The developing infant gut microbiota: mathematical predictions of the effects of oligosaccharides

Versluis, D.M.

The developing infant gut microbiota: Mathematical predictions of the effects of oligosaccharides

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 23 april 2024
klokke 13:45 uur

door

David Mattias Versluis

geboren te Gorinchem

in 1996

Promotores:

Prof. dr. R.M.H. Merks	
Prof. dr. H.P. Spaink	

Promotiecommissie:

Prof. dr. A. H. Meijer	
Prof. dr. V. van Noort	
Prof. dr. D. Claessen	
Prof. dr. I. Thiele	University of Galway
Dr. K. Faust	KU Leuven
Dr. C. Belzer	Wageningen University & Research

The developing infant gut microbiota: Mathematical predictions of the effects of oligosaccharides

Cover design: *Laura Versluis*

Funding provided by FrieslandCampina

Printed by IPSKAMP Printing

It is for many of us incomparably more interesting to be wandering in a half-light uncovering hidden truths, bumping into others similarly occupied and adding to a body of knowledge to which one's friends and contemporaries are contributing, than to be reading a complete account of established facts.
Marjory Stephenson, *Bacterial Metabolism*, Longmans, Green and Co. (1948)

Contents

Chapter 1

Introduction

The human gut starts entirely or nearly entirely free of bacteria [1, 2], but is quickly colonised by many bacterial species [3, 4]. In a matter of days the bacterial population grows to around 10^{10} bacterial cells per gram feces [5, 6]. Together, these bacteria form a dynamic community in the gut. This community is known as the infant gut microbiota. Influenced by factors such as nutrition and the environment, each infant develops a unique microbiota. These microbiotas are very different from those typically found in adults, and deserve to be studied on their own. The infant gut microbiota has been recognised as important for infant health since the 19th century [7], but until recently it was difficult to identify and quantify the major bacterial species involved [8]. With metagenomics it is now possible to identify hundreds of different species in infant feces [3, 4, 9], and to analyse their metabolic capacity [10, 11], but it is still unclear why the infant gut microbiota reaches the different compositions that it does, and how this process is influenced by nutrition [12]. To develop new hypotheses, we develop mathematical models in this thesis. For these models, we make use of genome-scale metabolic models (GEMs) [10, 11]. GEMs are reconstructions of bacterial metabolic networks, based on bacterial genetics and enzyme functions [13]. Each GEM consist of a list of reactions that a particular organism can perform [13]. Each reaction has a set of input and output molecules. Together, these reactions and molecules form a network model. Constraint-based modelling can generate predictions for the flow of molecules through this network [13]. A GEM can be used to create a stoichiometric matrix that includes the stoichiometry of each reaction. Constraint-based modelling then applies additional constraints to these

stoichiometric matrices [13]. These constraints can be based on our knowledge of the bacterial environment, such as by constraining some reactions based on the availability of specific nutrients. By using these constraints, constraint-based modelling can create predictions for what flows of molecules through the network are plausible under different environmental conditions [13]. Constraint-based modelling predicts metabolism as sets of flux rates for input fluxes, internal fluxes, and output fluxes. These represent the flow of nutrients and metabolites into, through, and out of the cell. One internal flux rate is usually taken as a proxy for the biomass production rate. We use constraint-based modelling to make predictions on individual cells, and integrate many of these predictions to create predictions for the mechanisms behind the infant gut microbiota. These predictions may be used to guide future *in vitro* and *in vivo* research into the infant gut microbiota, and inform potential nutritional interventions. In this chapter we will first explore the typical ecology of the infant gut, followed by how constraint-based modelling works and how it can be used to provide hypotheses on the mechanisms of the infant gut microbiota.

1.1 The infant gut microbiota

1.1.1 Infant health

The infant gut microbiota plays an important role in infant health [14]. Bacteria of the genus *Bifidobacterium* are very common in the infant gut, and are associated with health benefits for infants, such as a lower chance to experience colic [15] or be underweight [16], and reduced inflammation [17]. These positive effects of *Bifidobacterium* are partially explained through its direct interaction with the immune system [17], and partially by the metabolites it produces [18, 19]. Other bacterial species are also associated with improved infant health. For example, butyrate production by species such as *Anaerobutyricum hallii* is associated with a reduced chance to develop allergies or colic [20, 21, 22, 23, 24]. Butyrate is the preferred source of nutrition for the infant's gut colonocytes [25], which may explain its beneficial effects. More generally, many species in the infant gut, including *Bifidobacterium* and butyrate producers, also release other acids into the gut [18, 26]. Acids produced by the microbiota can prevent the growth of pathogens such as some pathogenic *Escherichia coli* strains [19, 18]. Finally, some species in the infant gut can consume the mucin that is secreted by the gut wall [27], which may also affect infant health [28]. This

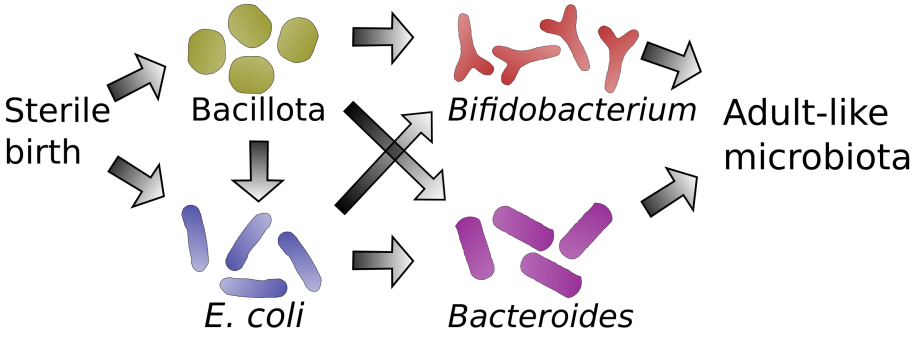


Figure 1.1: Typical temporal progression of dominant bacterial groups in the infant gut microbiota
Arrows indicate common transitions, but other transitions can also occur, and more community states exist. The presence and prevalence of trajectories other than *E. coli* to *Bifidobacterium* to adult-like vary greatly across studies. The trajectories presented here are based on [31, 32, 3, 33, 34]

consumption is associated with increased pathogen susceptibility in mice [28], possibly because mucin consumption exposes the gut wall directly to the bacteria in the gut [29]. A high abundance of *Bifidobacterium* may protect against mucin consumption in breastfed infants [30]. We will address the mechanics behind this in chapter 4. Because of these health effects, we are particularly interested in examining how and why the microbiota reaches the composition that it does, and how this composition may be influenced.

1.1.2 Microbiota composition

Studies of the healthy infant gut microbiota are necessarily limited to fecal samples [35]. These function as a proxy for the actual infant gut microbiota in the infant gut, but are not an unbiased representation of the whole gut microbiota [36, 35, 37]. Species that attach more strongly to the mucin may be under-represented in feces, for example [36, 35]. Initial studies into the composition of the infant gut microbiota were also limited to easily culturable organisms from feces. Though *E. coli* was first isolated and cultured from infant feces [7], other species were harder to culture. Even with the introduction of qPCR, incorrect primers caused a systematic under-reporting of *Bifidobacterium* species in feces until recently [8]. Metagenomic techniques became available to study the gut microbiota in the mid-2000's, and allow

for the analysis of all genomic material in a fecal sample [38]. This provided a culture and primer-independent view of the infant gut microbiota for the first time. However, bias is still present with this method, for example due to species-specific differences in DNA extraction efficiency [39], and some species that are detected with culture-based methods may not be detected by metagenomics [37]. Some general temporal patterns have been observed from the metagenomic analysis of infant fecal samples. These patterns vary greatly due to many factors, only some of which are known. Nonetheless, analysis of fecal samples currently provides the best available view on the infant gut microbiota. These analyses predict that for the first days after birth the infant gut microbiota in fecal samples is typically dominated by Enterobacteriaceae, particularly *E. coli* [32, 16, 31, 40]. This is typically followed by a succession to a microbiota dominated by *Bifidobacterium* species, particularly *Bifidobacterium longum* [32, 16, 31, 40]. However, this succession does not always occur. Various possible paths are visualised in Fig. 1.1. The gut microbiota of some infants (e.g. up to 17.6% in [33]) are instead dominated by Bacillota species such as *Lactobacillus* [33] or *Staphylococcus* [32], and some other infant gut microbiotas are dominated by *Bacteroides* species [34, 40, 3], such as *Bacteroides vulgatus* [3]. Regardless of community type, the microbiota typically shifts to an adult-like composition after several months. This shift often coincides with weaning [41, 40]. Such an adult-like microbiota is usually dominated by Bacillota or *Bacteroides* species [3, 40].

