**, the model needs to be simplified into 30 reactions, which means that we lose a lot of information. And the reason for model reduction is to calculate the EFMs and further to obtain the yield space. There are two types of methods to obtain the yield space from GEMs. One is to calculate all feasible pathways in a network like elementary flux modes (EFMs) and elementary flux vectors (EFVs). The other type of method is to identify the boundaries of the yield space by yield optimization. Generally, the second method can save computing sources for complex models because of few pathways for calculation and focus on the boundaries of the yield space.

Due to the high computational demand for calculating EFMs, applying the HCM approach on conventional genome-scale metabolic models is still a challenge. The number of EFMs is exponentially growing with the number of reactions in a network, and this is therefore a bottleneck for applying the HCM approach at the genome-scale directly (Vilkhovoy et al., 2016). According to the metabolic yield analysis (MYA) by Song and Ramkrishna, EFMs provide a master yield space and convex hull for pathways selection, and selected pathways' yield values are located on the boundaries of the yield space (Song & Ramkrishna, 2009). Therefore, to fill the gap between GEMs and HCMs without EFMs, alternative approaches should be able to obtain yield spaces from GEMs. Some studies have attempted to replace the EFMs in the HCM approach by EFVs or FBA modes (Ahamed et al., 2021; Vilkhovoy et al., 2016). EFVs are alternatives to EFMs but can account for inhomogeneous constraints (Klamt et al., 2017, 2018; Müller & Regensburger, 2016). Under inhomogeneous constraints, the EFVs can provide a theoretically correct yield space that is better than EFM. Although EFVs are more reliable for computing the yield space, the computational demand is still high (Ahamed et al., 2021). FBA modes cannot provide a complete yield space because they optimize the output rate of target metabolites, not yield (Klamt et al., 2018; Vilkhovoy et al., 2016). Therefore, algorithms for yield optimization need to be developed.

3.2. Opt-yield-FBA and HCM strategy

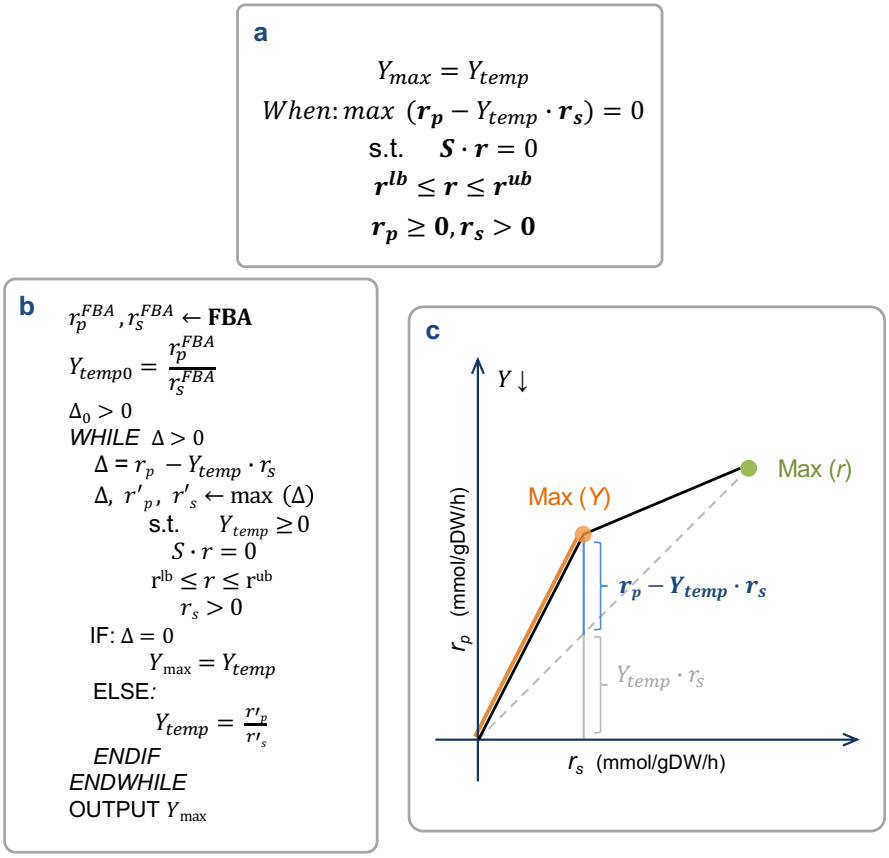


Figure 2.6| Opt-yield FBA mathematical description and pseudo-codes and illustrations. **a)** The objective function is a linear form with a temp value of Y_{temp}. The initial $Y_{temp} = Y_{FBA}$ and the object is looking for a larger Y_{temp} until finding the maximum yield. All the processes subject to steady-state ($\boldsymbol{S \cdot r = 0}$) capacity and irreversibility ($r^{lb} \le r \le r^{ub}$), and the substrate should be absorbed ($r_s > 0$). **b)** pseudo-codes of opt-yield FBA. c) The approximate solution of the optimal yield. Objective functions of $\boldsymbol{r_p - Y_{temp} \cdot r_s}$ meaning for yield optimization.

To obtain the yield space from GEMs, we developed a method to optimize the yield. The optimal yield solution is calculated by a series of iterations of the FBA. In each iteration, an assumed yield is defined as Y_{temp} and the optimized objective function ($r_p - Y_{temp} \cdot r_s$) is set as the differences between r_p and the product of Y_{temp} and r_s. The initial value of Y_{temp} is Y_{FBA} , which is the yield value when rate optimized by FBA. The return of each iteration is a vector of flux distribution that updates Y_{temp} for the next iteration. The iterations will be terminated when the optimization return zero and the maximum yield is found. In **Figure 2.6c**, is the relationships of the substrate uptake rate (r_s) and the production rate (r_p). The maximal r_p value from FBA is shown as a green dot and the maximal yield is shown with an orange dot or line. The max(r) and max(y) points do not overlap and the maximal Y requires multi optimal calculations. The objective function of $r_p - Y_{temp} \cdot r_s$ is shown by the blue line and text.

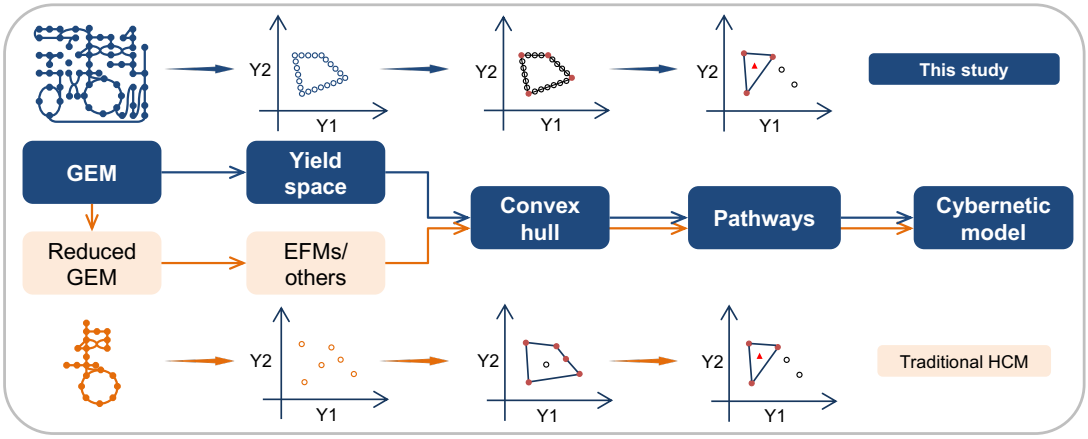


Figure 2.7| The HCM strategy and comparison with the traditional HCM. The HCM strategy in this study (top part) and the traditional HCM strategy (bottom part). In our HCM strategy, the yield space is calculated by the opt-yield-FBA from a complete GEM. All yield values of pathways from opt-yield-FBA are located on the boundaries of the yield space. The traditional HCM strategy processes small or medium size models and calculates EFMs to obtain the yield space. The yield values of EFMs are located both on the boundaries of space and inside of the yield space. The pathways from opt-yield-FBA or EFMs provide the master yield space for pathways selection by MYA and convex hull. Experimental data can be used to further select the active pathway. The HCM method can be applied based on the selected pathways.

Based on opt-yield-FBA, the new HCM strategy can avoid the calculation difficulties associated with identifying EFMs. By maximizing and minimizing a single target product, a range of yield values can be determined. For pairs of target products, the yields space can also be calculated by sampling. As shown in **Figure 2.7**, after obtaining the maximum and minimum yield values of Y1 ($Y1_{max}$, $Y1_{min}$, X-axis), any values between $Y1_{min}$ and $Y1_{max}$ can be sampled as constraints to optimize the Y2 (Y-axis). After calculating the maximum and minimum yield values of Y2 in different sampled intervals from $Y1_{max}$ and $Y1_{min}$, the yield space of Y1 and Y2 can be obtained. In comparison to EFM or EFV, opt-yield-FBA allows direct access to yield space. Since FBA usually processes the model through linear programming, the opt-yield-FBA algorithm can be implemented by the FBA framework and is compatible with most current modeling tools or solvers such as COBRApy. With the replacement of EFMs by opt-yield-FBA, the HCM strategy makes it possible to process complex GEMs directly and apply the HCMs approach to describe metabolic dynamic GEMs.

3.3. The application of HCB at the small scale

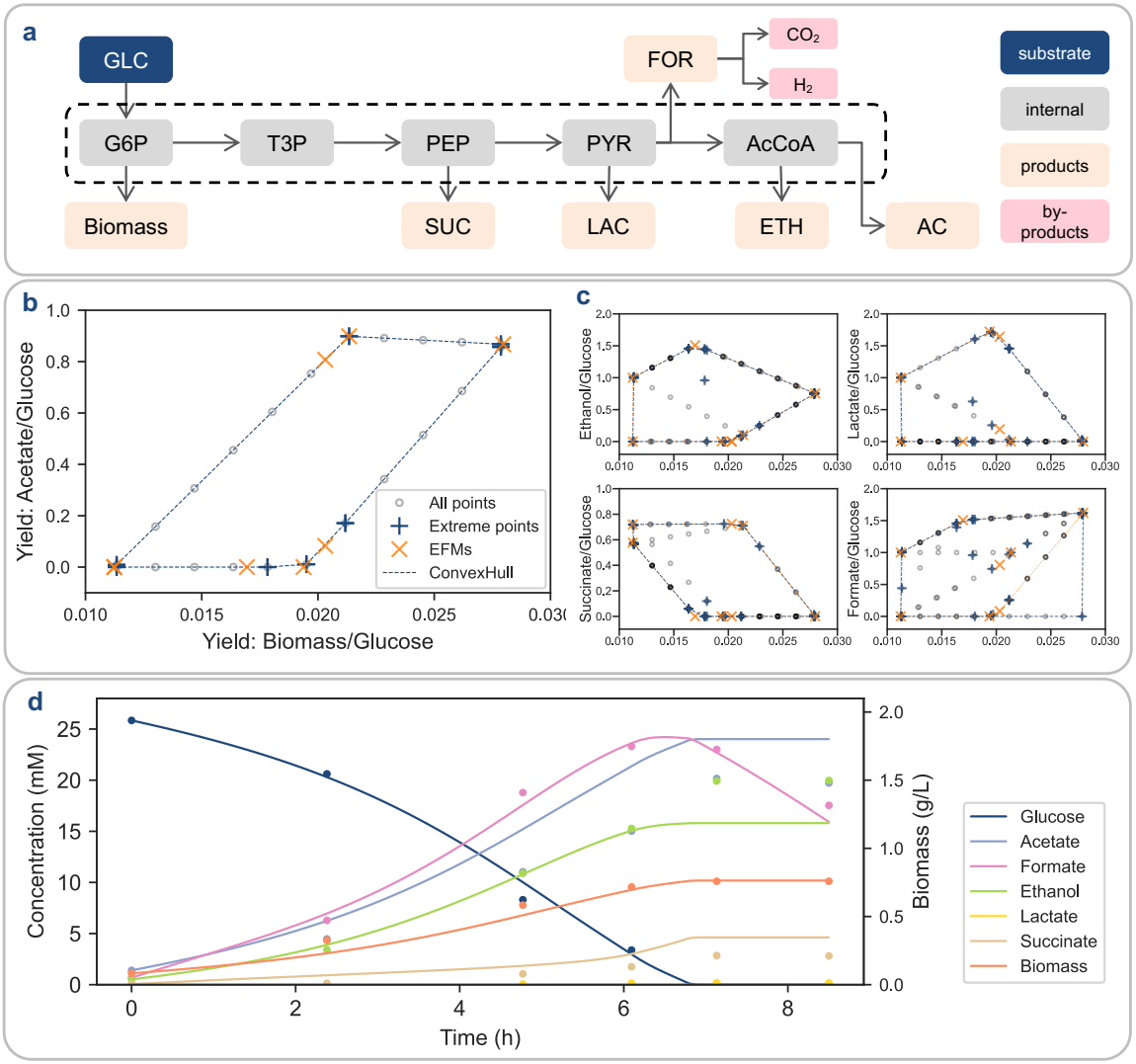


Figure 2.8| The HCM strategy applied on a reduced *E. coli* metabolic network. **a)** Overview of the receded *E. coli* network, which contains 12 reactions and 14 metabolites. GLC (glucose), G6P (glucose 6-phosphate), T3P (fructose 1,6-biphosphate), PEP (phosphoenolpyruvate), PYR (pyruvate), AcCoA (acetyl coenzyme-A), SUC (succinate), FOR (formate), ACT (acetate), LAC (lactate), ETH (ethanol), B (biomass), CO_2 and H_2. **b)** The yield distributions are calculated by EFMs and opt-yield-FBA for acetate, ethanol, lactate, succinate, and formate. The grey dot indicates the pathway generated by opt-yield-FBA, most of them located on the boundaries. The blue '+' marker indicates the extreme point of opt-yield FBA pathways, and the dashed line indicates the convex hull of the yield space. The orange 'x' marker indicates the pathway by EFMs. **c)** Simulation of biomass and metabolic dynamics under anaerobic conditions. The dot indicates experimental data, and the lines are obtained from the simulation.