In addition to these very abundant species, many other species are present in the infant gut microbiota at all timepoints [3, 31]. These may include populations of the health-promoting butyrate producing bacteria [3, 20], but may also include pathogens [42] and species such as *Ruminococcus gnavus* and *Cutibacterium avidum*, whose health effects are unclear [43, 44].

Next to the temporal variability of the microbiota there is likely also spatial variation in the abundance of particular species within the microbiota. Such a spatial distribution has been observed in piglets [45] and adult humans [46, 47, 48, 49]. These studies revealed a complex distribution of species and bacterial interactions across the gut. For example, in piglets the relative abundance of *Lactobacillus* is highest proximally [46, 47], while *Bacteroides* is more abundant distally [47]. In adult humans *Bacteroides* species are also more abundant in the distal colon, and *Enterococcus* species are more abundant in the proximal colon [48]. Similarly, Proteobacteria are more abundant close to the colon wall in adult humans, while Bacillota are more abundant in the lumen of the colon [49]. Similar patterns may exist in the infant

gut, but there is no specific knowledge on how the microbiota is spatially organized in the infant gut.

1.1.3 Causes of variation in microbiota composition

The above described patterns are generalisations for the infant gut microbiota, and it is important to note that the presence and timing of these patterns vary greatly across infants [32, 33, 31]. Some of this variation may be explained by differences in the techniques that are used to analyse the microbiota [8], but even within the same studies a great variability is usually found between individuals of the same age [32, 33, 31]. Many studies have attempted to explain some of this variation, and it seems that many different factors interact [50, 51]. We will highlight some of these factors: cesarean or vaginal birth [3], antibiotic usage [52], infant nutrition [53], and region of birth [54]. Together, these can explain some, but not all variation in the observed infant gut microbiota.

Analysis of fecal samples has shown that vaginally born infants typically acquire some bacterial species from the mother's gut microbiota and vaginal microbiota [3, 55]. Although the influence of these sources fades over time, they can greatly influence the species present in the infant gut microbiota in the first days and weeks [56, 55]. Due to the the selection pressures in the infant gut, the effect of birth type on relative abundances of species is typically much smaller. The composition of the infant gut microbiota represents neither the mother's gut microbiota nor the vaginal microbiota [3, 56]. In infants born by cesarean section the relation between the mother's gut and vaginal microbiota is disturbed, and the infant's gut microbiota is richer in species acquired from the skin microbiota and other sources in the environment [3]. This typically leads to a microbiota that is dominated by *E. coli* for longer, and a later transition to *Bifidobacterium* [16]. Cesarean section is also associated with a lower abundance of *Bacteroides* species in the infant gut microbiota [3] and an overall lowered bacterial diversity [55]. However, these are only general patterns, and some infants born by cesarean section have a microbiota indistinguishable from the typical microbiota of vaginally born infants [51]. Altogether, infants born vaginally are expected to have a microbiota that is, on average, more beneficial for the infant's health than those that are born by cesarean section [3]. Breastfeeding may partially compensate for the negative effects on the microbiota of a cesarean section birth [50].

Infants are sometimes treated with antibiotics, which can greatly affect the gut

microbiota even when not administered orally [57]. Antibiotics cause a decrease in *Bifidobacterium* [57, 58], an increase in *E. coli* [59, 60, 58], and a decrease in total diversity [57, 58]. This effect is strongest in vaginally born infants [52]. The microbiota usually recovers, but often only partially [57]. Changes in nutrition may aid recovery [61], but it is unclear how consistently effective these are.

A very important difference in nutrition between infants is the type and quantity of prebiotic oligosaccharides [62]. Prebiotic oligosaccharides are short chains of sugars that cannot be digested by the infant, but can be digested by certain species in the infant gut microbiota, primarily *Bifidobacterium* species [63, 64, 65]. When prebiotic oligosaccharides are present, *Bifidobacterium* species can become very abundant [53]. All human milk contains prebiotic oligosaccharides, but the type and quantity varies based on age and genetics [66, 67]. Milk oligosaccharides also occur in nearly all mammal species [68], even those that do not produce lactose, such as the Virginia opossum (*Didelphis virginiana*)[69, 70]. Milk oligosaccharides are hypothesised to have evolved specifically to stimulate a beneficial infant gut microbiota [65, 70]. Many, but not all, human infant formulas (artificial infant milk) also contain prebiotic oligosaccharides, which may be identical to a type present in human milk, or unique to infant formula [71]. When present, these oligosaccharides also stimulate *Bifidobacterium* species [53]. This creates variability in the microbiota between infants fed formula with oligosaccharides, formula without oligosaccharides, and breastfed infants [53]. Additional variation is introduced by the variable presence of *Bifidobacterium* species. Of the *Bifidobacterium* species, *Bifidobacterium longum* ssp. *infantis* is considered the best adapted to consuming human milk oligosaccharides [65]. However, infants are only likely to acquire *B. longum* ssp. *infantis* in regions with historically high rates of breastfeeding [54].

The birth conditions, nutrition and environment of the infant are known to be important for the composition of the gut microbiota, but it remains unclear how they are mediated. We must turn to bacterial metabolism and ecology to answer why and how these factors ultimately contribute to the observed compositions of the microbiota.

1.1.4 Bacterial metabolism and ecology

Because bacteria are excreted from the gut regularly, it is assumed that most bacterial populations in the gut, or at least the bacterial populations found in feces,

need to grow rapidly to maintain a constant population in the gut [72]. Bacteria need nutrients for growth, which can be derived from three sources in the infant gut: nutrition undigested by the host, such as lactose from milk, nutrients released by the host, such as intestinal mucin, and metabolites produced by other bacteria, such as lactate. The type and amount of each of these sources greatly influence the abundance of bacterial species [53, 73, 74, 75]. Because bacteria can derive nutrition from multiple sources, this also leads to complex interactions between bacteria. Bacteria may, for example, simultaneously compete with each other for host-derived nutrition and benefit from consuming each other's metabolites. We will discuss the sources of nutrition and the possible interactions these allow.

Human milk and infant formula contain carbohydrates, proteins, and fats that can be taken up by the small intestine of the infants [66]. A small quantity of these compounds will inevitably evade uptake, and be available to the microbiota[76, 77]. It is unclear exactly how much lactose is available to the microbiota, but it is estimated to be around 2% [78]. As infant feces does not typically contain lactose [79], the microbiota likely consumes all lactose that reaches the colon [80]. The presence of lactose in infant formula greatly increases the abundance of *Bifidobacterium* [75, 81]. This indicates that at least some of the sugars in infant formula become available to the microbiota. The presence of lactose in the gut may lead to competitive relationships between bacteria.

Human milk and many infant formulas also contain prebiotic oligosaccharides. Prebiotic oligosaccharides serve a unique role because they are resistant to uptake by the infant, and so are nearly entirely available to the microbiota. Specialized enzymes are required to break down these oligosaccharides, so they are not consumed by as many different species as lactose [63, 64]. As mentioned before, prebiotic oligosaccharides are consumed primarily by species of the genus *Bifidobacterium* [63, 64]. Both intracellular and extracellular enzymes for degradation of prebiotic oligosaccharides exist in this genus [82, 83], and the enzymes used depend on both the *Bifidobacterium* species and the oligosaccharide that is digested [82, 83, 84, 85, 86]. Competition occurs between *Bifidobacterium* species for uptake of these oligosaccharides [83].

A second way in which oligosaccharides become available to the microbiota in the colon is through mucin. Mucin is continuously released from the gut wall by dedicated goblet cells, and consists largely of oligosaccharides [29]. The composition of these oligosaccharides depends on age and genetics [29, 73]. Several species, including *Bifidobacterium bifidum* and *Bacteroides* species, can consume these oligosaccharides

[87, 27]. This typically occurs extracellularly, allowing other species to profit from the breakdown products too [27]. Some infants have a microbiota that consumes very little mucin [30]. This indicates that it is possible for this source of nutrition to remain unused.

Extracellular digestion of oligosaccharides, both from mucin and from prebiotics, often causes some breakdown products to be lost to the environment. This makes the products available to other species as 'public goods' [88, 27]. This is also known as substrate cross-feeding [89], and may negatively affect the species that performs the extracellular digestion [89, 88]. Many bacterial species in the infant gut also excrete metabolites, such as lactate, that can serve as nutrition for other species [90, 74], which is known as metabolite cross-feeding [89]. This type of cross-feeding is typically commensalist, as it does not negatively affect the producer of metabolites [89]. Cross-feeding is an important factor in shaping the composition of the infant gut microbiota [27, 91]. Many species, including the main producers of the beneficial butyrate, rely on cross-feeding to acquire nutrients [92]. These species can use both substrate cross-feeding and metabolite cross-feeding [27]. Conversely, cross-feeding is also used by some pathogens to establish themselves in the infant gut microbiota [93, 94]. Some species in the infant gut, such as *Veillonella dispar*, even lack the ability to consume sugars altogether [95]. This species relies entirely on alternative carbon sources such as lactate [74]. The production of cross-feeding substrates is variable - for example *Bifidobacterium* only produces the cross-feeding substrate lactate when sugars are abundant [90]. Other cross-feeding substrates can also be produced by *Bifidobacterium*, such as propane-1,2-diol and fucose, but these depend on the composition of the oligosaccharides consumed [96, 82].

Because our knowledge of the infant gut is limited to *in vitro* experiments and analysis of fecal samples it remains unclear how much production and consumption of cross-feeding substrates occurs *in vivo*, and by what species. This makes it difficult to design, for example, nutritional interventions that stimulate species that cross-feed, such as the butyrate producers associated with improved infant health [20]. *In vitro*, *Bifidobacterium* may consume a specific oligosaccharide and produce metabolites that are consumed by butyrate-producing bacteria. These butyrate-producing bacteria are associated with improved colonocyte health *in vivo*. However, *in vivo* the metabolites produced by *Bifidobacterium* may instead be taken up by other species, that are not beneficial (Fig. 1.2). Even when specific combinations of species can be studied *in vitro*, it is not feasibly to study many combinations due to constraints on time and

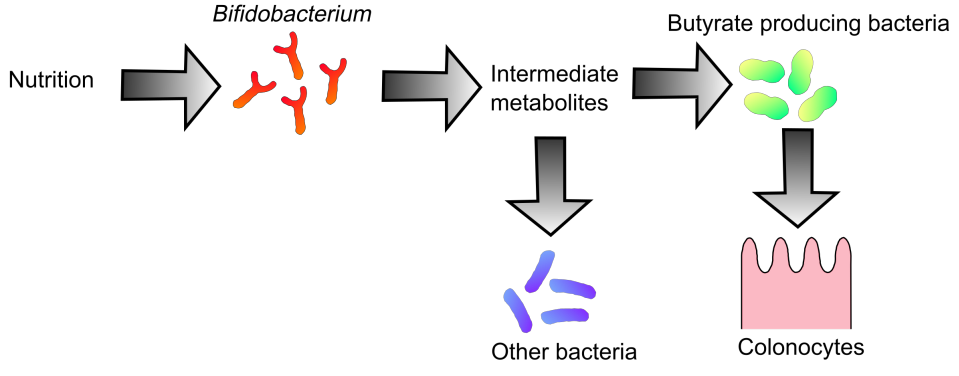


Figure 1.2: Schematic of potential cross-feeding in the infant gut
Nutrition, particularly prebiotic oligosaccharides, is consumed by *Bifidobacterium*, which produces intermediate metabolites such as lactate. These may be taken up by butyrate-producing bacteria, which produce butyrate that can be consumed by colonocytes. However, they may also be taken up by other bacteria that do not produce butyrate.

materials. More generally, it remains unclear how and why the ecology of the infant gut is formed, and what factors are most important in shaping them. Why are these particular species dominant, and others present at a low abundance? Can we shape the microbiota to have a certain composition? We turn to modelling of bacterial metabolism and ecology to try to provide more predictions and hypothesis to guide future experiments.

1.2 Modelling the infant gut microbiota

In this thesis we will use constraint-based modelling techniques to make predictions and create hypotheses for the mechanics of the infant gut. We will first present an overview of the methods used to model bacterial metabolism, followed by modelling of metabolic interactions over time and space.

1.2.1 Modelling bacterial metabolism

Many techniques exist for modelling bacterial metabolism and bacterial interactions, including systems of partial differential equations [97], inferred-network based methodology [98], and constraint-based modelling [99]. We will focus on constraint-based modelling, because of its ability to model very complex metabolisms and metabolic interactions without precise kinetic parameters.

The constraint based modelling techniques we discuss here start with a metabolic network reconstruction [13, 100]. A metabolic network reconstruction consists of a list of reactions, performed by enzymes, each of which has a set of input and output molecules. For example, the beta-galactosidase reaction takes the disaccharide lactose as input, and produces the monosaccharides glucose and galactose. Another reaction may convert galactose to glucose. Together, these reactions and molecules form a reconstructed metabolic network [13, 100] (Fig. 1.3A). From this network model, a stoichiometric matrix can be created that contains the stoichiometric coefficients of each reaction [13]. Many possible configurations of fluxes could flow through this system - we call these configurations solutions. Together, these solution form a solution space. Constraint-based modelling places constraints on the stoichiometric matrix to represent, for example, physical or chemical constraints [13]. This limits the solution space. The shape of the resulting solution space informs what set or sets of fluxes are plausible [13]. Concretely, constraint-based modelling can, as we will discuss in detail later, make predictions for the metabolic inputs and outputs of a bacterial population, what internal reactions are used in the process (Fig. 1.3 B), and how all this depends on the available inputs.

The constraint-based study of bacterial metabolic networks emerged from the study of capacitated flow networks [100]. Capacitated flow networks are graphs where each edge has a certain capacity. Capacitated flow networks were first used in an American study of Soviet railway networks [104]. These networks, where cities are nodes and railway lines are edges, could be analysed to show which lines were most vital to the functioning of the transportation network [104]. Similar analysis can be performed on graph representations of metabolic network models. Here, the nodes of a graph represent enzymes, and the edges linking them represent the molecules that the enzymes can convert into other molecules. Capacity limits can be imposed to represent the limited availability of certain molecules. Early analysis of bacterial metabolic networks models focused on analyzing very simplified networks that con-

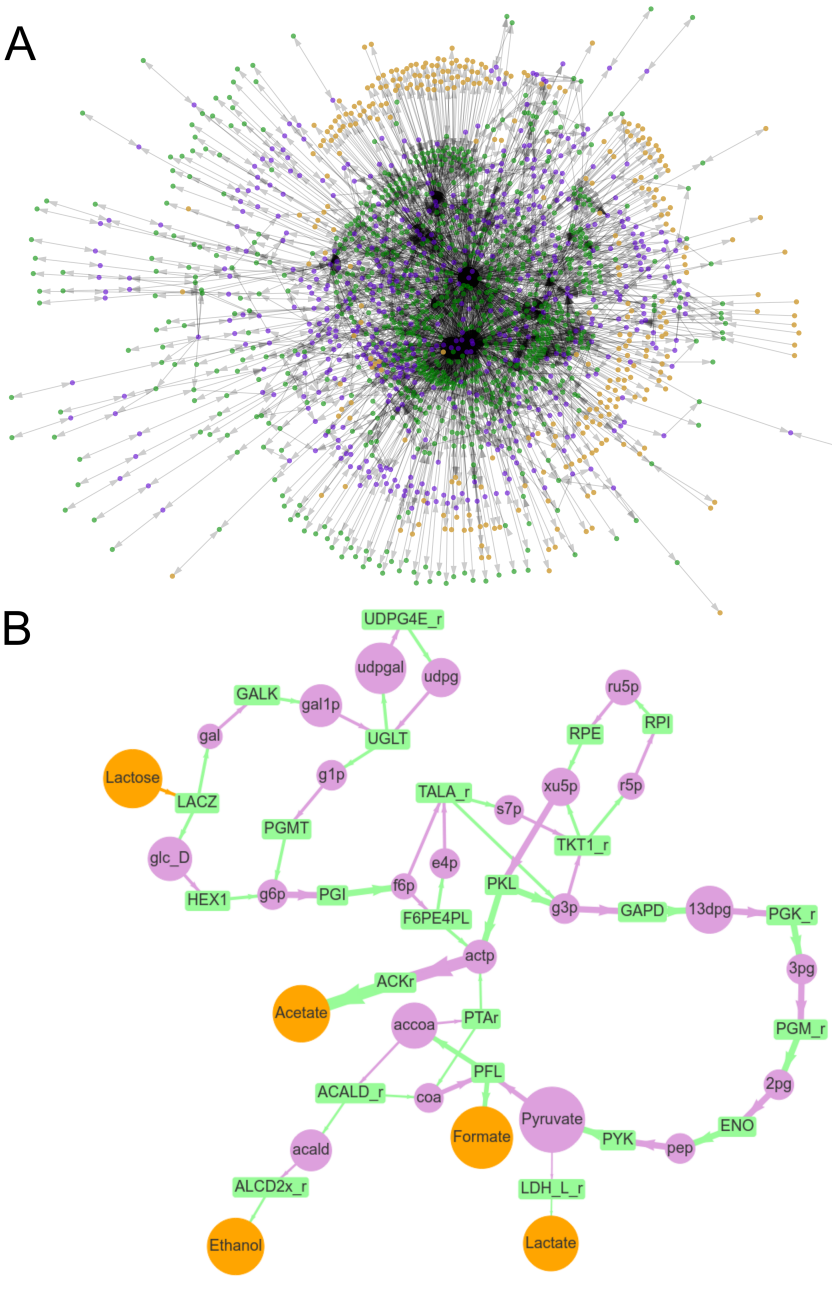


Figure 1.3: Visualisations of a genome-scale metabolic network model of *Bifidobacterium longum* Purple represents intermediate metabolites, orange represents metabolites that can be exchanged with the environment, and green represents reactions. Arrows indicate whether metabolites are consumed or produced by the reactions. GEM and names used from [101]. (A) Visualisation of the whole metabolic network model, created with ModelExplorer [102]. (B) Visualisation of a flux balance analysis solution of the network of A, when optimising ATP production with lactose and water as only inputs, and an enzymatic constraint [103]. Transport reactions and co-factors are not depicted.