We firstly applied the HCM strategy on a reduced *E. coli* metabolic network under anaerobic conditions (J. Kim et al., 2008). This small-scale network contains 12 internal reactions and 19 metabolites (J. Kim et al., 2008). The main metabolic pathways under anaerobic conditions are shown in **Figure 2.8a**, which include one substrate (glucose), five internal metabolites, and eight products. The cybernetic model simulates seven external metabolites, including glucose, succinate, formate, acetate, lactate, ethanol, and biomass. The first step in the strategy is to calculate the yield space. The biomass yield is defined as the flux of biomass

divided by the flux of glucose consumption, written as $Y_{biomass} = {}^{r_{biomass}}/_{r_{\mathrm{glucose}}}$. The biomass yield range was calculated by opt-yield-FBA. For each biomass yield interval value, we set the biomass yield as a constraint of the model, and search for the minimum and maximum acetate yield (Y_{ac}.). We obtained 22 pathways for each target metabolite and selected 38 pathways at extreme points to cover the five complete yield spaces. In comparison with the yield space of EFM, the yield space of opt-yield-FBA has the same shape as the yield space of EFM. The parameters were estimated by the least squares method and experimental data. Finally, the cybernetic model simulation shows a good agreement with the experimental observations.

3.4. The performance is robust for the medium scale network

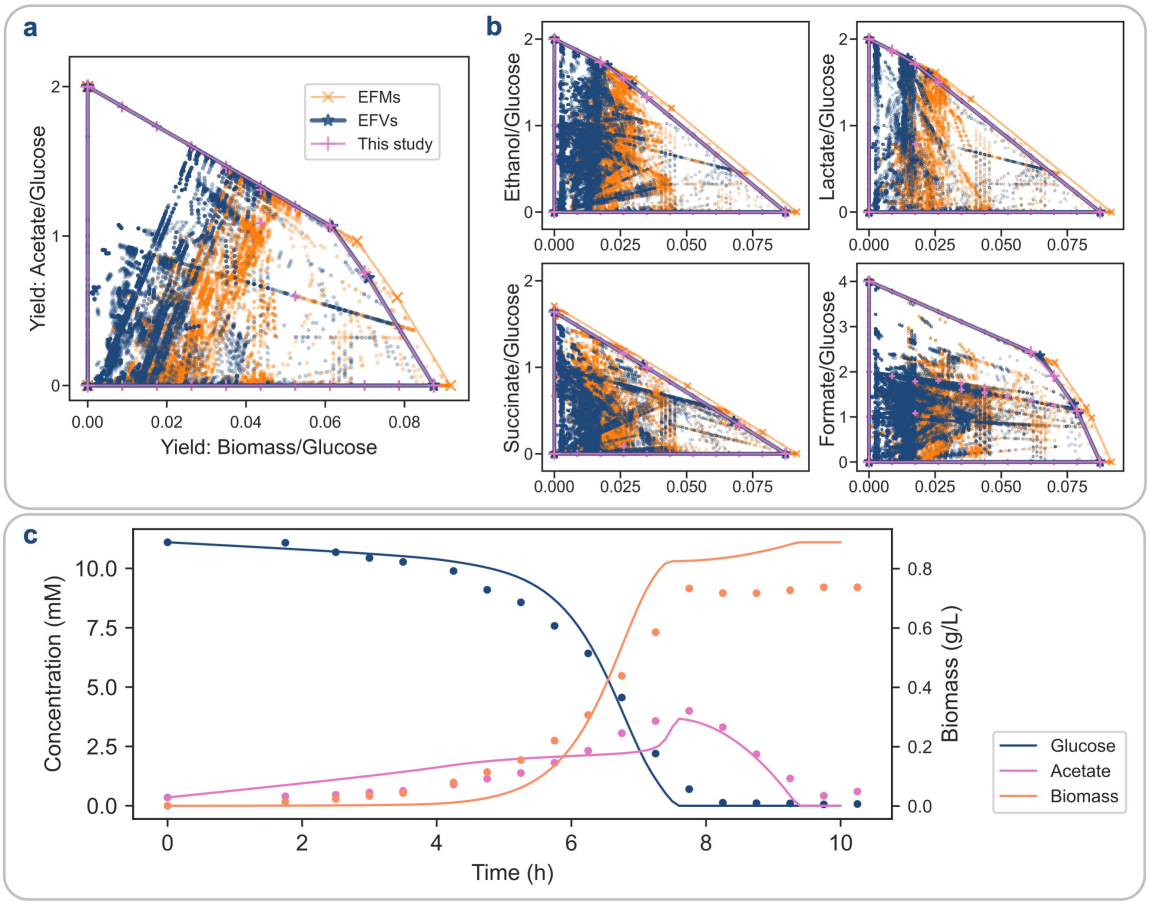


Figure 2.9| The HCM strategy applied on a *E. coli* core metabolic model. a, b) Yield spaces of acetate, ethanol, lactate, succinate, and formate. Calculated by EFMs, EFVs, and opt-yield-FBA. The X-axis indicates the biomass yields, and the Y-axis indicates the product yields. The dots and markers indicate the distribution of the yield values, and the line indicates the convex hull in a yield space. The colors orange, blue, and pink correspond to the EFM, EFV, and opt-yield-FBA respectively. The yield space of the EFVs is considered a theoretically correct result. The yield space from opt-yield-FBA is the same as that of EFVs. EFMs have a larger yield space because inhomogeneous constraints in the model are not considered. **c)** The simulation of fermentation. The dots denote the experimental observations, while the lines denote the simulation of the cybernetic model.

To compare the yield space of opt-yield-FBA and other methods. The *E. coli* core GEM under aerobic conditions was calculated. The GEM containing 95 reactions and 60 metabolites (Orth et al., 2010) and calculated 100,273 EFMs and 95,106 EFVs. We also calculated the pairwise yield space of five metabolites and biomass by opt-yield-FBA, and 110 opt-yield-FBA pathways were generated. The *E. coli* core model contains many such constraints like minimal maintain energy bounds and the FEVs could consider inhomogeneous constraints better than EFMs. Therefore, the yield space from EFVs is considered as a theoretically correct result which is smaller than that from EFMs. The difference between the EFMs yield space (orange lines in **Figure 2.9 a** and **b**) and the EFVs yield space (blue lines) may be related to the inhomogeneous constraints. The pathways from opt-yield-FBA are significantly less than the other two methods and are mainly located at the boundaries of the yield space. The yield space of opt-yield-FBA is almost the same as the EFVs' with 99.888% $\pm$ 0.098% and 100% similarity in the convex hull area and parameters. The results show that the yield space of opt-yield-FBA is similar to the correct yield space of EFV. There are 110 pathways from opt-yield-FBA could represent the same yield space of 95,106 EFVs and a more accurate yield space of 100,273 EFMs. Finally, we identified five active pathways using experimental data to construct a cybernetic model to simulate metabolic dynamics under aerobic conditions (Varma & Palsson, 1994).

3.5. The metabolic dynamics prediction at the genome scale

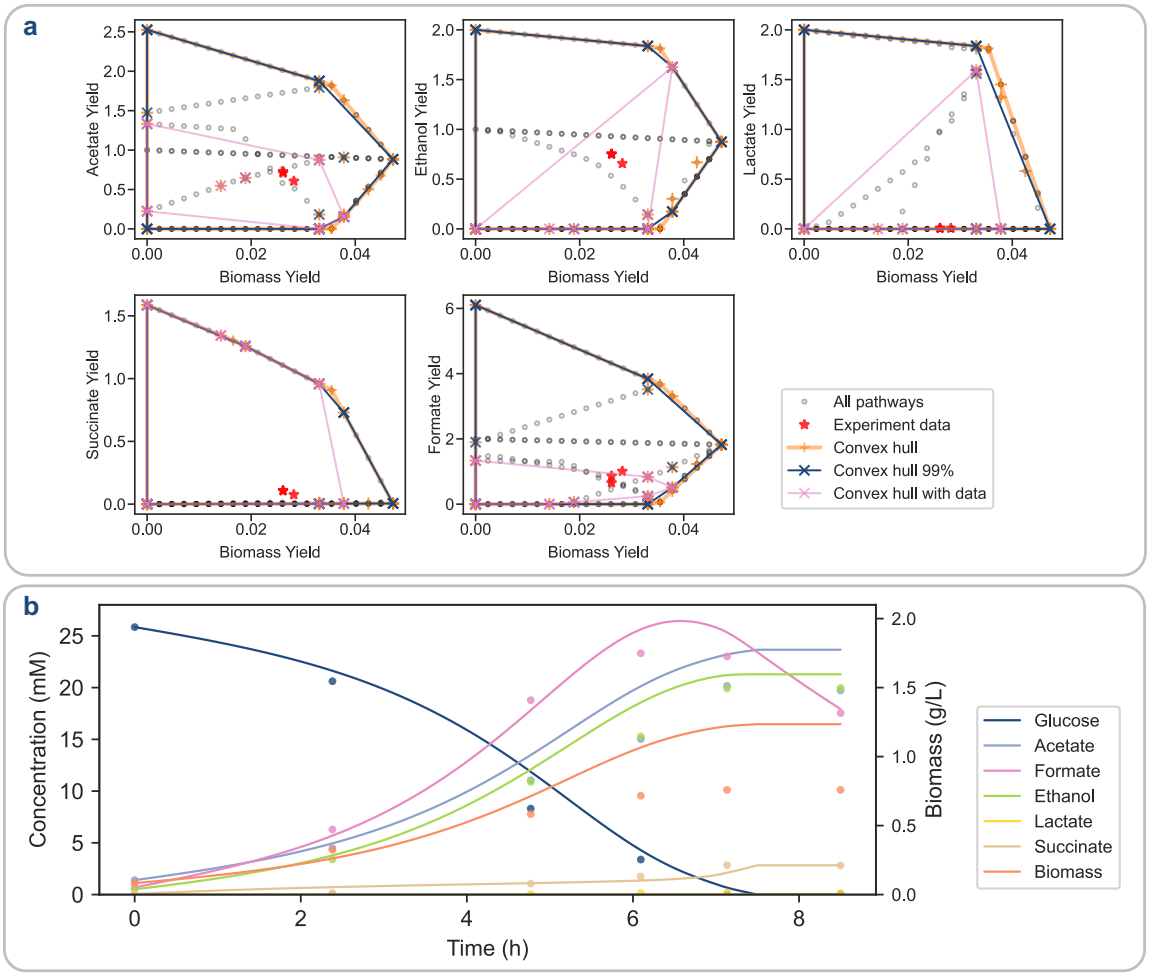


Figure 2.10| Cybernetic modeling for the *E. coli* *i*ML1515 model. a) The yield space of acetate, ethanol, lactate, succinate, and formate. The X-axis indicates the biomass yields, and the Y-axis indicates the product yields. The dots and markers indicate the yield distribution of the pathways, and the lines indicate the yield space. The grey dots are all pathways generated by opt-yield-FBA and the red stars indicate experimental data. The orange and blue lines are convex hulls which covering complete and 99% yield spaces. The pink line indicates the convex hull of active pathways identified by MYA with experimental data. **b)** The cybernetic model simulation is under anaerobic conditions. The dots indicate experimental observations and the lines indicate simulations of cybernetic models.

To evaluate the performance in full-size GEMs, we applied opt-yield-FBA on *E. coli* *i*ML1515, which contains 1,515 genes, 2,719 reactions, and 1,192 metabolites (Monk et al., 2017). In this case, the calculation of EFMs or EFVs is challenging due to the large computational demand. The opt-yield-FBA is an FBA based method and can be used to calculate *i*ML1515 yield space of succinate, formate, acetate, lactate, and ethanol (**Figure 2.10a**). There are 220 pathways generated in total. A following convex hull yield analysis was used to reduce the number of pathways and 37 pathways were selected to represent the five pairwise yield spaces. We also calculated a set of pathways that covers 99% of the volume of the original convex hull (orange line and * in **Figure 2.10a**). Based on the experimental data (red stars in **Figure 2.10a**) (J. Kim et al., 2008), we identified the final active pathways by MYA (red lines

and 'x' in **Figure 2.10a**), these are a minimal set of pathways that could cover the experimental yield. We collected the experimental data of six metabolites and biomass concentrations under anaerobic conditions. After estimation of the parameters, the cybernetic model was used to simulate dynamic conditions as shown in **Figure 2.10b**. All metabolites concentration simulations agree with experimental data, especially formate degradation.

3.6. The simulation of multi- species interactions

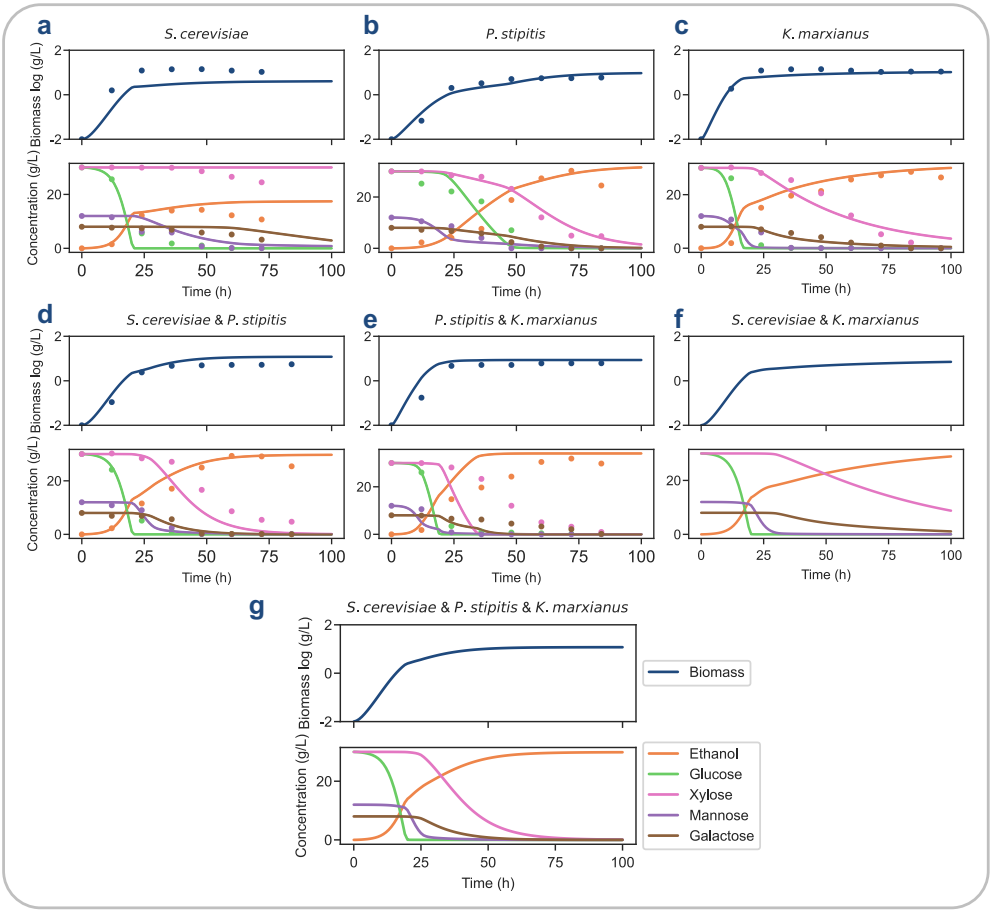


Figure 2.11| HCM strategy for three yeast species interactions. a-c) The simulation of monoculture including *Saccharomyces cerevisiae, Pichia stipites* and *Kluyveromyces marxianus*. Both the biomass (log g/L) and metabolites concentrations (g/L) were shown. **d)** Simulation of two species coculture of *S. cerevisiae* and *P. stipites.* **e)** The simulation of two species co-culture of *P. stipites* and *K. marxianus.* **f)** The prediction of three species co-culture. The dots denote the experimental observations, while the lines denote results generated by the cybernetic model simulation.