tained only a few reactions [100]. More recent techniques allow for the reconstruction of genome-scale metabolic models (GEMs) that may contain thousands of reactions, a large part of the metabolism of a species [10]. These GEMs are generated by analysing the complete genome of a bacterial species, identifying the enzymes and metabolic functions present, and assigning appropriate reactions [10]. Further gap-filling and curation allows for the creation of metabolic models that match much, but not all, of the available *in vitro* data on, for example, which carbohydrates could be fermented by which bacterial species[10]. We use the AGORA database in this work, which contains over 800 GEMs of species within the human gut microbiota, including all major infant gut microbiota species [10]. These models encompass the consumption and creation of sugars, some polysaccharides, and amino acids, but not the prebiotic oligosaccharides that are common in the infant gut, or factors such as protein regulation. The lack of protein regulation means that the constitutive expression of enzymes, which occurs in many bacteria [105, 106], is not represented. These models thus provide a robust basis for modelling, but not a complete representation of metabolism. The models need to be extended to address some more specific questions, such as questions about prebiotic oligosaccharides.

From a GEM, bacterial metabolism can then be analysed using several constraint-based techniques, such as flux variability analysis [107], elementary flux mode analysis [108], and flux balance analysis (FBA) [13]. We will focus on FBA. FBA can create predictions for what fluxes flow through the reactions of a metabolic network under certain conditions. Concretely, FBA predicts a solution for the network. This solution consists of a set of fluxes $\vec{f}$, which includes input fluxes, internal fluxes, output fluxes, and flux through a reaction that is assumed to be a proxy for biomass production. To do this, FBA uses a stoichiometric matrix S generated from the GEM that contains the stoichiometry of all metabolic reactions [13]. Together, all these reactions are assumed to be in steady state:

$$S \cdot \vec{f} = 0,$$

where $\vec{f}$ is a vector of the metabolic fluxes through each reaction in the network, in quantity per time unit per population unit. One reaction is marked as the 'biomass reaction'. FBA assumes that the bacterium regulates its metabolism in such a way that flux through the biomass reaction is optimised. Given these assumptions, FBA uses linear programming to calculate what set or sets of fluxes $\vec{f}$ optimise this biomass

reaction. This set of fluxes is the prediction FBA makes of the bacteria's metabolism. It contains the input and output of metabolites, as well as a growth rate. To place the model in a more specific context, many additional constraints can be used in FBA. For example, reactions that can take up metabolites from the the environment, F_{in}, can be constrained by an upper bound F_{ub} that represents the availability of metabolites from the environment:

$$\vec{F_{in}} \leq \vec{F_{ub}}. \tag{1.1}$$

Additional limits can be set on the network to represent limited enzymatic capacity, such as by limiting the total amount of flux [100, 109]:

$$\sum \vec{f} \leq a \tag{1.2}$$

where the enzymatic constraint a is in quantity per time unit per population unit. Many additional techniques exist to constrain metabolism. For example, thermodynamic constraints can be introduced [110], or membrane occupancy constraints [111]. These constraints allow for more realistic predictions of bacterial metabolism by allowing for, for example, substrate concentration-dependent shifts in metabolism [112]. Specific reactions can also be added or disabled to represent mutated bacterial populations [13].

1.2.2 Modelling bacterial metabolic interactions

Initial models of bacterial metabolism represented a single population of bacteria exchanging nutrients and metabolites with the environment (Fig 1.4A). FBA was used to calculate a steady state, or a set of steady states, which could then be analysed [113, 114]. However, bacterial species are rarely alone in their ecosystem. Later models incorporated multiple species, which could each have their own network of reactions [115]. These populations were coupled to each other by their shared use of the same metabolite pool (Fig. 1.4B). Again, FBA was used to calculate steady states which could be analysed. This allowed for the modelling of competition and cooperation between species. However, many ecological processes are dynamic, and do not maintain a single steady state. To represent these conditions dynamic FBA (dFBA) was developed [116]. In dFBA the model is run in timesteps, and a steady state is calculated for each timestep. Within each timestep the populations are

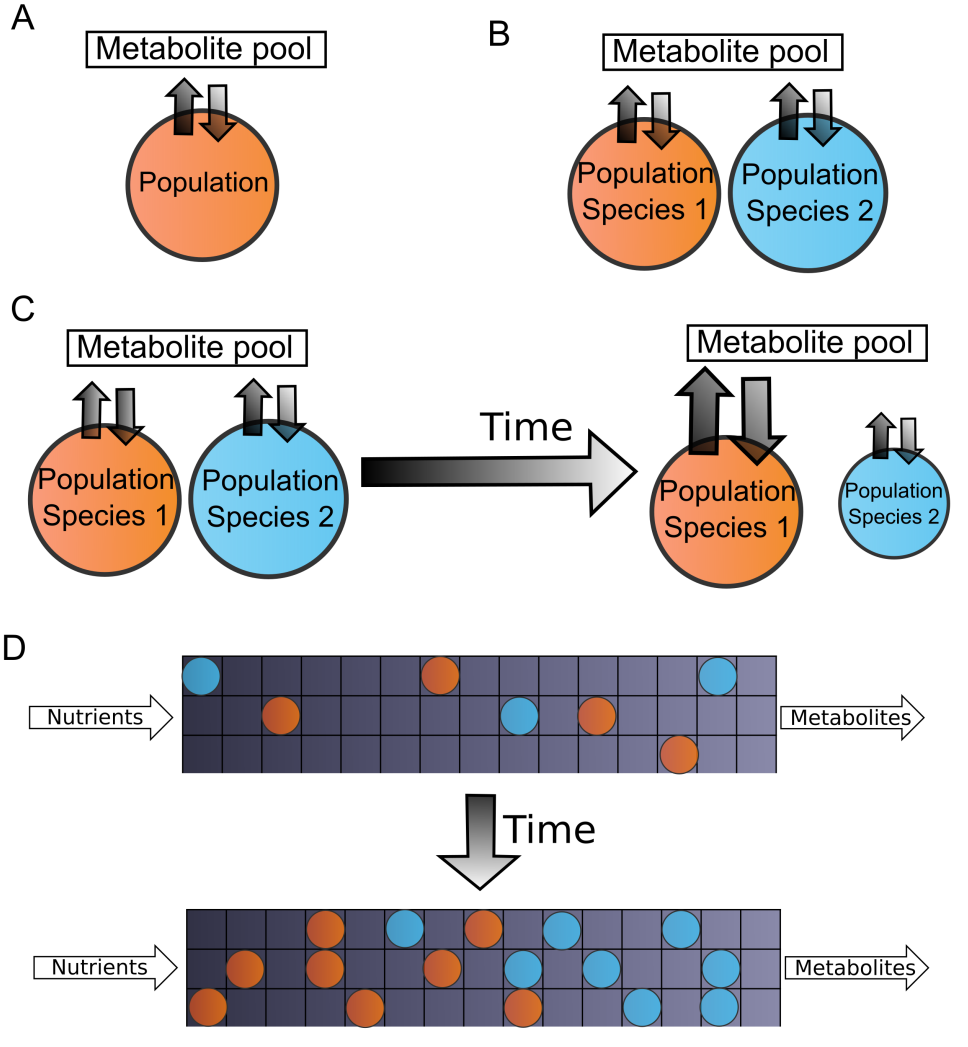


Figure 1.4: Schematics of several ways of modelling bacterial metabolism
(A) A single population exchanging metabolites with a metabolite pool in steady state (B) Two populations exchanging metabolites with the same metabolite pool in steady state (C) Two populations exchanging metabolites with the same metabolite pool in non-steady state, i.e. with dynamic FBA (D) Two species, each with many small populations, on a spatial grid that regulates the exchange of metabolites in non-steady state, i.e. with dynamic FBA. Populations may also move across the grid.

assumed to be in a quasi-steady state [117]. This means that there is no accumulation of nutrients or metabolites within the cells, so that all nutrients taken up are converted into either growth or metabolites. By updating the metabolite pool and population sizes each timestep based on the quasi-steady state solution, and then recalculating the FBA, changes over time can be modelled for one ore more populations [118] (Fig. 1.4C). When multiple species are used this is known as dynamic multi-species metabolic modeling (DMMM)[118]. DMMMs are capable of capturing ecological processes that vary over time, but the spatial dimension is also important in many ecological contexts [119, 46, 120]. To represent this, spatial DMMMs have been developed [121, 122] (Fig. 1.4D). These models also include a spatial dimension that influences how bacterial populations can interact over time [121, 122]. Bacterial populations have specific locations, and exchange of metabolites between populations depends on diffusion and flow [121]. Bacterial populations can also move across locations, so that spatial separation can occur. The addition of spatial mechanics allows the system to model complex competitive relationships between bacterial colonies that would not exist in a well-mixed condition [121]. This approach also allows for the modelling of the spatial dimension of the gut, which we will discuss next.

1.2.3 Modelling the gut microbiota

From the spatial DMMM approach, several models have been developed that provide explanations or predictions for observed effects in the gut microbiota. Van Hoek and Merks [122] based all species in their DMMM on a *Lactobacillus plantarum* genome-scale metabolic model enhanced with other common gut bacteria pathways. They let this model evolve to fill several niches in a spatial simulation of the adult gut [122]. These evolved species also displayed spatial variation, as cross-feeding took place more distally in the gut than primary consumption. The BacArena framework was the first to model the gut microbiota with a variety of species-specific GEMs [123]. This allows the model to make predictions on the metabolic and spatial roles of specific species in the microbiota. Specifically, BacArena predicted that the presence of intestinal mucins drives bacterial diversity, and that a gradient of mucin consumption exists in the gut [123]. Later work with the SteadyCom framework reproduced the spatial variation between facultative and strict anaerobes in the adult human gut, but not the Bacillota/*Bacteroides* ratio typically observed [124]. In this thesis, we will apply the DMMM modelling approach to the infant gut microbiota to create

hypothesis and make predictions specific to the infant gut. We base our model on the spatial model of Van Hoek and Merks [122], which we combine with the GEMs of the AGORA database [10].

1.3 Thesis outline

The thesis is structured as follows. In **chapter 2** we introduce the multiscale mathematical model of the infant gut microbiota. We use the model to examine the role of factors such as enzyme limitations and oxygen availability in shaping succession in the infant gut. In **chapter 3** we expand the model to include the prebiotic oligosaccharides galacto-oligosaccharides (GOS) and 2'-fucosyllactose (2'-FL), and examine how the addition of these prebiotics influence the model predictions. We focus in particular on the possible effects on the health-promoting *Bifidobacterium* species and butyrate producers. In **chapter 4** we further expand the model by including mucins, and the extracellular digestion of mucins, GOS, and 2'-FL. This extracellular digestion allows for cross-feeding to occur in new ways. We examine why breastfeeding may reduce mucin consumption in the infant gut, and explain the different abundance of species through differences in their metabolism, in particular whether or not they produced public goods. In **chapter 5** we examine the effects of disturbances by factors such as antibiotics on the infant gut microbiota, how the microbiota recovers from these disturbances, and how this may be influenced by prebiotic oligosaccharides. Finally, in **chapter 6** we discuss the methods and findings of the thesis in a broader context. We discuss the assumptions and weaknesses of our modelling approach, and give recommendations on future improvements and explorations.

Chapter 2

A multiscale spatiotemporal model including a switch from aerobic to anaerobic metabolism reproduces succession in the early infant gut microbiota

Authors

David M. Versluis[a], Ruud Schoemaker[b], Ellen Looijesteijn[b], Daniël Muysken[d], Prescilla V. Jeurink[b], Marcel Paques[b], Jan M. W. Geurts[b], Roeland M. H. Merks[a,c]

Affiliations

[a] Leiden University, Institute of Biology, Leiden, The Netherlands

[b] FrieslandCampina, Amersfoort, the Netherlands

[c] Leiden University, Mathematical Institute, Leiden, The Netherlands

[d] Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands

Published in mSystems Vol. 7 No. 5 (2022)

doi.org/10.1128/msystems.00446-22

2

Abstract

The human intestinal microbiota starts to form immediately after birth and is important for the health of the host. During the first days, facultatively anaerobic bacterial species generally dominate, such as Enterobacteriaceae. These are succeeded by strictly anaerobic species, particularly *Bifidobacterium* species. An early transition to *Bifidobacterium* species is associated with health benefits; for example, *Bifidobacterium* species repress growth of pathogenic competitors and modulate the immune response. Succession to *Bifidobacterium* is thought to be due to consumption of intracolonic oxygen present in newborns by facultative anaerobes, including Enterobacteriaceae. To study if oxygen depletion suffices for the transition to *Bifidobacterium* species, here we introduced a multiscale mathematical model that considers metabolism, spatial bacterial population dynamics, and cross-feeding. Using publicly available metabolic network data from the AGORA collection, the model simulates *ab initio* the competition of strictly and facultatively anaerobic species in a gut-like environment under the influence of lactose and oxygen. The model predicts that individual differences in intracolonic oxygen in newborn infants can explain the observed individual variation in succession to anaerobic species, in particular *Bifidobacterium* species. *Bifidobacterium* species became dominant in the model by their use of the bifid shunt, which allows *Bifidobacterium* to switch to suboptimal yield metabolism with fast growth at high lactose concentrations, as predicted here using flux balance analysis. The computational model thus allows us to test the internal plausibility of hypotheses for bacterial colonization and succession in the infant colon.

2.1 Introduction

The human infant microbiota starts forming directly after birth, and differs greatly from the adult microbiota [16]. Three main community types are observed in the infant: (a) a Proteobacteria-dominated microbiota; (b) an Actinobacteria-dominated microbiota; and, less frequently (12-14% in one study [33]), (c) a Bacilli-dominated microbiota [16]. The three main community types are established after an ecological succession of early communities. This ecological succession is thought to be controlled by nutrition and early oxygen in the colon. To develop hypotheses on the potential mechanisms and controlling factors of the initial development of the human microbiota, in particular the role of early oxygen, we introduced a computational model.

In the first 24 to 48 hours after birth, Proteobacteria, including *Escherichia coli* and *Enterobacter cloacae*, and Bacilli, including *Streptococcus*, *Lactobacillus*, and *Staphylococcus*, are the most common [8, 16]. In the following weeks Proteobacteria are often replaced by anaerobic Actinobacteria, mainly *Bifidobacterium* species, whereas Bacilli are succeeded by either Proteobacteria or Actinobacteria. Other anaerobic species are also found, but typically do not dominate. Actinobacteria generally persist as the dominant group until weaning [16]. A possible trigger for the replacement of the Enterobacteriacea by *Bifidobacterium* spp., is depletion of a hypothesized initial amount of oxygen [125, 14]. Oxygen diffuses into the gut lumen from the body and is taken up by bacteria, by colonocytes, and by non-biological chemical processes in the cecal contents [126, 127, 128] leading to oxygen depletion. The relative importance of these processes is still under debate [127, 128].

Bifidobacterium species [16] are associated with a range of health benefits. For example, early succession to a *Bifidobacterium* dominated microbiota is correlated with reduced probabilities to be underweight at 18 months of age or to experience colic [33, 15]. Acetate and lactate produced by *Bifidobacterium* also contribute to the acidification of the infant gut, which suppresses many pathogens [18]. A key question, therefore, is if and how *Bifidobacterium* can be promoted. The composition and development of the infant microbiota is thought to be determined by many factors, including nutrition [3, 129, 53], i.e. human milk or infant formula. The presence of lactose, the most abundant carbohydrate in human milk and most infant formulas, is essential for acquiring high levels of *Bifidobacterium* [75]. This suggests that nutrition can play a crucial role in promoting *Bifidobacterium*.