To evaluate our modeling approach in multi-species fermentation, we tested multi-species models to test whether our method can be used to simulate the dynamics of microbial consortia. The first case includes a system consisting of three yeast, *Saccharomyces cerevisiae, Pichia stipites,* and *Kluyveromyces marxianus* (Geng et al., 2012). This simple system was simulated by a reduced model by the traditional HCM, and we performed our HCM strategy with a complete yeast GEM (Geng et al., 2012; Lu et al., 2019). These three yeast species have similar fermentation characteristics and most interspecific relationships

are competitive to absorb the same carbon source (Rouhollah et al., 2007). For simplicity, we used the same GEM for the three species, and the GEM were adopted from the Yeast8 model, which contains 3,989 reactions and 2,693metabolites (Lu et al., 2019). 88 pathways were first identified to provide the master yield spaces and after the convex hull yield analysis, 17 pathways were selected to cover 99% of the master yield space. We identified different sets of active pathways for the three species using MYA with different fermentation data. For *S. cerevisiae* mono-culture, nine active pathways are selected to cover the experimental yield space; for *P. stipites* 10 active pathways were selected, and 10 active pathways were selected for *K. marxianus*. We estimated a set of the parameters that could simulate the mono- culture and two species co-culture. Next the same parameters were used to predict three-species co-cultures. Because the experimental date only observed the total biomass, for the culture of muti species, the biomass of different species was assumed to be the same.

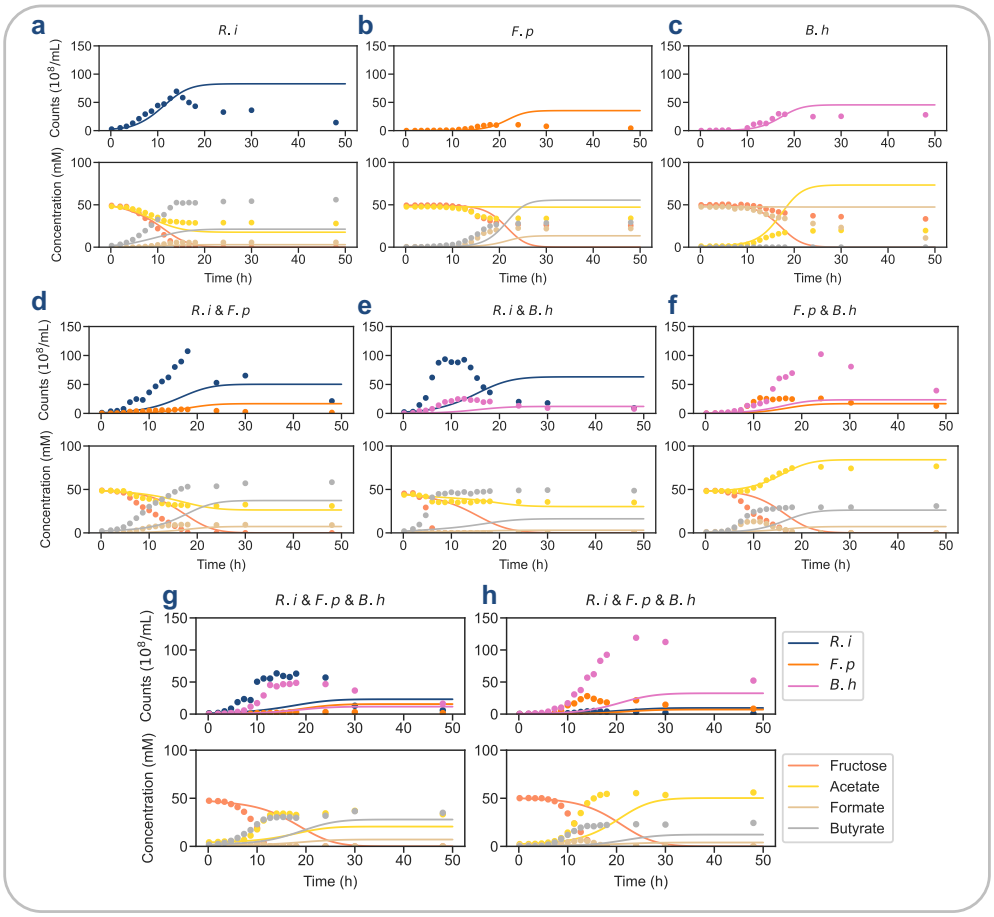


Figure 2.12| The simulation of human gut microbial consortia with three members. a-c) The simulation of single species mono-culture of *Roseburia intestinalis (R.i), Faecalibacterium prausnitzii (F.p)*, and *Blautia hydrogenotrophica (B.h)* **d-f)** The simulation of two species co-cultures. **g)** The simulation of three species co-cultures. The dots denote the experimental measurements, while the lines denote model simulations. Both the cells count (upper part), and metabolites concentrations (mM, lower part) are shown.

To validate the ability of our strategy to simulate gut microbial dynamics, a synthetic community of three species was selected (D'hoe et al., 2018). These three species co-cultures including representative gut microbial species *Roseburia intestinalis* L1-82, *Faecalibacterium*

prausnitzii A2-165 and *Blautia hydrogenotrophica* S5a33 (D'hoe et al., 2018). All three species are capable of using fructose as the main carbon source. *B. hydrogenotrophica* is an acetate producer and formate consumer, while *R. intestinalis* and *F. prausnitzii* are acetate consumers. The *B. hydrogenotrophica* GEM contains 1,719 reactions and 1,195 metabolites; the *F. prausnitzii* GEM contains 1,358 reactions and 989 metabolites; the *R. intestinalis* GEM contains 1,724 reactions and 1,215 metabolites. Because the three species have different substrate usage, the production yields were generated based on different substrates and different rates. All the yield calculations and pathways selection considered fructose as the carbon source, and acetate or formate as an alternative carbon source accordingly. After analyzing the convex hull yield, we finally selected 4, 5, and 5 active pathways for *B. hydrogenotrophica, R. intestinalis,* and *F. prausnitzii* separately.

First, the model parameters were estimated using mono-culture experimental data, and then applied to predict two species and three species co-cultures. For mono-cultures, shown in **Figure 2.12, the concentration of** all biomass and most metabolites agrees well with the experimental data, but some mismatches were identified for the mono-culture of *F. prausnitzii* (**Figure 2.12b**) where the acetate concentration was not well simulated. For two species co-cultures, shown in **Figure 2.12d-f**, parts of biomass and metabolites concentration agree well with experimental data. The relationship between *R. intestinalis* and *F. prausnitzii* is competition and their dual-culture simulation in **Figure 2.12** was better than the other two dual-culture in **Figure 2.12e** and **Figure 2.12f**. For three species co-cultures, shown in **Figure 2.12g-h**, two different states are obtained, which may be due to dependency on the initial species abundance. Model simulations for different data are basically the same and are not sensitive to the initial abundance of species. Even though the model does not fit quantitatively to all the experimental data, some species and metabolites are matching the experimental data or at least trends for the three species culture simulations such as mono-culture and *R. intestinalis* and *F. prausnitzii* dual-culture. We tried to improve model performance by integrating different experimental data sets to estimate the parameters.

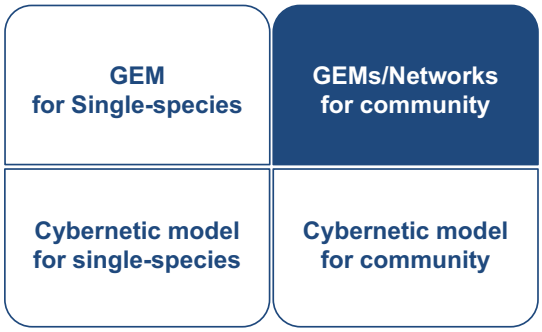
Table 2.2| Number of pathways in different models

GEMs	Reactions	Metabolites	EFM/EFV	Opt-Yield-FBA	Convex hull	Active path/final path
Receded *E. coli*	12	19	8/-	110	38	6
E. coli core	95	60	100,273/ 95,106	110	37	5
*i*ML1515	2,712	1,192	-/-[1]	210	20 (99% convex hull)	8
Yeast *S. cerevisiae*	3989	2693	-/-[1]	88	17 (99%)	9
P. stipitis	3989	2693	-/-[1]	88	17 (99%)	10
K. marxianus	3989	2693	-/-[1]	88	17 (99%)	10
B. hydrogenotrophica	1,719	1,195	-/-[1]	132	14 (99%)	4
F prausnitzii	1,358	989	-/-[1]	132	22 (99%)	5
R. intestinalis	1,724	1,215	-/-[1]	132	16 (99%)	5

[1] -: EFMs/EFVs cannot be calculated.

EFMs and EFVs face challenges when processing large-scale models with many inhomogeneous constraints, requiring extensive computational resources to perform a large number of pathway calculations. The opt-yield-FBA can be an alternative of EFMs in the HCM approach because it can obtain yield space using GEMs with lower computational demands. As shown in **Table 2.2**, opt-yield-FBA calculated 110 pathways for the *E. coli* core model, while EFMs and EFVs calculated ~100,000 pathways.

4. Part III (Paper IV & V): GEMs and metabolic networks for community



Related Papers:

Paper IV: Metagenomic analysis of type 2 diabetes datasets identifies cross-cohort microbial and metabolic signatures

Paper V: The metabolic network inference framework for shotgun metagenomics

Type 2 diabetes mellitus (T2D) is a multi-factor disease and one of the fastest increasing diseases. To identify microbial and metabolic features associated with T2D, we performed a systematic analysis of four published metagenomic studies. A total of 1,779 metagenomic shotgun sequencing (MGS) files of individuals with different glycemic status were analyzed. We construct genome-scale metabolic models (GEMs) for microbial species and corresponding communities. We simulated the metabolic capability of these communities, such as the fluxes of producing SCFAs and amino acids. Using machine learning, several SCFAs producing bacterial species and metabolic reactions were consistently identified to be associated with T2D status across studies.

Due to the complexity of community modeling, we developed the Analyzer for Metabolic Networks (AMN) toolbox for reconstructing and analyzing the metabolic networks form MGS. Unlike GEMs, topological metabolic networks are simpler and suitable for rapid reconstruction to provide the metabolic capacity of samples. The AMN tool is a Python package that helps users quickly reconstruct metabolic networks, analyze the structure and properties of networks.

4.1. GEMs for type 2 diabetes microbiota communities

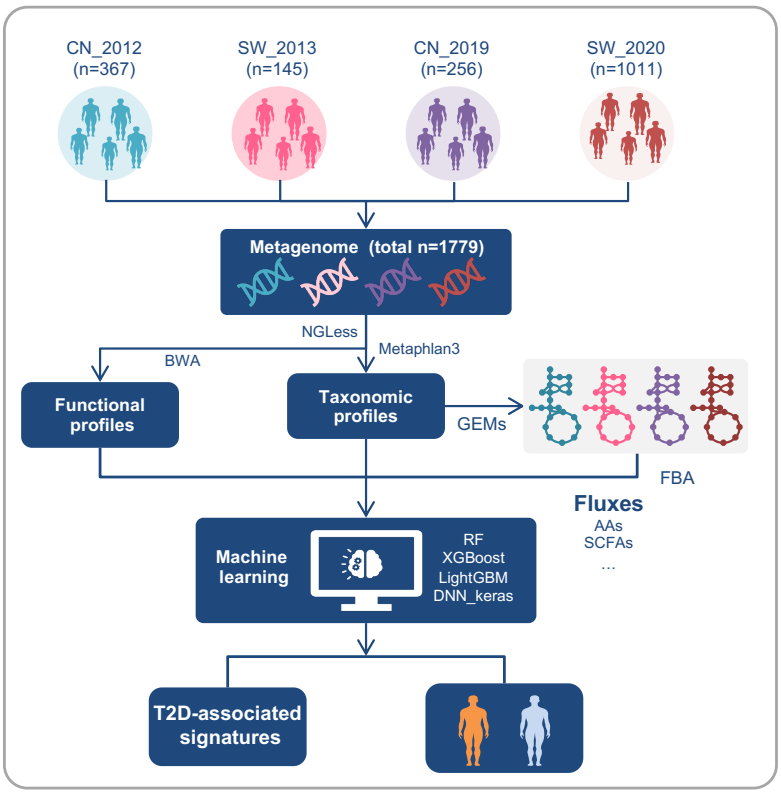


Figure 2.13| The workflow of integrating machine learning with genome-scale metabolic models (GEMs) of the gut microbiota.

The whole workflow of this section is illustrated in **Figure 2.13**. We performed a systematical analysis of the published gut metagenomic data, using machine learning and community level metabolic models. 1,779 MGS datasets from four previous studies were collected, consisting of 848 normal glucose tolerance (NGT), 571 prediabetes (Pre-DM) and 360 individuals with T2D. A standardized bioinformatics pipeline is used to extract microbial features, including functional characteristics and taxonomy. To simulate the metabolic capacity of microbial communities, we constructed 827 GEMs for gut microbial species and 1,779 community-level GEMs for all individuals. Furthermore, we devised different decision tree-based machine learning models to predict the status of T2D, based on the microbial compositional and metabolic features of the gut microbiota as well as potential confounding factors. By using machine learning integrated with community level metabolic models, we identified a number of bacterial species and metabolic reactions that produce SCFAs. The importance of bacterial species and metabolic responses in predicting T2D status has kept consistent across studies. However, classification models based on gut microbial features have limited performance in distinguishing T2D from NGT in independent cohorts, suggesting that specific microbial-based models should be developed for certain types of cohorts.