To obtain more insight into the potential ecological and metabolic mechanisms

underlying bacterial colonization and succession in the infant colon, here we propose a multiscale, spatial computational model of the infant gut microbiota. Earlier work has introduced similar models of the adult gut microbiota [130, 131, 123, 122]. Those models make use of flux balance analysis (FBA) to model metabolism. Given a number of constraints, including substrate availability and enzyme availability, FBA uses genome-scale metabolic network models (GEMs) to predict optimal biomass or energy production and the exchange of metabolites with the environment. In spatial FBA, FBA models representing small subpopulations of bacteria are coupled together to model a microbial, metabolizing ecosystem. While some approaches assume the whole system is in steady-state, allowing them to use the optimization approach for the whole ecosystem [131], here we build upon dynamic approaches such as BacArena [123], COMETS [132], and the approach by Van Hoek and Merks [122]. In these models, the metabolite exchange rates given by FBA are used to dynamically change metabolite concentrations in the environment, and predict the resulting effects on the bacterial populations. These properties make these dynamic, spatial FBA approaches highly suitable for our purpose of modeling the formation of the infant gut microbiota.

To analyze succession of bacterial species in a dynamic ecosystem such as the infant gut, it is key to accurately describe the dependency of bacterial growth rate and metabolite exchange on environmental metabolite concentrations. Previous work has shown that these are well described by constraints on the metabolic capacity of bacterial cells. These are given by an 'enzymatic constraint' which poses a limit on the summed metabolic flux through the whole cell [100], or through more advanced methods such as flux balance analysis with molecular crowding (FBAwMC) [133], which weighs the flux constraint according to the efficiency and volume of the individual enzymes [112]. Our model is, therefore, based upon the previous approach by Van Hoek and Merks [122], which uses FBAwMC. However, because such detailed data on enzyme efficiency and volume is unavailable for the majority of GEMs, here we have replaced FBAwMC with the simpler 'enzymatic constraint' approach [100]. This approach allows us to model metabolic limitations and substrate-dependent switches in metabolism.

The model represents the first 21 days of development of the infant microbiota in a dynamic, spatially extended, gut-like environment, and considers three scales. At the microscopic scale, the model simulates individual metabolites. At the mesoscopic scale, metabolism and growth of bacteria occurs as a function of the local availability

of metabolites. The local environment is depleted by the bacterial models, and receives the metabolites they deposit. This will influence the availability of metabolites in the next timesteps. The macroscopic scale represents the whole ecosystem. At this level a diffusion model simulates the spread of metabolites between adjacent lattice sites, and a local mixing model simulates mobility of bacterial populations. The metabolites also undergo advection distally to represent luminal movements, while to represent adherence to the mucus the bacterial populations do not advect. Bacterial populations are also randomly deleted to represent local extinction, and they can form new populations in their immediate environment. The system currently considers 15 bacterial species based on 15 GEMs taken from the AGORA collection [10], a database of 773 semi-automatically reconstructed GEMs of human gut bacteria that includes all major infant species [10].

The infant gut microbiota is partially limited by carbohydrates. Prebiotic carbohydrates increase population sizes *in vitro* and in mice inoculated with infant gut bacteria [62, 134]. The primary carbohydrate in human milk and nearly all infant formula is lactose, and infant formulas without lactose lead to a different infant microbiota with greatly reduced *Bifidobacterium* abundance [75]. Therefore to a first approximation, we ignore key nutrient sources such as human milk oligosaccharides, fats, protein residues, or intestinal mucus and focus on carbohydrate metabolism, here taking lactose as the only input carbon source in the model.

In comparison with the previous approach [122] upon which the present work builds, the technological advance is that the present framework simulates a dynamic ecology of a large diversity of bacterial species. To this end, the modeling framework can, with minor modifications, import any GEM written in SBML, such as those in the AGORA database that we use in this project [10]. The previous system [122] could only describe a 'superbacterial' metabolic network representative of the adult microbiota, while bacterial diversity was simulated by activating or repressing individual pathways. In the present work we studied a selection of species representative of the infant microbiota. The new system also allows us to analyze and visualise the flow of fluxes through the emergent network of dynamically interacting spatially distributed populations. The conceptual advance following from this technological advance is that we can now predict the effect of environmental conditions (oxygen, food, etc.) on the relative abundance of individual bacterial species. This advance has allowed us to use the system to demonstrate the mechanistic consistency of the hypothesis [125, 14] that a transition from an Enterobacteriaceae-dominated micro-

biota to a *Bifidobacterium* spp.-dominated microbiota is explained through gradual depletion of oxygen leading to anaerobiosis. The model suggests that high levels of initial oxygen can lead to prolonged dominance by Enterobacteriaceae, even after oxygen has been depleted, possibly explaining observed variation in the composition of the infant microbiota [16, 53], and reveals possible cross-feeding interactions between the species. Altogether the model proved to be a useful tool for qualitatively evaluating hypotheses on the dynamics of the infant gut microbiota.

2.2 Results

2.2.1 Model outline

The model of the infant colon is discussed in detail in Section Methods. Briefly, given a typical infant colon length of 45 cm [135] and diameter of approximately 1.5 cm [136] the colon is described on a regular square lattice of 225×8 boxes of 2 $mm \times 2$ mm each, thus reaching a good balance between computational efficiency and accuracy (Fig. 2.1A). Because bacterial populations can vary at a much smaller scale [137, 138], the boxes are purely used as a computational discretization. Each lattice site can contain a population of one species, representing only the locally dominant bacterial species. It is described by a GEM from a list of 15 species prevalent in the earliest infant gut microbiota ([139, 3] selected from the AGORA collection [10]) (Table 2.1, Methods section' Species composition').

A simulation proceeds as follows. We initialize the system by placing approximately 540 small populations of the 15 species randomly across the lattice (Fig. 2.1B). We run the model in timesteps representing three minutes of real time, these timesteps proceed as follows: To mimic carbohydrates entering the colon from the small intestine, a pulse of lactose is introduced every 60 timesteps to the lattice sites of the six most proximal columns (Fig. 2.1A-1). Then we predict the metabolism of local bacterial populations using FBA. To mimic metabolic trade-offs, we use an enzymatic constraint that sets the maximum summed flux for each FBA solution. The uptake bounds for each population are set according to the local concentrations of metabolites (Fig. 2.1C&D). FBA returns a solution that maximizes the rate of ATP production, which is a suitable proxy for the rate of biomass production [140] given the focus on carbohydrate metabolism. This biomass production rate (indicated by brightness in Fig. 2.1B) is used to update the local population size. Populations

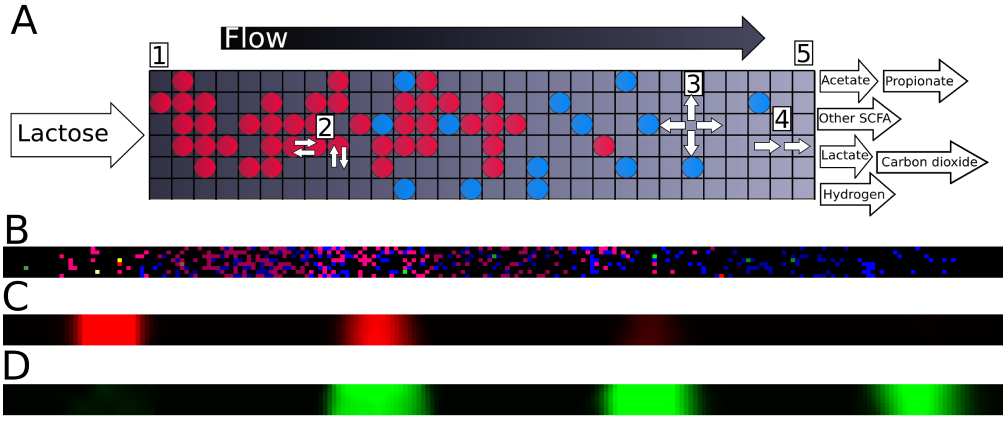


Figure 2.1: The multi-scale metabolic model
(A) Schematic of the model at a system level. Circles represent bacterial populations, color represents species. Flow through the tube is from left (proximal) to right (distal). Lactose is placed at the proximal side. Output metabolites are examples, and depend on bacterial metabolism. Lattice dimensions and ratio are schematic. (B) Screenshot of the bacterial layer of the model at a single time point. Color indicates species, brightness indicates growth rate on the current timestep. (C,D) Screenshots showing, respectively, lactose and lactate concentration of the model corresponding to B. Intensity indicates concentration.

spread into an available adjacent lattice site once the population size has exceeded a threshold. The FBA solution also returns a set of metabolite exchange rates, which are used to update the local metabolite concentrations (Fig. 2.1D). In total, 723 extracellular metabolites are described in the model, but only 17 of these are typically found in more than micromolar amounts in our simulations (Methods, Table 2.2). Bacterial populations diffuse through swaps of the content of adjacent lattice sites (Fig. 2.1A-2). Metabolites diffuse between adjacent lattice sites (Fig. 2.1A-3) and advect to more distal sites (Fig. 2.1A-4) over one lattice per timestep, reaching transit times of effectively 11 hours [141, 142]. Metabolites and populations that reach the distal side of the system (Fig. 2.1A-5) are removed. To mimic the appearance of new bacterial populations (e.g. from ingested spores or introduction from intestinal crypts [143]) each empty lattice site now has a small probability to be filled with a new bacterial population selected at random from all available species (Table 2.1). This completes the timestep, after which we record the metabolite concentrations and bacterial population in the whole model. All parameters are given in Section Parameters. We have also performed a sensitivity analysis for several pa-

rameters around the default parameter set shown in Table 2.3 (Section Sensitivity analysis).

We performed several checks to ensure the biochemical and thermodynamic plausibility of the model (Methods section 'Checking the validity of the GEMs'). After the resulting corrections to the GEMs (S1 Table), no atoms were generated or lost by the FBA solutions, there was no flux through the objective function without a substrate, and 99.999% of all fluxes of Fig. 2.3A contained less free energy in the output than in the input. The remaining 0.001% all had no energy increase and growth rates less than 0.001% of the average growth rate. We therefore concluded that the system conserved mass and energy.

2.2.2 Enzymatic constraint reproduces metabolic switching in *Bifidobacterium*

Many microorganisms switch between metabolic pathways of higher and lower yield depending on substrate availability. These so called metabolic switches can be reproduced with FBA through an enzymatic constraint [100, 144, 109, 112] that limits the summed metabolic flux of a bacterium. *Bifidobacterium* also displays a metabolic switch through a pathway known as the bifid shunt. At low extracellular sugar concentrations, *Bifidobacterium* produces a mixture of acetate, ethanol, and formate in

Table 2.1: Species included in our model, all from the AGORA collection[10]. Color indicates color used in figures

Species	Family	Anaerobic? [10]	Reactions
Bacteroides vulgatus	Bacteroidaceae	Yes	2474
Bifidobacterium breve	Bifidobacteriaceae	Yes	1987
Bifidobacterium longum ssp. *infantis*	Bifidobacteriaceae	Yes	1005
Bifidobacterium longum ssp. *longum*	Bifidobacteriaceae	Yes	2043
Blautia hansenii	Lachnospiraceae	Yes	2055
Collinsella aerofaciens	Coriobacteriaceae	Yes	920
Dorea formicigenerans	Lachnospiraceae	Yes	2072
Parabacteroides distasonis	Porphyromonadaceae	Yes	2519
Ruminococcus gnavus	Lachnospiraceae	Yes	2225
Veillonella dispar	Veillonellaceae	Yes	1124
Enterococcus faecalis	Enterococcaceae	Facultative	1174
Enterobacter cloacae	Enterobacteriaceae	Facultative	1800
Escherichia coli SE11	Enterobacteriaceae	Facultative	2356
Lactobacillus gasseri	Lactobacillaceae	Facultative	1163
Streptococcus salivarius	Streptococcaceae	Facultative	1190

a 1:1:2 ratio from pyruvate. At increased sugar concentrations the bifid shunt instead reroutes pyruvate to lactate. In addition acetate is produced from acetyl-phosphate independently, such that at low concentration of sugars *Bifidobacterium* produces acetate, ethanol and formate in a 4:1:2 ratio, and at increased sugar concentrations *Bifidobacterium* switches to production of acetate and lactate in a 3:2 ratio [145, 90].

To test if the enzymatic constraint sufficed to explain the metabolic switch in *Bifidobacterium*, we performed simulations of a single timestep in increasing lactose concentrations (Fig. 2.2A). These correctly predicted the metabolic switch. The relative production rate of each metabolite (Fig. 2.2A) before, during, and after the metabolic switch, as well as the pathways used (S1 Figure), matched the experimental observations [145, 90]. Simulations without the constraint only reproduced the situation at low concentrations of lactose (S2 Figure A), showing that the enzymatic constraint is required. The model also predicted an unobserved metabolic switch on lactose in *E. coli* (S2 Figure B) and reproduced the metabolic switch in *Bacteroides vulgatus* [146] (S2 Figure C). Another effect of the enzymatic constraint is that the growth rate saturates at a species-dependent concentration (Fig. 2.2B, S2 Figure D). As a consequence the relative growth rates of the species depended on the lactose concentration.

2.2.3 Model predicts *Bifidobacterium* dominance through metabolism under anaerobic conditions

After having established the effect of the enzymatic constraint in models of individual bacterial populations, we studied the behavior of the full multiscale model for anaerobic conditions. After 21 days (10080 timesteps) *Bifidobacterium* spp. had become the most abundant in 27 of 30 simulations (Fig. 2.3A, S1 Video), a typical situation found in infants of a corresponding age [3, 33]. Also in agreement with *in vivo* data, the model predicted the presence of populations of *E. coli*. Our model also predicted populations of *Blautia hansenii*, which is only occasionally present in the infant colon [3]. We will investigate the niche of this species further in the section 'Non-*Bifidobacterium* species benefit from lactate consumption'. All species other than the aforementioned had abundances either lower or not significantly higher than their initial abundance after 21 simulated days ($p>0.05$, two-sample t-test). To test if these species were required for the formation of the community we ran an additional 30 simulations in which we only included species that had an abundance

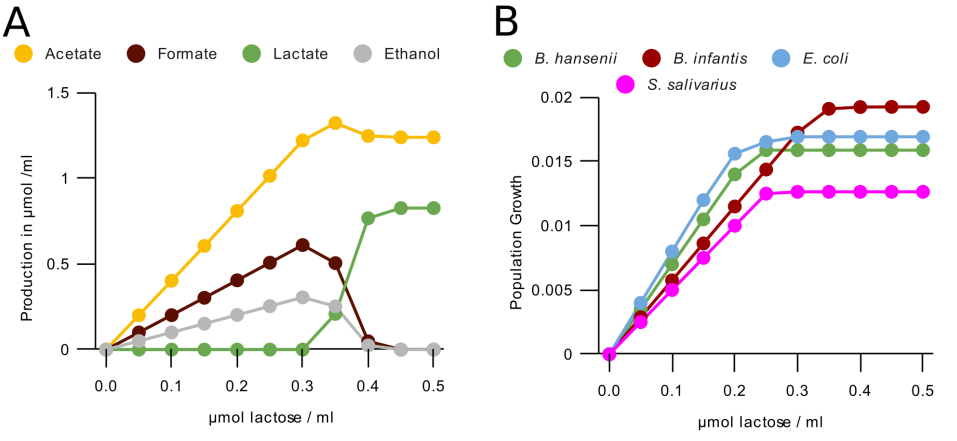


Figure 2.2: Enzymatic constraint reproduces metabolic switching in *Bifidobacterium*
(A) Production of metabolites per timestep by a *B. longum infantis* population of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) using our FBA with enzymatic constraint method. (B) Growth per timestep by lactose concentration for populations of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) of some major bacterial species using our FBA with enzymatic constraint method. *B. longum infantis* is used as the *Bifidobacterium* representative.

significantly larger than their initial abundance at the end of other simulations. This led to a very similar set of outcomes (S3 Figure A). In these simulations none of the species abundances differed from the original simulations when all species were included ($p>0.05$, two-sample t-test). In the remaining simulations, all species were included in order to prevent an initial bias, thus leaving in the possibility that any of the rare species become abundant under one of the conditions considered. In practice, this did not occur.

We next analysed the metabolites diffusing or advecting out of the distal end of the gut over the last two days of the simulations of Fig. 2.3A. We found considerable amounts of acetate, formate, lactate, hydrogen, ethanol, and succinate (Fig. 2.3B). All other metabolites were present in much smaller quantities. The high abundance of acetate and lower abundance of lactate agree with the fecal composition of three month old formula-fed infants [147], and is characteristic of microbiota dominated by *Bifidobacterium*. Formate, hydrogen, ethanol and succinate are also all reported in infant feces [31, 148, 149, 150].