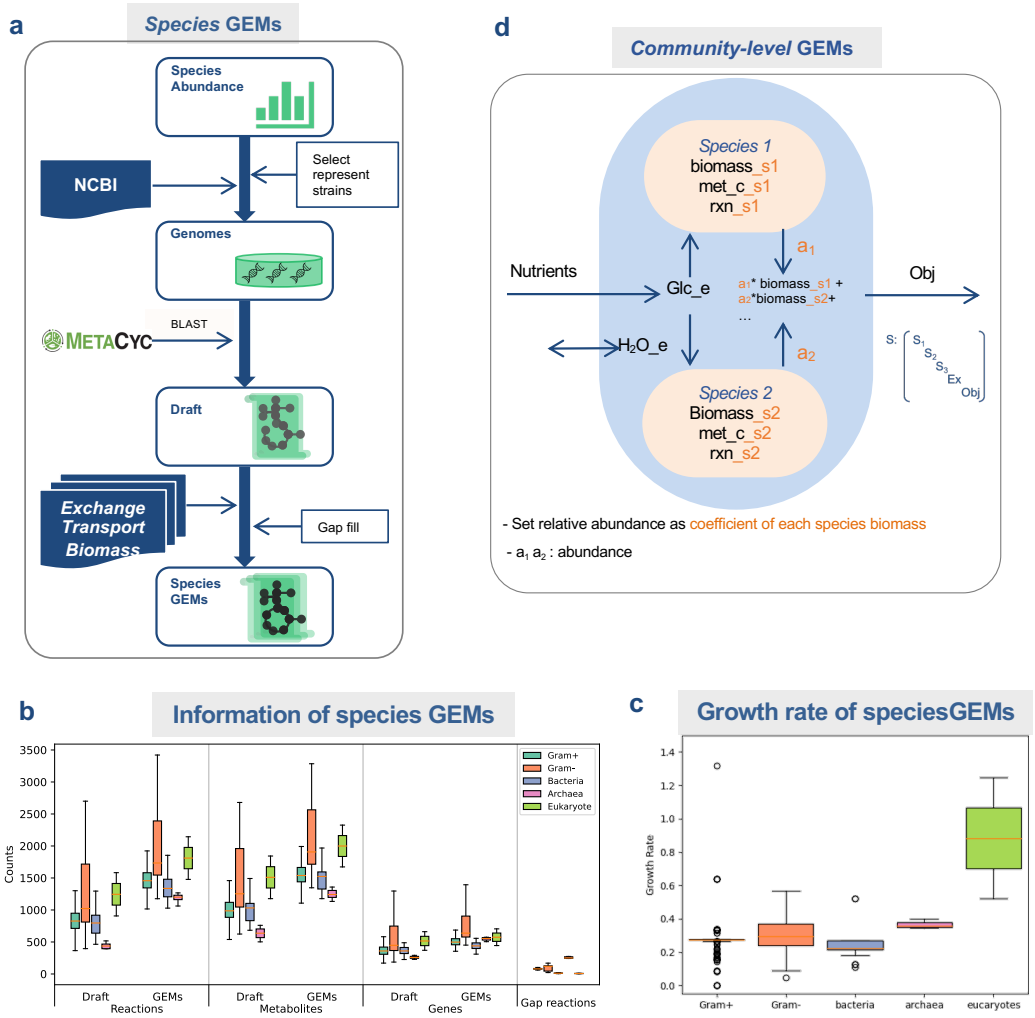


Figure 2.14| Reconstruction and characteristics of GEMs. a) The species GEM reconstruction pipeline based on the MetaCyc database. **b)** Characteristics of species GEMs, counts distributions of reactions, metabolites, genes, and gaps reactions of species GEMs. All species are divided into five categories, gram-positive, gram-negative, bacteria that missing gram information, archaea and eukaryotes. **c)** Growth rate of species GEMs, Most GEMs have completed metabolic functions. **d)** The pipeline for the construction of community GEMs by creating different compartments to separate intercellular metabolites and allow extracellular metabolites to be transported between species. The overall biomass μ was set as the weighted combination of the biomasses μ_i of all species GEMs. The relative abundances of species a_i were used as coefficients for the biomass of each species. For the simulation, the community-level metabolic network is defined as a stoichiometric coefficient matrix S, which combines the stoichiometric coefficient matrixes S_i of all species. pFBA is used to simulate metabolic fluxes at a steady state when maximizing an objective function under given condition.

To investigate the potential metabolic capabilities of the identified species in the gut microbiota composition, we first constructed genome-scale metabolic models of 827 individual species, which include 456 gram-positive species, 331 gram-negative species, 30 bacteria with missing gram information, eight archaea and two eukaryotes. All the individual species GEMs were constructed by a semiautomatic pipeline based on the MetaCyc database. The distributions of reactions, metabolites, genes and gap reactions of all species GEMs are shown in **Figure 2.14**. After adding 86.25 ± 38.24 gap reactions, all species GEMs could grow under the dGMM+LAB medium, which is a mixture of the defined gut microbiota medium

(GMM) and the LAB medium, where the gut organisms could grow more than in other mediums. Furthermore, we construct community-level metabolic models of individual gut microbiota by considering each microbial GEM as one component of the entire metabolic model. Additionally, the overall biomass reaction of the community-level metabolic model was set as the weighted combination of the biomasses of all species GEMs, where the species abundance was used as the coefficients for each species biomass.

We first simulated the potential production capacity of some representative metabolites, such as SCFAs (lactic and acetic acids) and ethanol, in each community-level metabolic model by optimizing the biosynthesis of metabolites. In addition, we calculated the flux distribution under the maximization of biomass growth for the community level metabolic model by performing pFBA. From community level metabolic models, we also counted the reactions in different MetaCyc pathways, and the proportion of the counted reactions to all reactions in one MetaCyc pathway could present the completeness of the pathway. The sum of the exchange reaction fluxes, and the internal reaction fluxes that appeared in different species compartments was used as the final characteristic. Through simulation analysis of community-level metabolic models, we obtained the metabolic features, that allow for providing insight into metabolism or interrelationships in T2D-associated gut microbiota.

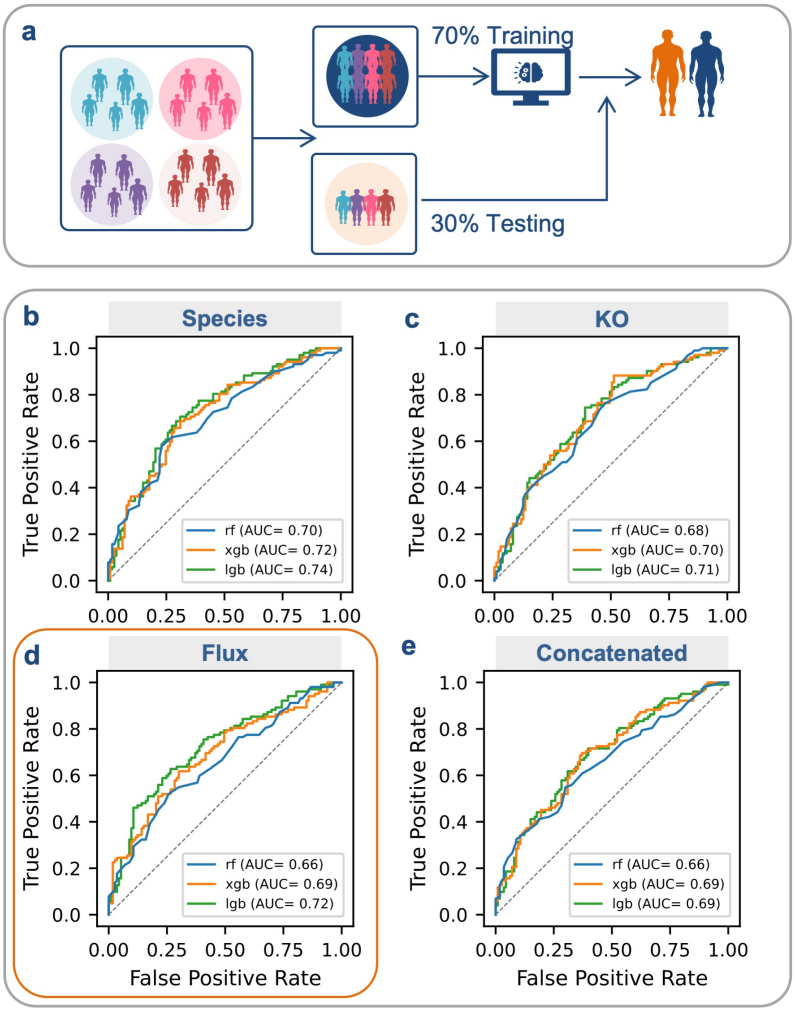


Figure 2.15| Training and testing dataset process and classification model for T2D state prediction based on gut microbial features. a) The pooled data from the four included studies was split into two parts, including 70% training dataset and 30% testing dataset. **b)** using species abundances. Performance of the predictive models validated on the testing dataset when pooling all data from the four included studies. **c)** KO profiles. **d)** the simulated reaction fluxes of the gut microbial community. **e)** The concatenated features of species, KOs and fluxes. The models were trained using three decision tree-based ML methods, including the light gradient boosting machine (LightGBM), extreme gradient boosting decision trees (XGBoost) and random forest (RF), as well as were adjusted by covariates age, Body mass index (BMI), ethnicity and gender.

All samples were split into two parts, 70% training dataset and 30% testing dataset. The prediction models discriminating the NGT and the T2D are first trained and evaluated by five-fold cross-validation, and then applied to predict the T2D state of a new sample in the test dataset. Predictive models adjusted for variables such as gender, age, and BMI had better classification performance for NGT and T2D identification in most cases compared to classifiers that did not adjust for covariates. It is strongly suggestive that confounding factors should be considered when studying the gut microbiota related to diabetic disease.

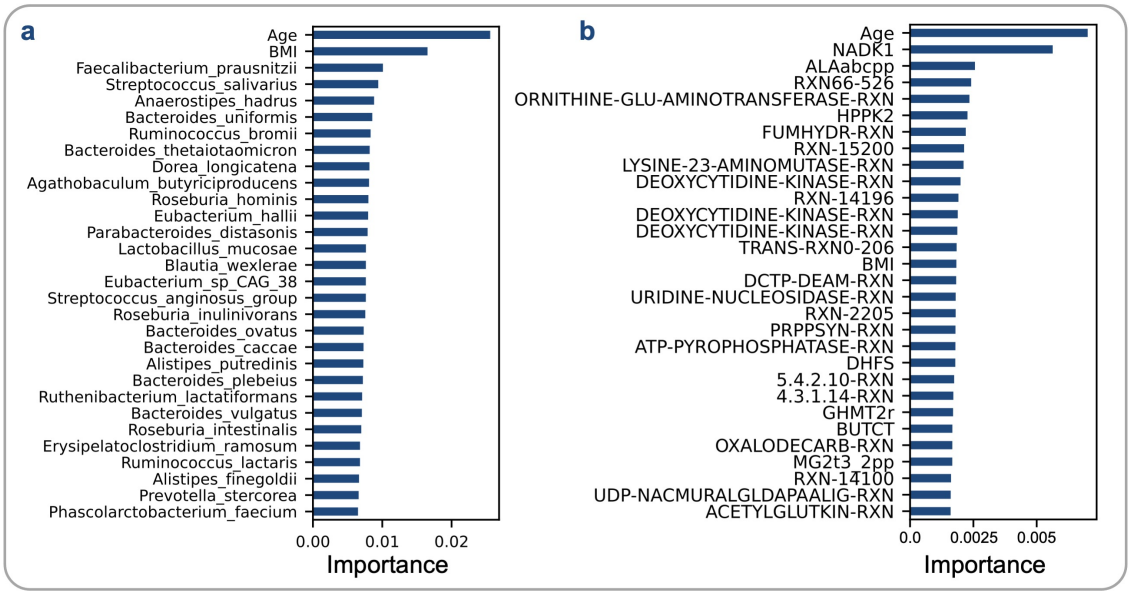


Figure 2.16| Important microbial features to distinguish **NGT from T2D. a)** The top 30 important microbial species. **b)** The top 30 important reaction fluxes. They were identified by training a random forest model based on the pooled data with adjustment by covariates age, BMI, ethnicity and gender.

The important microbial features in the classification were evaluated and ranked by the metric of mean decrease in Gini impurity (**Figure 2.16**). Age and BMI are considered the two most important factors in the classification of NGT and T2D, which may be significantly correlated with T2D and confound the relationship between gut microbiota and T2D. *Faecalibacterium prausnitzii*, three species of *Roseburia* (*Roseburia intestinalis, Roseburia hominis, Roseburia inulinivorans*), three species of *Bacteroides* (*Bacteroides uniformis*, *Bacteroides caccae* and *Bacteroides vulgatus*) and two species of *Ruminococcus* (*Ruminococcus bromii, Ruminococcus lactaris*) were identified as important for prediction of T2D status, which in agreement with previous studies. Part of them have been suggested to be butyrate-producing microbial species that have a beneficial effect on T2D. When using the pooled reaction fluxes, age, BMI, gender and ethnicity to train the model for predicting T2D risk. butyrate-producing related reactions were identified to be important for prediction of T2D status, such as NADK1 (NAD kinase GTP catalyzed by NAD kinase 1), ALAabcpp (L-alanine transport via ABC system), RXN-15200 (involved in L-phenylalanine biosynthesis III pathway), two reactions LYSINE-23-AMINOMUTASERXN and 4.3.1.14-RXN (involved in L-lysine fermentation to acetate and butyrate), and BUTCT (catalyzed by the acetyl-CoA: butyrate-CoA transferase), GHMT2r (catalyzed by glycine hydroxymethyltransferase). Therefore, ML combined with community-level metabolic models has the potential to enable the identification of novel gut microbial signatures associated with T2D.

4.2. The metabolic network inference framework for shotgun metagenomics

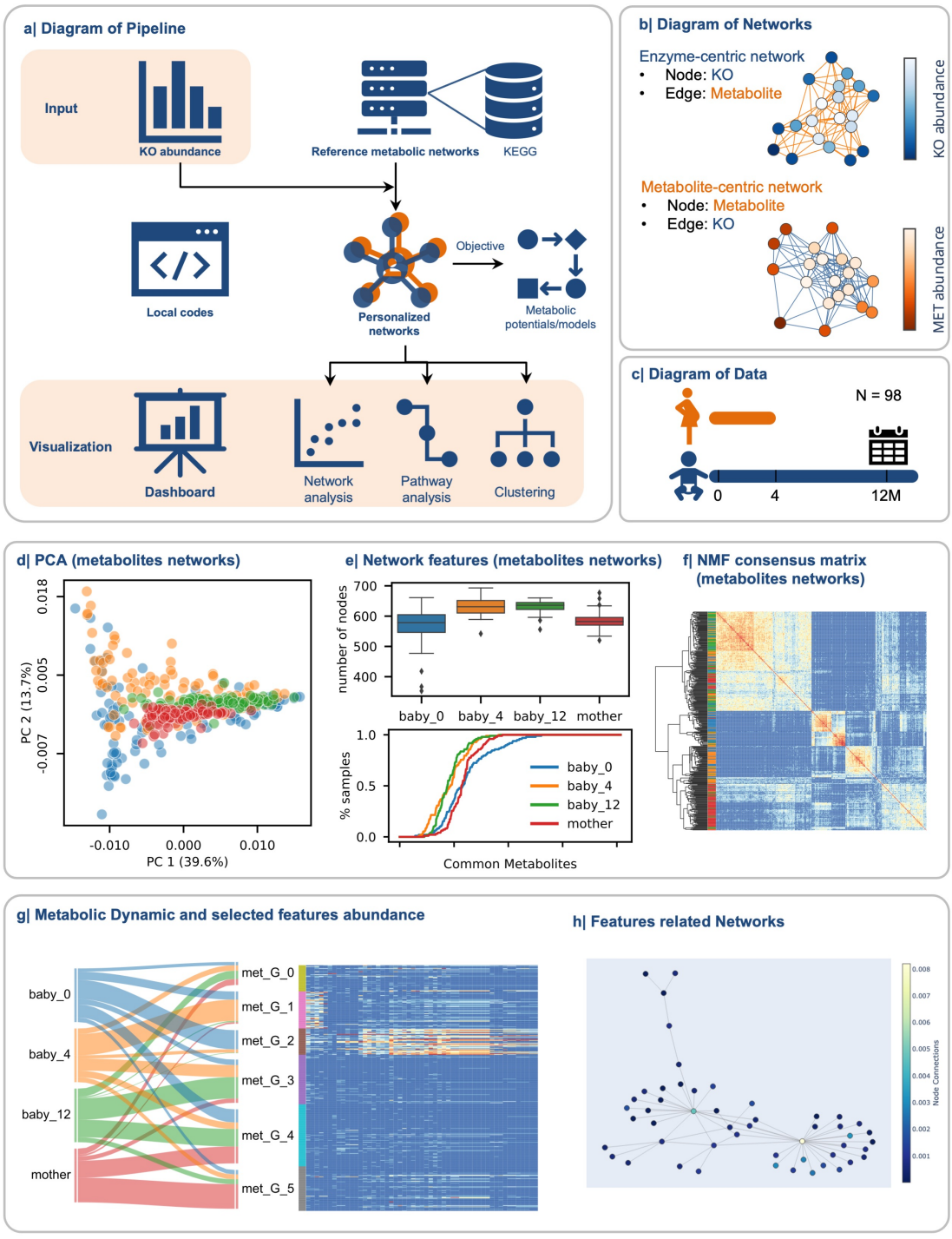


Figure 2.17| The implementation of Analyzer for Metabolic Networks (AMN) and results of the infant data set. a) Diagram of the pipeline. The AMN generates networks from our reference network, which is based on the KEGG database. After analysis and calculation, the results will be saved as local files or dashboard reports. The results demonstrate and visualize network structure, characteristics, pathways, clustering, etc. **b)** Diagram of networks. The ANM generates two types of networks; one is an enzyme-centric network and the other is a metabolite-centric network. The former uses enzymes as nodes and metabolites as edges, and the latter is the opposite. All subsequent analyses are performed in parallel on both networks. **c)** Diagram of the infant dataset. The data set contains four groups (born, four-month, 12-month, and mother groups) and 98 samples in each group. **d)** Diagram of PCA (for metabolites networks).

Preliminary check of the data for discrepancies. **e)** Features of the networks. Network properties such as the number of nodes and common nodes in four groups can be shown. **f)** Consensus matrix of the NMF results. For the metabolite abundances data, AMN estimated rank-6 to be the best rank and plotted the consensus matrix after 10,000 runs. **g)** Classification of data and abundance of selected features. The dynamics from the initial metabolic group (born, four-month, 12-month, and mother groups) to NMF predicted groups (met_G_1-6). The heatmap shows the NMF selected feature metabolites and their abundance. **h)** Networks of selected metabolites. Users can select nodes according to their requirements or NMF analysis.

In this section, we developed an effective tool to rapidly build metabolic networks and analyze microbiome data. The Analyzer for Metabolic Networks (AMN) toolbox is a Python programming language-based package. AMN provides both the command line and a graphical interface for constructing and analyzing metabolic networks. AMN can generate two types of directed and weighted networks; one is an enzyme-centric network and the other is a metabolite-centric network. The enzyme-centric network uses enzymes as nodes and metabolites as edges, and the metabolite-centric network uses metabolites as nodes and enzymes as edges. Both networks come from reference networks based on the KEGG database. Local figures or dashboard reports will visualize network structures and properties. On the interactive dashboard, users can display any property according to their requirements. To implement our package, we introduced a shotgun metagenomic dataset of infants and their mothers with four groups and 98 samples in each group. Data diagrams and principal component analysis (PCA) can be found in **Figure 2.17**. The number of metabolites is increasing from the baby_0 group to the baby_12 group, and the baby_0 group contains less coverage of common nodes. Nonnegative matrix factorization (NMF) and clustering analysis also be applied to investigate potential features and classification. The AMN could estimate the best rank for NMF and select key features and samples. Their consensus matrices that help to show and measure the stability of the clusters can be plotted.

5. Conclusions

In this thesis, I focus on modeling gut microbiota for steady state and dynamic systems, from single species to communities. Through modeling, I studied the metabolism of gut microbiota and its impact on human health.

In the first part (**Paper I &II**), I focused on *L. reuteri*, reconstructed a GEM of *L. reuteri* ATCC PTA 6475 to simulate the metabolic capabilities and growth rates under different media. I further reconstructed core and pan- GEMs of 35 *L. reuteri* strains to study their metabolism diversity. All the GEMs provided a reliable basis to investigate the metabolism of *L. reuteri* in detail and their potential benefits on host health. Furthermore, I investigated the effects of *L. reuteri* on bone metabolism with limited clinical data. The results indicated that supplementation with *L. reuteri ATCC PTA 6475* could have beneficial effects on bone metabolism.

In the second part (**Paper III**), I attempted to model the dynamic behaviors of gut microbes and developed a methodology to link GEM and HCM. The new HCM strategy with the yield analysis algorithm (opt-yield-FBA) can simulate metabolic dynamics at the genome-scale. The opt-yield-FBA is an FBA based method that can calculate optimal yield solutions and yield space for GEMs. Finally, I illustrated the strategy to simulate the dynamics of microbial communities.

In the third part (**Paper IV &V**), I reconstructed the models of microbial communities for the T2D cohort and simulated their metabolic capabilities. By integrating machine learning with metabolic modeling, a number of SCFAs producing bacterial species and metabolic reactions were consistently identified to be associated with T2D status. I further developed a toolbox for reconstructing metabolic networks and analyzing the structure and properties of networks. A shotgun metagenomic dataset of infants was analyzed as a case study for implementation.

In conclusion, this thesis has achieved most of our aims. We successfully modeled the gut microbiota under steady and dynamic states, simulated its behaviors, and explored the relationships with human health. At the same time, our modeling of the dynamics of microbial communities still needs to be improved. How to apply the dynamic model to communities and enhance the robustness of the model is what we need to do afterward.

6. Future perspectives

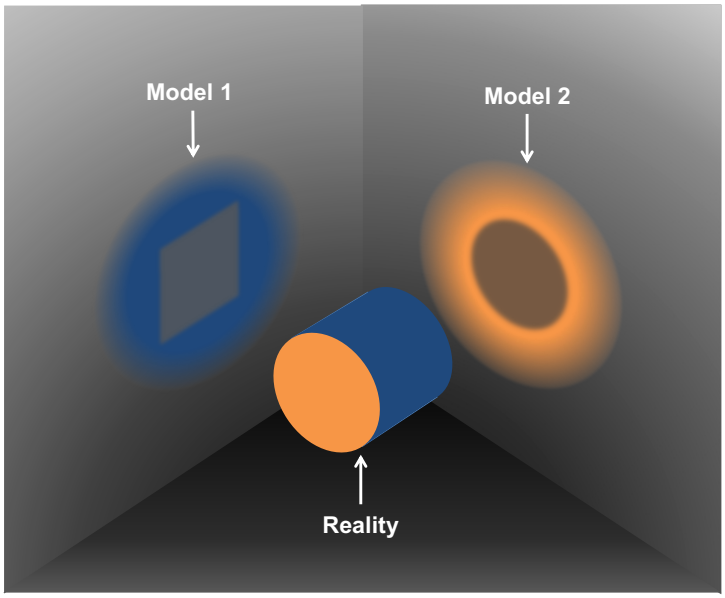


Figure 2.18| Model and reality.

How far apart are model and reality? I don't know, but I believe the models are approaching reality. In my own opinion, I believe that biological systems are one of the most complex systems in the world. Before accumulating enough knowledge, we just try to learn from this complex system in a crude way. A model is an essential tool to learn from the world and helps us simulate and approach reality. It is also an induction of the available data to explore the mechanism behind the phenomenon.

This thesis is an attempt to explore the metabolic modeling of gut microbes and there remain topics to be explored.

The quantity and quality of GEMs. Depending on the research questions, some studies continuously improve the quality of a single species GEM and train the model with more experimental data, for instance, yeast and *E. coli* GEMs. Some studies construct a large number of models based on limited experimental data and ignore the quality of a single model. How to keep the balance between quality and quantity is what needs to be considered for gut microbiota modeling. The GEM method comes from systems biology and metabolic engineering. Some industrial strains have a lot of experimental data to support their GEMs and form a design-test-learning cycle. To improve the model's ability, some studies have integrated more information in GEMs, such as transcriptome, proteome, and enzyme data. These experimental data-based experiences may hardly help us to improve the quality of all individual GEMs in the microbial community. However, we can first focus on the quality of a few representative strains of GEMs and establish a design-test-learn cycle. There are

challenges and opportunities in the development of modeling tools, quality control, and format standardization for modeling a large number of species.

The purpose of modeling is to reveal mechanisms or find correlations. For revealing mechanisms, interaction models within humans should be considered; for prediction, black-box models such as machine learning help us locate important features from lots of data.

In dynamic system modeling, there is still a long way to go. Kinetic models with complete enzyme data are trustworthy but are limited by computational power. Some studies reduce the model to limit the number of parameters and avoid overfitting. In the meantime, the simplification of metabolism limits the usage of the model in specific situations. Giving the parameters a defined biological meaning and a mechanism and verifying them from the experimental data is a way to keep the balance between overfitting and the number of parameters.

7. Acknowledgements

Out of four years of doctoral study, I went through three years of pandemics, and almost one year of war and inflation. That's my unique experience as a PhD student. Honestly, it hasn't been a perfect four years, but I've been lucky to meet many lovely and warm people along the way.

I am very grateful to meet my supervisor, Prof. Jens Nielsen. Thanks for the opportunity to join Sysbio, for the exciting projects, and for a lot of support and freedom. I not only learned knowledge and efficiency from Jens, but also felt his patience, his understanding, and his encouragement. Without his trust, it would have been impossible to finish my studies. I also appreciated the opportunity to visit BII, where I saw a passionate work environment. These experiences and inspirations will be valuable for my future career planning and work passion. I am so lucky and proud to be a student of Jens. I'd also like to thank my co-supervisor, Boyang. I am very happy to be in the same office, where we have had many discussions and brainstorming. Thank him for his advice, guidance, and encouragement whenever I encounter difficulties. I would also like to thank my co-supervisor Aleksej and examiner Verena, who have been very helpful in my PhD study. I learned a lot from both of you.

Big thanks to all the people in Sysbio. They are sweet, friendly, and full of joy and respect. Thanks to Peishun, I've learned a lot from him in the seven years since I've known him. He is intelligent, practical and hardworking, all of those I should learn. Thanks to Hao.W Jun and Gang, I enjoyed working with them and got a lot of inspiration from their perspectives. Thanks to Xin, Lei, Yating, Boyang, Gang and Yanyan, for enjoying food together. Food is one of the most important reasons I love the world, and it is so nice to have food with you. It is also super fun to play video games with Gang, Yating, Yu and Jichen. I traveled halfway across the earth and feel so lucky to know Mihail, Demi, Angelo, Parizad, Promi, Fariba, Cheewin, Andrea and all the friends in Sysbio. Your greetings and smiles, like sunshine, make the Swedish winter warm and happy. I hope that the group will grow better and better.

Thanks to my friends Jing, Yuchong, Francy, and Jian for eating together, going to the gym together, traveling together, and making the boring life in Sweden colorful. I also thank my friends who I haven't seen for many years in China, Yujie, Yuanyuan, Yan, Te and Huan. Your caring makes me full of hope for life. I would also like to thank fysiken gym, although I don't go there very often; thanks to the "tuo kou xiu da hui", "qi pa shuo" and other shows that bring me joy; thank "One Piece" for reminding I am still a kid; and thank the broadcast "mo mo dao lai" which I've been listening to for over ten years, it’s sweet and healing.

Thank you to my family. Their love and support never change. I'm sorry I wasn't there for my parents when bad things happened in the last four years. Thank you to my two sisters, who gave me great encouragement and support all the way. I love you.

There are too many joys and stories to list here (too difficult to translate into English). Thanks for everything in the past four years. Whether they are happy or sad things, all of these have shaped me as I am today. These stories are very ordinary, but they will never be forgotten. They are carefully placed in the corner, or deposited in my memory box, waiting for the next coincidental opening.

Hao Luo

Acknowledgements (in Chinese)
致谢：

读博四年，三年疫情，一年的战争与通膨，不可思议的事情集中发生在我读博期间，组成了我奇妙的读博生活。老实说，这不是完美的四年，但我很幸运，一路走来，遇到了很多可爱而温暖的人。

我很感激遇到我的导师 Jens 教授。感谢他给予我加入 Sysbio 的机会，给予我一个有趣的课题，并给予足够的支持与自由。在他身上，我不仅学到渊博的智慧与极高的效率，最重要的是（是老延太帅了！），无数的耐心，包容与鼓励，没有这些，我很难坚持走下来。我也很感谢让我去 BII 访学，在那里我看到充满激情员工和公司，也从导师身上看到他对事业的热爱。这些经历与启发，将会是我未来职业规划与工作激情的宝贵财富。能成为老延的学生，实为幸运与自豪。我也很感谢我的副导师博阳师兄，在同一个办公室的日子里，自由的讨论是十分幸运且难得的经历。他的观点和意见对我的的课题都有至关重要的帮助，也感谢他在我遇到困难时候的开导与鼓励。也感谢副导师 Aleksej 和 Verena，在我博士研究中都有很大帮助，与他们交流过程中，学到很多。

感谢 Sysbio 中一起工作的小伙伴，他们可爱，友善，充满欢乐与尊敬。感谢顺子师兄，与你认识的七年中，我学会了很多，聪明，踏实，勤奋，这些都对我有莫大的帮助。感谢汪浩师兄和李刚师兄耿俊师姐，很开心能跟你们一起工作，也我很享受跟你们一起谈论课题，在你们的观点中受到很多启发。感谢欣姐无数次的投喂，还有石磊师姐，雅婷师姐，博阳师兄，李刚师兄，岩岩师姐，一起吃饭的时光很快乐，美食是我爱上世界的最重要原因之一，能与你们一起分享，十分开心。还有跟刚哥，雅婷，包包，陈禹一起开黑的日子也超级开心。我很幸运，跨越半个地球遇见了 Demi，Angelo，Mihail，Parizad，Prmi，Fariba，Rassol，Cheewin，Andrea 以及所有 Sysbio 的小伙伴，你们的问候和微笑，就像阳光一样，让瑞典的冬天充温暖和快乐。也希望课题组发展的越来越好。

感谢我的朋友陈静，羽翀，法兰西，张健，一起吃饭，一起去健身房，一起旅游，让瑞典枯燥的生活变得丰富多彩。还有远方多年未见的小伙伴们，驴姐，圈圈，张岩，特特，欢。。你们时不时的关心让我对生活充满希望。

还要感谢 fysiken 健身房，虽然没去几次；感谢《脱口秀大会》《奇葩说》等综艺带给我欢乐；感谢尾田老师的《海贼王》让我记得自己还是个少年；还有听了好多年的广播《默默道来》，温馨且治愈。

感谢我的家人，虽然四年间，家中发生了很多变故，但是你们从来没有减少对我的关心，对于这段时间没在家中陪伴我很抱歉和愧疚。感谢两位姐姐，成长路上给予我莫大的鼓励与支持，也成为四年漂泊不能陪伴父母的一丝自我安慰。爱你们！

感谢所有人四年的同行或相遇，无数欢乐与庆幸不胜枚举，在此不能一一述说（要翻译成英语太难了。。。）。拥抱四年遇到很多人和事，开心的，不开心的，这一切塑造了一个现在的我，在此一并谢过。这些故事，其实很平常，却念不忘，它们放在角落，或存入记忆的盒子，等待下一次偶然开启。

我爱这个世界，希望世界会更好😊。

罗浩

8. References

Aagaard, K., Ma, J., Antony, K. M., Ganu, R., Petrosino, J., & Versalovic, J. (2014). The placenta harbors a unique microbiome. *Science Translational Medicine*, *6*(237). https://doi.org/10.1126/SCITRANSLMED.3008599/SUPPL_FILE/6-237RA65_SM.PDF

Aagaard, K., Riehle, K., Ma, J., Segata, N., Mistretta, T. A., Coarfa, C., Raza, S., Rosenbaum, S., van den Veyver, I., Milosavljevic, A., Gevers, D., Huttenhower, C., Petrosino, J., & Versalovic, J. (2012). A Metagenomic Approach to Characterization of the Vaginal Microbiome Signature in Pregnancy. *PLOS ONE*, *7*(6), e36466. https://doi.org/10.1371/JOURNAL.PONE.0036466

Abubucker, S., Segata, N., Goll, J., Schubert, A. M., Izard, J., Cantarel, B. L., Rodriguez-Mueller, B., Zucker, J., Thiagarajan, M., Henrissat, B., White, O., Kelley, S. T., Methé, B., Schloss, P. D., Gevers, D., Mitreva, M., & Huttenhower, C. (2012). Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. *PLOS Computational Biology*, *8*(6), e1002358. https://doi.org/10.1371/JOURNAL.PCBI.1002358

Ahamed, F., Song, H., & Ho, Y. K. (2021). Modeling Coordinated Enzymatic Control of Saccharification and Fermentation by Clostridium thermocellum During Consolidated Bioprocessing of Cellulose. *Biotechnology and Bioengineering*, bit.27705. https://doi.org/10.1002/bit.27705

Alayande, K. A., Aiyegoro, O. A., Nengwekhulu, T. M., Katata-Seru, L., & Ateba, C. N. (2020). Integrated genome-based probiotic relevance and safety evaluation of Lactobacillus reuteri PNW1. *PLOS ONE*, *15*(7), e0235873. https://doi.org/10.1371/journal.pone.0235873

Aric A. Hagberg, Daniel A. Schult, & Pieter J. Swar. (2008). *Exploring Network Structure, Dynamics, and Function using NetworkX*. http://conference.scipy.org/proceedings/SciPy2008/paper_2

Association, A. D. (2011). Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care*, *34*(Supplement_1), S62–S69. https://doi.org/10.2337/DC11-S062

Avershina, E., Storrø, O., Øien, T., Johnsen, R., Pope, P., & Rudi, K. (2014). Major faecal microbiota shifts in composition and diversity with age in a geographically restricted cohort of mothers and their children. *FEMS Microbiology Ecology*, *87*(1), 280–290. https://doi.org/10.1111/1574-6941.12223

Bäckhed, F. (2011). Programming of Host Metabolism by the Gut Microbiota. *Annals of Nutrition and Metabolism*, *58*(Suppl. 2), 44–52. https://doi.org/10.1159/000328042

Bauer, E., & Thiele, I. (2018). From Network Analysis to Functional Metabolic Modeling of the Human Gut Microbiota. *MSystems*, *3*(3). https://doi.org/10.1128/MSYSTEMS.00209-17/ASSET/E9898F05-B6E8-4993-B827-59B91C7C961A/ASSETS/GRAPHIC/SYS0031822180003.JPEG

Bengmark, S. (1998). Ecological control of the gastrointestinal tract. The role of probiotic flora. *Gut*, *42*(1), 2–7. https://doi.org/10.1136/GUT.42.1.2

Biagi, E., Candela, M., Turroni, S., Garagnani, P., Franceschi, C., & Brigidi, P. (2013). Ageing and gut microbes: Perspectives for health maintenance and longevity. *Pharmacological Research*, *69*(1), 11–20. https://doi.org/10.1016/J.PHRS.2012.10.005

Biagi, E., Nylund, L., Candela, M., Ostan, R., Bucci, L., Pini, E., Nikkïla, J., Monti, D., Satokari, R., Franceschi, C., Brigidi, P., & de Vos, W. (2010). Through Ageing, and Beyond: Gut Microbiota and Inflammatory Status in Seniors and Centenarians. *PLOS ONE*, *5*(5), e10667. https://doi.org/10.1371/JOURNAL.PONE.0010667

Biedermann, L., Zeitz, J., Mwinyi, J., Sutter-Minder, E., Rehman, A., Ott, S. J., Steurer-Stey, C., Frei, A., Frei, P., Scharl, M., Loessner, M. J., Vavricka, S. R., Fried, M., Schreiber, S., Schuppler, M., & Rogler, G. (2013). Smoking Cessation Induces Profound Changes in the Composition of the Intestinal Microbiota in Humans. *PLOS ONE*, *8*(3), e59260. https://doi.org/10.1371/JOURNAL.PONE.0059260

Bosma, E. F., Forster, J., & Nielsen, A. T. (2017). Lactobacilli and pediococci as versatile cell factories – Evaluation of strain properties and genetic tools. In *Biotechnology Advances* (Vol. 35, Issue 4, pp. 419–442). Elsevier Inc. https://doi.org/10.1016/j.biotechadv.2017.04.002

Britton, R. A., Irwin, R., Quach, D., Schaefer, L., Zhang, J., Lee, T., Parameswaran, N., & McCabe, L. R. (2014). Probiotic L. reuteri Treatment Prevents Bone Loss in a Menopausal Ovariectomized Mouse Model. *Journal of Cellular Physiology*, *229*(11), 1822–1830. https://doi.org/10.1002/jcp.24636

Browne, H. P., Forster, S. C., Anonye, B. O., Kumar, N., Neville, B. A., Stares, M. D., Goulding, D., & Lawley, T. D. (2016). Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature 2016 533:7604*, *533*(7604), 543–546. https://doi.org/10.1038/nature17645

Buchner, B. A., & Zanghellini, J. (2021). EFMlrs: a Python package for elementary flux mode enumeration via lexicographic reverse search. *BMC Bioinformatics*, *22*(1), 1–21. https://doi.org/10.1186/S12859-021-04417-9/FIGURES/9

Carbonero, F., Benefiel, A. C., Alizadeh-Ghamsari, A. H., & Gaskins, H. R. (2012). Microbial pathways in colonic sulfur metabolism and links with health and disease. *Frontiers in Physiology*, *3 NOV*, 448. https://doi.org/10.3389/FPHYS.2012.00448/BIBTEX

Claesson, M. J., Cusack, S., O'Sullivan, O., Greene-Diniz, R., de Weerd, H., Flannery, E., Marchesi, J. R., Falush, D., Dinan, T., Fitzgerald, G., Stanton, C., van Sinderen, D., O'Connor, M., Harnedy, N., O'Connor, K., Henry, C., O'Mahony, D., Fitzgerald, A. P., Shanahan, F., … O'Toole, P. W. (2011). Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(SUPPL. 1), 4586–4591. https://doi.org/10.1073/PNAS.1000097107/SUPPL_FILE/PNAS.201000097SI.PDF

Cocquyt, T., Zhou, Z., Plomp, J., & van Eijck, L. (2021). Neutron tomography of Van Leeuwenhoek's microscopes. *Science Advances*, *7*(20), eabf2402. https://doi.org/10.1126/sciadv.abf2402

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., Ling, A. v., Devlin, A. S., Varma, Y., Fischbach, M. A., Biddinger, S. B., Dutton, R. J., & Turnbaugh, P. J. (2013). Diet rapidly and reproducibly alters the human gut microbiome. *Nature 2013 505:7484*, *505*(7484), 559–563. https://doi.org/10.1038/nature12820

D'hoe, K., Vet, S., Faust, K., Moens, F., Falony, G., Gonze, D., Lloréns-Rico, V., Gelens, L., Danckaert, J., De Vuyst, L., & Raes, J. (2018). Integrated culturing, modeling and transcriptomics uncovers complex interactions and emergent behavior in a three-species synthetic gut community. *ELife*, *7*. https://doi.org/10.7554/eLife.37090

Donaldson, G. P., Lee, S. M., & Mazmanian, S. K. (2015). Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology 2015 14:1*, *14*(1), 20–32. https://doi.org/10.1038/nrmicro3552

Dore, M. P., Bibbò, S., Pes, G. M., Francavilla, R., & Graham, D. Y. (2019). Role of probiotics in helicobacter pylori eradication: Lessons from a Study of Lactobacillus reuteri Strains DSM 17938 and ATCC PTA 6475 (Gastrus®) and a Proton-Pump Inhibitor. *Canadian Journal of Infectious Diseases and Medical Microbiology*, *2019*. https://doi.org/10.1155/2019/3409820

Fang, X., Lloyd, C. J., & Palsson, B. O. (2020). Reconstructing organisms in silico: genome-scale models and their emerging applications. *Nature Reviews Microbiology 2020 18:12*, *18*(12), 731–743. https://doi.org/10.1038/s41579-020-00440-4

Finley, S. D., & Hatzimanikatis, V. (2021). Editorial Overview: Mathematical modeling: It's a matter of scale. *Current Opinion in Systems Biology*, *28*. https://doi.org/10.1016/J.COISB.2021.100360

Geng, J., Song, H.-S., Yuan, J., & Ramkrishna, D. (2012). On enhancing productivity of bioethanol with multiple species. *Biotechnology and Bioengineering*, *109*(6), 1508–1517. https://doi.org/10.1002/bit.24419

Gensollen, T., Iyer, S. S., Kasper, D. L., & Blumberg, R. S. (2016). How colonization by microbiota in early life shapes the immune system. *Science*, *352*(6285), 539–544. https://doi.org/10.1126/SCIENCE.AAD9378/ASSET/2F91909F-0DF5-4161-88BB-078F2A92D227/ASSETS/GRAPHIC/352_539_F3.JPEG

Gill, S. R., Pop, M., DeBoy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., Gordon, J. I., Relman, D. A., Fraser-Liggett, C. M., & Nelson, K. E. (2006). Metagenomic analysis of the human distal gut microbiome. *Science*, *312*(5778), 1355–1359. https://doi.org/10.1126/SCIENCE.1124234/SUPPL_FILE/GILL.SOM.PDF

Gustafsson, J. K., Ermund, A., Johansson, M. E. V., Schütte, A., Hansson, G. C., & Sjövall, H. (2012). An ex vivo method for studying mucus formation, properties, and thickness in human colonic biopsies and mouse small and large intestinal explants. *American Journal of Physiology - Gastrointestinal and Liver Physiology*, *302*(4), 430–438. https://doi.org/10.1152/AJPGI.00405.2011/ASSET/IMAGES/LARGE/ZH30031261100007.JPEG

Henry, C. S., Bernstein, H. C., Weisenhorn, P., Taylor, R. C., Lee, J. Y., Zucker, J., & Song, H. S. (2016). Microbial Community Metabolic Modeling: A Community Data-Driven Network Reconstruction. *Journal of Cellular Physiology*, *231*(11), 2339–2345. https://doi.org/10.1002/JCP.25428

Hugon, P., Dufour, J. C., Colson, P., Fournier, P. E., Sallah, K., & Raoult, D. (2015). A comprehensive repertoire of prokaryotic species identified in human beings. *The Lancet Infectious Diseases*, *15*(10), 1211–1219. https://doi.org/10.1016/S1473-3099(15)00293-5

Jakobsson, H. E., Abrahamsson, T. R., Jenmalm, M. C., Harris, K., Quince, C., Jernberg, C., Björkstén, B., Engstrand, L., & Andersson, A. F. (2014). Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by Caesarean section. *Gut*, *63*(4), 559–566. https://doi.org/10.1136/GUTJNL-2012-303249

Jernberg, C., Löfmark, S., Edlund, C., & Jansson, J. K. (2007). Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *The ISME Journal 2007 1:1*, *1*(1), 56–66. https://doi.org/10.1038/ismej.2007.3

Karlsson, F. H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C. J., Fagerberg, B., Nielsen, J., & Bäckhed, F. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature 2013 498:7452*, *498*(7452), 99–103. https://doi.org/10.1038/nature12198

Kim, G. B., Kim, W. J., Kim, H. U., & Lee, S. Y. (2020). Machine learning applications in systems metabolic engineering. *Current Opinion in Biotechnology*, *64*, 1–9. https://doi.org/10.1016/J.COPBIO.2019.08.010

Kim, J., Varner, J. D., & Ramkrishna, D. (2008). A hybrid model of anaerobic E. coli GJT001: Combination of elementary flux modes and cybernetic variables. *Biotechnology Progress*, *24*(5), 993–1006. https://doi.org/10.1002/btpr.73

Kim, Y., Gu, C., Kim, H. U., & Lee, S. Y. (2020). Current status of pan-genome analysis for pathogenic bacteria. *Current Opinion in Biotechnology*, *63*, 54–62. https://doi.org/10.1016/J.COPBIO.2019.12.001

Klamt, S., Müller, S., Regensburger, G., & Zanghellini, J. (2018). A mathematical framework for yield (vs. rate) optimization in constraint-based modeling and applications in metabolic engineering. *Metabolic Engineering*, *47*, 153–169results. https://doi.org/10.1016/j.ymben.2018.02.001

Klamt, S., Regensburger, G., Gerstl, M. P., Jungreuthmayer, C., Schuster, S., Mahadevan, R., Zanghellini, J., & Müller, S. (2017). From elementary flux modes to elementary flux vectors: Metabolic pathway analysis with arbitrary linear flux constraints. *PLOS Computational Biology*, *13*(4), e1005409. https://doi.org/10.1371/journal.pcbi.1005409

Koh, A., Molinaro, A., Ståhlman, M., Khan, M. T., Schmidt, C., Mannerås-Holm, L., Wu, H., Carreras, A., Jeong, H., Olofsson, L. E., Bergh, P. O., Gerdes, V., Hartstra, A., de Brauw, M., Perkins, R., Nieuwdorp, M., Bergström, G., & Bäckhed, F. (2018). Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell*, *175*(4), 947-961.e17. https://doi.org/10.1016/J.CELL.2018.09.055

Kristjansdottir, T., Bosma, E. F., Branco dos Santos, F., Özdemir, E., Herrgård, M. J., França, L., Ferreira, B., Nielsen, A. T., & Gudmundsson, S. (2019). A metabolic reconstruction of Lactobacillus reuteri JCM 1112 and analysis of its potential as a cell factory. *Microbial Cell Factories*, *18*(1), 186. https://doi.org/10.1186/s12934-019-1229-3

Lane, N. (2015). The Unseen World: Reflections on Leeuwenhoek (1677) "Concerning Little Animal." *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1666), 20140344. https://doi.org/10.1098/rstb.2014.0344

LeBlanc, J. G., Milani, C., de Giori, G. S., Sesma, F., van Sinderen, D., & Ventura, M. (2013). Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Current Opinion in Biotechnology*, *24*(2), 160–168. https://doi.org/10.1016/J.COPBIO.2012.08.005

Li, G., Ji, B., & Nielsen, J. (2019). The pan-genome of Saccharomyces cerevisiae. *FEMS Yeast Research*, *19*(7), 64. https://doi.org/10.1093/FEMSYR/FOZ064

Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., Arumugam, M., Kultima, J. R., Prifti, E., Nielsen, T., Juncker, A. S., Manichanh, C., Chen, B., Zhang, W., Levenez, F., Wang, J., Xu, X., Xiao, L., Liang, S., … Mende, D. R. (2014). An integrated catalog of reference genes in the human gut microbiome. *Nature Biotechnology 2014 32:8*, *32*(8), 834–841. https://doi.org/10.1038/nbt.2942

Louis, P., & Flint, H. J. (2017). Formation of propionate and butyrate by the human colonic microbiota. *Environmental Microbiology*, *19*(1), 29–41. https://doi.org/10.1111/1462-2920.13589

Louis, P., Hold, G. L., & Flint, H. J. (2014). The gut microbiota, bacterial metabolites and colorectal cancer. *Nature Reviews Microbiology 2014 12:10*, *12*(10), 661–672. https://doi.org/10.1038/nrmicro3344

Lu, H., Li, F., Sánchez, B. J., Zhu, Z., Li, G., Domenzain, I., Marcišauskas, S., Anton, P. M., Lappa, D., Lieven, C., Beber, M. E., Sonnenschein, N., Kerkhoven, E. J., & Nielsen, J. (2019). A consensus S. cerevisiae metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. *Nature Communications*, *10*(1), 1–13. https://doi.org/10.1038/s41467-019-11581-3

Lu, H., Li, F., Yuan, L., An Domenzain, I., Yu, R., Wang, H., Li, G., Chen, Y., Ji, B., Kerkhoven, E. J., & Nielsen, J. (2021). Yeast metabolic innovations emerged via expanded metabolic network and gene positive selection. *Molecular Systems Biology*, *17*(10), e10427. https://doi.org/10.15252/MSB.202110427

Luckey, T. D. (1972). Introduction to intestinal microecology. *The American Journal of Clinical Nutrition*, *25*(12), 1292–1294. https://doi.org/10.1093/AJCN/25.12.1292

Macfarlane, S., & Macfarlane, G. T. (2003). Regulation of short-chain fatty acid production. *Proceedings of the Nutrition Society*, *62*(1), 67–72. https://doi.org/10.1079/PNS2002207

Machado, D., Andrejev, S., Tramontano, M., & Patil, K. R. (2018). Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Research*, *46*(15), 7542–7553. https://doi.org/10.1093/NAR/GKY537

Martens, J. H., Barg, H., Warren, M., & Jahn, D. (2002). Microbial production of vitamin B12. *Applied Microbiology and Biotechnology 2001 58:3*, *58*(3), 275–285. https://doi.org/10.1007/S00253-001-0902-7

Maurice, C. F., Haiser, H. J., & Turnbaugh, P. J. (2013). Xenobiotics Shape the Physiology and Gene Expression of the Active Human Gut Microbiome. *Cell*, *152*(1–2), 39–50. https://doi.org/10.1016/J.CELL.2012.10.052

Mishra, S. K., Malik, R. K., Manju, G., Pandey, N., Singroha, G., Behare, P., & Kaushik, J. K. (2012). Characterization of a Reuterin-Producing Lactobacillus reuteri BPL-36 Strain Isolated from Human Infant Fecal Sample. *Probiotics and Antimicrobial Proteins*, *4*(3), 154–161. https://doi.org/10.1007/s12602-012-9103-1

Mizrahi-Man, O., Davenport, E. R., & Gilad, Y. (2013). Taxonomic Classification of Bacterial 16S rRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs. *PLOS ONE*, *8*(1), e53608. https://doi.org/10.1371/JOURNAL.PONE.0053608

Monk, J. M., Lloyd, C. J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., Feist, A. M., & Palsson, B. O. (2017). iML1515, a knowledgebase that computes Escherichia coli traits. *Nature Biotechnology*, *35*(10), 904–908. https://doi.org/10.1038/nbt.3956

Monod, J. (1978). THE PHENOMENON OF ENZYMATIC ADAPTATION And Its Bearings on Problems of Genetics and Cellular Differentiation. In *Selected Papers in Molecular Biology by Jacques Monod* (pp. 68–134). Elsevier. https://doi.org/10.1016/B978-0-12-460482-7.50017-8

Monod, Jacques. (1942). *Recherches sur la croissance des cultures bactériennes*. https://doi.org/10.3/JQUERY-UI.JS

Moore, W. E. C., & Holdeman, L. v. (1974). Human Fecal Flora: The Normal Flora of 20 Japanese-Hawaiians. *Applied Microbiology*, *27*(5), 961–979. https://doi.org/10.1128/AM.27.5.961-979.1974

Morrison, D. J., & Preston, T. (2016). Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Https://Doi.Org/10.1080/19490976.2015.1134082*, *7*(3), 189–200. https://doi.org/10.1080/19490976.2015.1134082

Mu, Q., Tavella, V. J., & Luo, X. M. (2018). Role of Lactobacillus reuteri in human health and diseases. *Frontiers in Microbiology*, *9*(APR), 1–17. https://doi.org/10.3389/fmicb.2018.00757

Müller, S., & Regensburger, G. (2016). Elementary vectors and conformal sums in polyhedral geometry and their relevance for metabolic pathway analysis. *Frontiers in Genetics*, *7*(MAY), 90. https://doi.org/10.3389/FGENE.2016.00090/BIBTEX

Musso, G., Gambino, R., & Cassader, M. (2010). Obesity, Diabetes, and Gut MicrobiotaThe hygiene hypothesis expanded? *Diabetes Care*, *33*(10), 2277–2284. https://doi.org/10.2337/DC10-0556

Narang, A., Konopka, A., & Ramkrishna, D. (1997). Dynamic analysis of the cybernetic model for diauxic growth. *Chemical Engineering Science*, *52*(15), 2567–2578. https://doi.org/10.1016/S0009-2509(97)00073-0

Neish, A. S. (2009). Microbes in Gastrointestinal Health and Disease. *Gastroenterology*, *136*(1), 65–80. https://doi.org/10.1053/J.GASTRO.2008.10.080

Nelson, K. E., Weinstock, G. M., Highlander, S. K., Worley, K. C., Creasy, H. H., Wortman, J. R., Rusch, D. B., Mitreva, M., Sodergren, E., Chinwalla, A. T., Feldgarden, M., Gevers, D., Haas, B. J., Madupu, R., Ward, D. v., Birren, B. W., Gibbs, R. A., Methe, B., Petrosino, J. F., … Zhu, D. (2010). A catalog of reference genomes from the human microbiome. *Science*, *328*(5981), 994–999. https://doi.org/10.1126/science.1183605

Nilsson, A. G., Sundh, D., Bäckhed, F., & Lorentzon, M. (2018). Lactobacillus reuteri reduces bone loss in older women with low bone mineral density: a randomized, placebo-controlled, double-blind, clinical trial. *Journal of Internal Medicine*, *284*(3), 307–317. https://doi.org/10.1111/joim.12805

Oh, P. L., Benson, A. K., Peterson, D. A., Patil, P. B., Moriyama, E. N., Roos, S., & Walter, J. (2010). Diversification of the gut symbiont Lactobacillus reuteri as a result of host-driven evolution. *The ISME Journal*, *4*(3), 377–387. https://doi.org/10.1038/ismej.2009.123

Orth, J. D., Fleming, R. M. T., & Palsson, B. Ø. (2010). Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide. *EcoSal Plus*, *4*(1). https://doi.org/10.1128/ECOSALPLUS.10.2.1/ASSET/AF2F042E-12AF-48A1-967D-9627A8D81014/ASSETS/GRAPHIC/10.2.1_FIG_022.GIF

Palau-Rodriguez, M., Tulipani, S., Queipo-Ortuño, M. I., Urpi-Sarda, M., Tinahones, F. J., & Andres-Lacueva, C. (2015). Metabolomic insights into the intricate gut microbial-host interaction in the development of

obesity and type 2 diabetes. *Frontiers in Microbiology*, *6*(OCT), 1151. https://doi.org/10.3389/FMICB.2015.01151/BIBTEX

Pedersen, H. K., Gudmundsdottir, V., Nielsen, H. B., Hyotylainen, T., Nielsen, T., Jensen, B. A. H., Forslund, K., Hildebrand, F., Prifti, E., Falony, G., le Chatelier, E., Levenez, F., Doré, J., Mattila, I., Plichta, D. R., Pöhö, P., Hellgren, L. I., Arumugam, M., Sunagawa, S., … Pedersen, O. (2016). Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature 2016 535:7612*, *535*(7612), 376–381. https://doi.org/10.1038/nature18646

Plovier, H., Everard, A., Druart, C., Depommier, C., van Hul, M., Geurts, L., Chilloux, J., Ottman, N., Duparc, T., Lichtenstein, L., Myridakis, A., Delzenne, N. M., Klievink, J., Bhattacharjee, A., van der Ark, K. C. H., Aalvink, S., Martinez, L. O., Dumas, M. E., Maiter, D., … Cani, P. D. (2016). A purified membrane protein from Akkermansia muciniphila or the pasteurized bacterium improves metabolism in obese and diabetic mice. *Nature Medicine 2016 23:1*, *23*(1), 107–113. https://doi.org/10.1038/nm.4236

Pompei, A., Cordisco, L., Amaretti, A., Zanoni, S., Matteuzzi, D., & Rossi, M. (2007). Folate production by bifidobacteria as a potential probiotic property. *Applied and Environmental Microbiology*, *73*(1), 179–185. https://doi.org/10.1128/AEM.01763-06/ASSET/B3681A48-DA64-4EFB-80BD-02E0F6F0B8EF/ASSETS/GRAPHIC/ZAM0010773810004.JPEG

Poretsky, R., Rodriguez-R, L. M., Luo, C., Tsementzi, D., & Konstantinidis, K. T. (2014). Strengths and Limitations of 16S rRNA Gene Amplicon Sequencing in Revealing Temporal Microbial Community Dynamics. *PLOS ONE*, *9*(4), e93827. https://doi.org/10.1371/JOURNAL.PONE.0093827

Proffitt, C., Bidkhori, G., Lee, S., Tebani, A., Mardinoglu, A., Uhlen, M., Moyes, D. L., & Shoaie, S. (2022). Genome-scale metabolic modelling of the human gut microbiome reveals changes in the glyoxylate and dicarboxylate metabolism in metabolic disorders. *IScience*, *25*(7), 104513. https://doi.org/10.1016/J.ISCI.2022.104513

Ramakrishna, R., Ramkrishna, D., & Konopka, A. E. (1996). Cybernetic Modeling of Growth in Mixed, Substitutable Substrate Environments: PreferentiaIand Simultaneous Utilization. *Biotechnology and Bioengineering*, *52*, 141–151. https://doi.org/10.1002/(SICI)1097-0290(19961005)52:1

Ramkrishna, D. (1983). *A Cybernetic Perspective of Microbial Growth* (pp. 161–178). https://doi.org/10.1021/bk-1983-0207.ch007

Ramkrishna, D., & Song, H.-S. (2018). Cybernetic Modeling for Bioreaction Engineering. In *Cybernetic Modeling for Bioreaction Engineering*. Cambridge University Press. https://doi.org/10.1017/9780511731969

Rana, P., Berry, C., Ghosh, P., & Fong, S. S. (2020). Recent advances on constraint-based models by integrating machine learning. *Current Opinion in Biotechnology*, *64*, 85–91. https://doi.org/10.1016/J.COPBIO.2019.11.007

Ridlon, J. M., Kang, D. J., Hylemon, P. B., & Bajaj, J. S. (2014). Bile Acids and the Gut Microbiome. *Current Opinion in Gastroenterology*, *30*(3), 332. https://doi.org/10.1097/MOG.0000000000000057

Rodríguez, J. M., Murphy, K., Stanton, C., Ross, R. P., Kober, O. I., Juge, N., Avershina, E., Rudi, K., Narbad, A., Jenmalm, M. C., Marchesi, J. R., & Collado, M. C. (2015). The composition of the gut microbiota throughout life, with an emphasis on early life. *Microbial Ecology in Health & Disease*, *26*(0). https://doi.org/10.3402/MEHD.V26.26050

Rouhollah, H., Iraj, N., Giti, E., & Sorah, A. (2007). Mixed sugar fermentation by Pichia stipitis, Sacharomyces cerevisiaea, and an isolated xylosefermenting Kluyveromyces marxianus and their cocultures. *African Journal of Biotechnology*, *6*(9).

Salminen, S., Gibson, G. R., McCartney, A. L., & Isolauri, E. (2004). Influence of mode of delivery on gut microbiota composition in seven year old children. *Gut*, *53*(9), 1388–1389. https://doi.org/10.1136/GUT.2004.041640

Santos, F., Teusink, B., Molenaar, D., van Heck, M., Wels, M., Sieuwerts, S., de Vos, W. M., & Hugenholtz, J. (2009). Effect of amino acid availability on vitamin B12 production in Lactobacillus reuteri. *Applied and Environmental Microbiology*, *75*(12), 3930–3936. https://doi.org/10.1128/AEM.02487-08

Santos, F., Vera, J. L., Lamosa, P., de Valdez, G. F., de Vos, W. M., Santos, H., Sesma, F., & Hugenholtz, J. (2007). Pseudovitamin B12 is the corrinoid produced by Lactobacillus reuteri CRL1098 under anaerobic conditions. *FEBS Letters*, *581*(25), 4865–4870. https://doi.org/10.1016/j.febslet.2007.09.012

Santos, F., Wegkamp, A., De Vos, W. M., Smid, E. J., & Hugenholtz, J. (2008). High-level folate production in fermented foods by the B12 producer Lactobacillus reuteri JCM1112. *Applied and Environmental Microbiology*, *74*(10), 3291–3294. https://doi.org/10.1128/AEM.02719-07

Sauer, M., Russmayer, H., Grabherr, R., Peterbauer, C. K., & Marx, H. (2017). The Efficient Clade: Lactic Acid Bacteria for Industrial Chemical Production. *Trends in Biotechnology*, *35*(8), 756–769. https://doi.org/10.1016/j.tibtech.2017.05.002

Savino, F., Pelle, E., Palumeri, E., Oggero, R., & Miniero, R. (2007). Lactobacillus reuteri (American Type Culture Collection Strain 55730) versus simethicone in the treatment of infantile colic: A prospective randomized study. *Pediatrics*, *119*(1), e124–e130. https://doi.org/10.1542/peds.2006-1222

Schepper, J. D., Collins, F. L., Rios-Arce, N. D., Raehtz, S., Schaefer, L., Gardinier, J. D., Britton, R. A., Parameswaran, N., & McCabe, L. R. (2019). Probiotic Lactobacillus reuteri Prevents Postantibiotic Bone Loss by Reducing Intestinal Dysbiosis and Preventing Barrier Disruption. *Journal of Bone and Mineral Research*, *34*(4), 681–698. https://doi.org/10.1002/jbmr.3635

Schroeder, B. O., & Bäckhed, F. (2016). Signals from the gut microbiota to distant organs in physiology and disease. *Nature Medicine 2016 22:10*, *22*(10), 1079–1089. https://doi.org/10.1038/nm.4185

Seaver, S. M. D., Liu, F., Zhang, Q., Jeffryes, J., Faria, J. P., Edirisinghe, J. N., Mundy, M., Chia, N., Noor, E., Beber, M. E., Best, A. A., DeJongh, M., Kimbrel, J. A., D'haeseleer, P., McCorkle, S. R., Bolton, J. R., Pearson, E., Canon, S., Wood-Charlson, E. M., … Henry, C. S. (2021). The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes. *Nucleic Acids Research*, *49*(D1), D575–D588. https://doi.org/10.1093/NAR/GKAA746

Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLOS Biology*, *14*(8), e1002533. https://doi.org/10.1371/JOURNAL.PBIO.1002533

Shoaie, S., Karlsson, F., Mardinoglu, A., Nookaew, I., Bordel, S., & Nielsen, J. (2013). Understanding the interactions between bacteria in the human gut through metabolic modeling. *Scientific Reports 2013 3:1*, *3*(1), 1–10. https://doi.org/10.1038/srep02532

Shoaie, S., & Nielsen, J. (2014). Elucidating the interactions between the human gut microbiota and its host through metabolic modeling. *Frontiers in Genetics*, *5*(APR). https://doi.org/10.3389/FGENE.2014.00086

Siebert, D., & Wendisch, V. F. (2015). Metabolic pathway engineering for production of 1,2-propanediol and 1-propanol by Corynebacterium glutamicum. *Biotechnology for Biofuels*, *8*(1), 91. https://doi.org/10.1186/s13068-015-0269-0

Song, H. S., Kim, S. J., & Ramkrishna, D. (2011). Synergistic Optimal Integration of Continuous and Fed-Batch Reactors for Enhanced Productivity of Lignocellulosic Bioethanol. *Industrial and Engineering Chemistry Research*, *51*(4), 1690–1696. https://doi.org/10.1021/IE200879S

Song, H. S., & Ramkrishna, D. (2009). Reduction of a set of elementary modes using yield analysis. *Biotechnology and Bioengineering*, *102*(2), 554–568. https://doi.org/10.1002/bit.22062

Song, H. S., & Ramkrishna, D. (2010). Prediction of metabolic function from limited data: Lumped hybrid cybernetic modeling (L-HCM). *Biotechnology and Bioengineering*, *106*(2), n/a-n/a. https://doi.org/10.1002/bit.22692

Song, H. S., & Ramkrishna, D. (2013). Complex Nonlinear Behavior in Metabolic Processes: Global Bifurcation Analysis of Escherichia coli Growth on Multiple Substrates. *Processes*, *1*(3), 263–278. https://doi.org/10.3390/PR1030263

Spinler, J. K., Taweechotipatr, M., Rognerud, C. L., Ou, C. N., Tumwasorn, S., & Versalovic, J. (2008). Human-derived probiotic Lactobacillus reuteri demonstrate antimicrobial activities targeting diverse enteric bacterial pathogens. *Anaerobe*, *14*(3), 166–171. https://doi.org/10.1016/j.anaerobe.2008.02.001

Sriramulu, D. D., Liang, M., Hernandez-Romero, D., Raux-Deery, E., Lünsdorf, H., Parsons, J. B., Warren, M. J., & Prentice, M. B. (2008). Lactobacillus reuteri DSM 20016 produces cobalamin-dependent diol dehydratase in metabolosomes and metabolizes 1,2-propanediol by disproportionation. *Journal of Bacteriology*, *190*(13), 4559–4567. https://doi.org/10.1128/JB.01535-07

Steuer, R. (2007). Computational approaches to the topology, stability and dynamics of metabolic networks. *Phytochemistry*, *68*(16–18), 2139–2151. https://doi.org/10.1016/J.PHYTOCHEM.2007.04.041

Stolyar, S., van Dien, S., Hillesland, K. L., Pinel, N., Lie, T. J., Leigh, J. A., & Stahl, D. A. (2007). Metabolic modeling of a mutualistic microbial community. *Molecular Systems Biology*, *3*, 92. https://doi.org/10.1038/MSB4100131

Straight, J. V., & Ramkrishna, D. (1991). Complex growth dynamics in batch cultures: Experiments and cybernetic models. *Biotechnology and Bioengineering*, *37*(10), 895–909. https://doi.org/10.1002/BIT.260371002

Su, M. S. W., Oh, P. L., Walter, J., & Gänzle, M. G. (2012). Intestinal origin of sourdough Lactobacillus reuteri isolates as revealed by phylogenetic, genetic, and physiological analysis. *Applied and Environmental Microbiology*, *78*(18), 6777–6780. https://doi.org/10.1128/AEM.01678-12

Tabák, A. G., Herder, C., Rathmann, W., Brunner, E. J., & Kivimäki, M. (2012). Prediabetes: a high-risk state for diabetes development. *Lancet (London, England)*, *379*(9833), 2279–2290. https://doi.org/10.1016/S0140-6736(12)60283-9

Tailford, L. E., Owen, C. D., Walshaw, J., Crost, E. H., Hardy-Goddard, J., le Gall, G., de Vos, W. M., Taylor, G. L., & Juge, N. (2015). Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. *Nature Communications 2015 6:1*, *6*(1), 1–12. https://doi.org/10.1038/ncomms8624

Terzer, M., & Stelling, J. (2008). Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics*, *24*(19), 2229–2235. https://doi.org/10.1093/bioinformatics/btn401

Thursby, E., & Juge, N. (2017). Introduction to the human gut microbiota. *Biochemical Journal*, *474*(11), 1823. https://doi.org/10.1042/BCJ20160510

Tyakht, A. v., Kostryukova, E. S., Popenko, A. S., Belenikin, M. S., Pavlenko, A. v., Larin, A. K., Karpova, I. Y., Selezneva, O. v., Semashko, T. A., Ospanova, E. A., Babenko, V. v., Maev, I. v., Cheremushkin, S. v., Kucheryavyy, Y. A., Shcherbakov, P. L., Grinevich, V. B., Efimov, O. I., Sas, E. I., Abdulkhakov, R. A., … Govorun, V. M. (2013). Human gut microbiota community structures in urban and rural populations in Russia. *Nature Communications 2013 4:1*, *4*(1), 1–9. https://doi.org/10.1038/ncomms3469

Vankerckhoven, V., Huys, G., Vancanneyt, M., Vael, C., Klare, I., Romond, M. B., Entenza, J. M., Moreillon, P., Wind, R. D., Knol, J., Wiertz, E., Pot, B., Vaughan, E. E., Kahlmeter, G., & Goossens, H. (2008). Biosafety assessment of probiotics used for human consumption: recommendations from the EU-PROSAFE project. In *Trends in Food Science and Technology* (Vol. 19, Issue 2, pp. 102–114). Elsevier Ltd. https://doi.org/10.1016/j.tifs.2007.07.013

Varma, A., & Palsson, B. O. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Applied and Environmental Microbiology*, *60*(10), 3724. https://doi.org/10.1128/AEM.60.10.3724-3731.1994

Vilkhovoy, M., Minot, M., & Varner, J. D. (2016). Effective Dynamic Models of Metabolic Networks. *IEEE Life Sciences Letters*, *2*(4), 51–54. https://doi.org/10.1109/lls.2016.2644649

von Kamp, A., & Schuster, S. (2006). Metatool 5.0: Fast and flexible elementary modes analysis. *Bioinformatics*, *22*(15), 1930–1931. https://doi.org/10.1093/bioinformatics/btl267

Walther, T., & François, J. M. (2016). Microbial production of propanol. In *Biotechnology Advances* (Vol. 34, Issue 5, pp. 984–996). Elsevier Inc. https://doi.org/10.1016/j.biotechadv.2016.05.011

Wang, H., Robinson, J. L., Kocabas, P., Gustafsson, J., Anton, M., Cholley, P. E., Huang, S., Gobom, J., Svensson, T., Uhlen, M., Zetterberg, H., & Nielsen, J. (2021). Genome-scale metabolic network reconstruction of model animals as a platform for translational research. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(30), e2102344118. https://doi.org/10.1073/PNAS.2102344118/SUPPL_FILE/PNAS.2102344118.SD04.XLSX

Wang, J., Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., Peng, Y., Zhang, D., Jie, Z., Wu, W., Qin, Y., Xue, W., Li, J., Han, L., … Wang, J. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature 2012 490:7418*, *490*(7418), 55–60. https://doi.org/10.1038/nature11450

Wiener, N. (1965). Perspectives in Cybernetics. *Progress in Brain Research*, *17*(C), 399–415. https://doi.org/10.1016/S0079-6123(08)60174-0

Woodmansey, E. J., McMurdo, M. E. T., Macfarlane, G. T., & Macfarlane, S. (2004). Comparison of compositions and metabolic activities of fecal microbiotas in young adults and in antibiotic-treated and non-antibiotic-treated elderly subjects. *Applied and Environmental Microbiology*, *70*(10), 6113–6122. https://doi.org/10.1128/AEM.70.10.6113-6122.2004/ASSET/44203252-DEC8-4B19-834B-BC66C0768B30/ASSETS/GRAPHIC/ZAM0100448740002.JPEG

Ye, C., Wei, · Xinyu, Shi, T., Sun, X., Xu, N., Cong Gao, ·, & Zou, W. (2022). Genome-scale metabolic network models: from first-generation to next-generation. *Applied Microbiology and Biotechnology 2022 106:13*, *106*(13), 4907–4920. https://doi.org/10.1007/S00253-022-12066-Y

Young, J. D. (2005). A system-level mathematical description of metabolic regulation combining aspects of elementary mode analysis with cybernetic control laws. *Theses and Dissertations Available from ProQuest*.

Yu, J., Zhao, J., Song, Y., Zhang, J., Yu, Z., Zhang, H., & Sun, Z. (2018). Comparative Genomics of the Herbivore Gut Symbiont Lactobacillus reuteri Reveals Genetic Diversity and Lifestyle Adaptation. *Frontiers in Microbiology*, *9*, 1151. https://doi.org/10.3389/fmicb.2018.01151

Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M. A. P., Harris, H. M. B., Mattarelli, P., O'toole, P. W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G. E., Gänzle, M. G., & Lebeer, S. (2020). A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. *International*

Journal of Systematic and Evolutionary Microbiology, *70*(4), 2782–2858. https://doi.org/10.1099/ijsem.0.004107

Zoetendal, E. G., Raes, J., van den Bogert, B., Arumugam, M., Booijink, C. C., Troost, F. J., Bork, P., Wels, M., de Vos, W. M., & Kleerebezem, M. (2012). The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *The ISME Journal 2012 6:7*, *6*(7), 1415–1426. https://doi.org/10.1038/ismej.2011.212

"When it comes to reading galley proofs, I always feel reminded of an awful sight once seen in a prisoner-of-war camp: a man slowly and deliberately eating his own vomit"

--Konrad Lorenz, 1973 Nobel Prize in Physiology or Medicine

Thanks for reading to the end