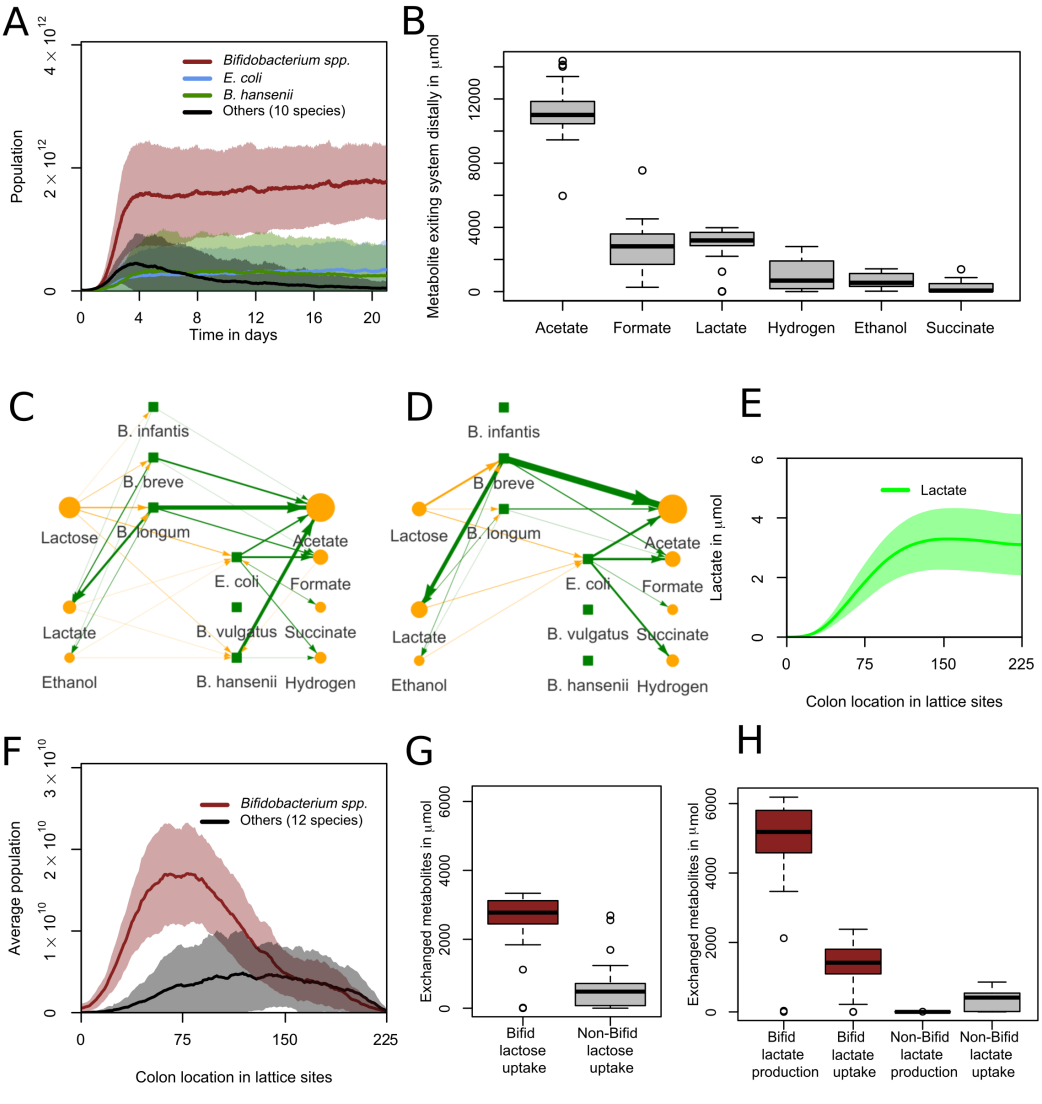


Figure 2.3: ***Bifidobacterium*** **spp become dominant anaerobically and mostly proximally with a unique metabolic profile**
(A) Abundances for grouped species over 21 days. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (B) Distribution of metabolites exiting the system distally over the last two days (960 timesteps) of the simulations (n=30). (C,D) Visualisation of metabolic interactions in the sample run of S1 Video. Green lines represent production, yellow lines represent consumption. Line width and intensity are proportional to the amount exchanged with the environment, with a threshold of 0.5 µmol, with no normalization. Metabolite circle size is relative to the most abundant metabolite, with a minimum displayed size of 26 pixels. Data from hour 90 to 93 (step 1800 to 1860) for C and last 3 hours (step 10020 to 10080) for D (E) Spatial distribution of lactate over the last two days (960 timesteps) of the simulation. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (F) Abundance per location for grouped species over the last two days (960 timesteps). Curve shows mean value and shaded area one standard deviation over n=30 simulations. (G) Metabolism of lactose by grouped species. All data from last two days. n=30. (H) Metabolism of lactate by grouped species. All data from last two days. n=30.

To examine how this metabolic pattern is formed, we analysed the metabolic production and uptake in a sample run (Fig. 2.3C). *Bifidobacterium* converts lactose into a mix of acetate, lactate, ethanol, and formate, as predicted (Fig. 2.2A), with a higher lactate production than ethanol or formate. Some of this lactate is reabsorbed in later timesteps and converted into the three other metabolites by *Bifidobacterium*. A mixture of lactose and cross-feeding products is converted by *E. coli* and *B. hansenii* into the remaining metabolites of Fig. 2.3B. The main cross-feeding nutrient absorbed is lactate, produced by *Bifidobacterium*, but there is also uptake of ethanol, formate, acetate, and succinate. These are also produced by the cross-feeders. The network has become simpler at the end of this run, as *B. hansenii* is at a very low abundance and *E. coli* now only consumes lactose, lactate, and ethanol, which it does not produce (Fig. 2.3D).

2.2.4 Cross-feeding emerges in simulations

The observed cross-feeding in the sample run (Fig. 2.3C&D) suggested that production of metabolites by *Bifidobacterium* spp. may be crucial for *E. coli* and *B. hansenii*. To further analyse such cross-feeding interactions we next analysed the spatial distribution of lactose and the Bifidobacterial products lactate and acetate. From day 19 to 21, lactose was present mostly proximally (S3 Figure B), lactate was most abundant in the middle of the colon (Fig. 2.3E), and acetate was most abundant distally (S3 Figure C). To analyse the role of *Bifidobacterium* spp. in shaping this metabolic pattern, we analysed the location of populations from day 19 to 21 (Fig. 2.3F). *Bifidobacterium* spp. was located proximally, whereas the other species were located more distally. These results suggested that separated metabolic niches had formed in the simulated colon. To quantify cross-feeding interactions we summed the flux of lactose and lactate for *Bifidobacterium* and for all non-*Bifidobacterium* species grouped together. *Bifidobacterium* spp. was the largest consumer of lactose (Fig. 2.3G) and both the largest producer and consumer of lactate (Fig. 2.3H). An additional consumption of lactate towards acetate, formate and ethanol takes place at lower lactose concentrations, when the enzymatic constraint is not saturated by lactose uptake. However, the non-*Bifidobacterium* groups consumed more lactate relative to their lactose consumption. To test the effect of lactate uptake by *Bifidobacterium*, additional simulations were run in which the conversion from lactate to pyruvate was blocked, preventing *Bifidobacterium* from consuming lactate. In these

simulations *Bifidobacterium* was the most abundant group in 20 of the 30 simulations, with the largest average abundance (S3 Figure D). Thus lactate consumption played a small role but was not the primary cause of *Bifidobacterium*'s dominance. We hypothesized that *Bifidobacterium*'s primary niche is as primary consumer, consuming lactose, while for other species niches exist as primary and/or secondary consumer around consuming lactose or lactate. Combined with the observation (Fig. 2.3E) that lactate concentration is lower in the distal colon and the recorded cross-feeding fluxes in our network visualisation (S1 Video), we can conclude that the population at the distal end of the simulated colon consists at least partially of such secondary consumers, i.e. cross-feeders.

2.2.5 Lactose uptake through the bifid shunt is essential for *Bifidobacterium* dominance

Next we determined whether the bifid shunt is essential to *Bifidobacterium* dominance in the model. We disabled the bifid shunt by blocking all flux through the reactions associated with fructose-6-phosphate phosphoketolase and ran 30 simulations. Fructose-6-phosphate phosphoketolase is unique to and necessary for the bifid shunt, and characteristic of *Bifidobacterium* as a genus [151]. *Bifidobacterium* spp. were dominant in none of these simulations (Fig. 2.4A, S2 Video). Instead, *E. coli*, *B. vulgatus*, and *B. hansenii* were dominant. This indicates that the bifid shunt is crucial to the dominance of *Bifidobacterium* spp. in our model. We also ran 30 simulations in which only the production of lactate was disabled in *Bifidobacterium* (S3 Figure F). 27 of these simulations still led to *Bifidobacterium* dominance. This indicates that the metabolic switch to lactate production was not essential to *Bifidobacterium* dominance in our model. *B. vulgatus* also had a higher average abundance, but this was not significant (p=0.27, two-sample t-test).

2.2.6 *B. hansenii* but not *E. coli* can sustain itself on cross-feeding

We observed that non-*Bifidobacterium* species appeared in the simulation behind a proximal population of *Bifidobacterium* spp. (Fig. 2.3F). These non-*Bifidobacterium* species consumed lactose (Fig. 2.3D&G), despite the fact that lactose concentrations have dropped compared to the proximal part (S3 Figure B). We therefore

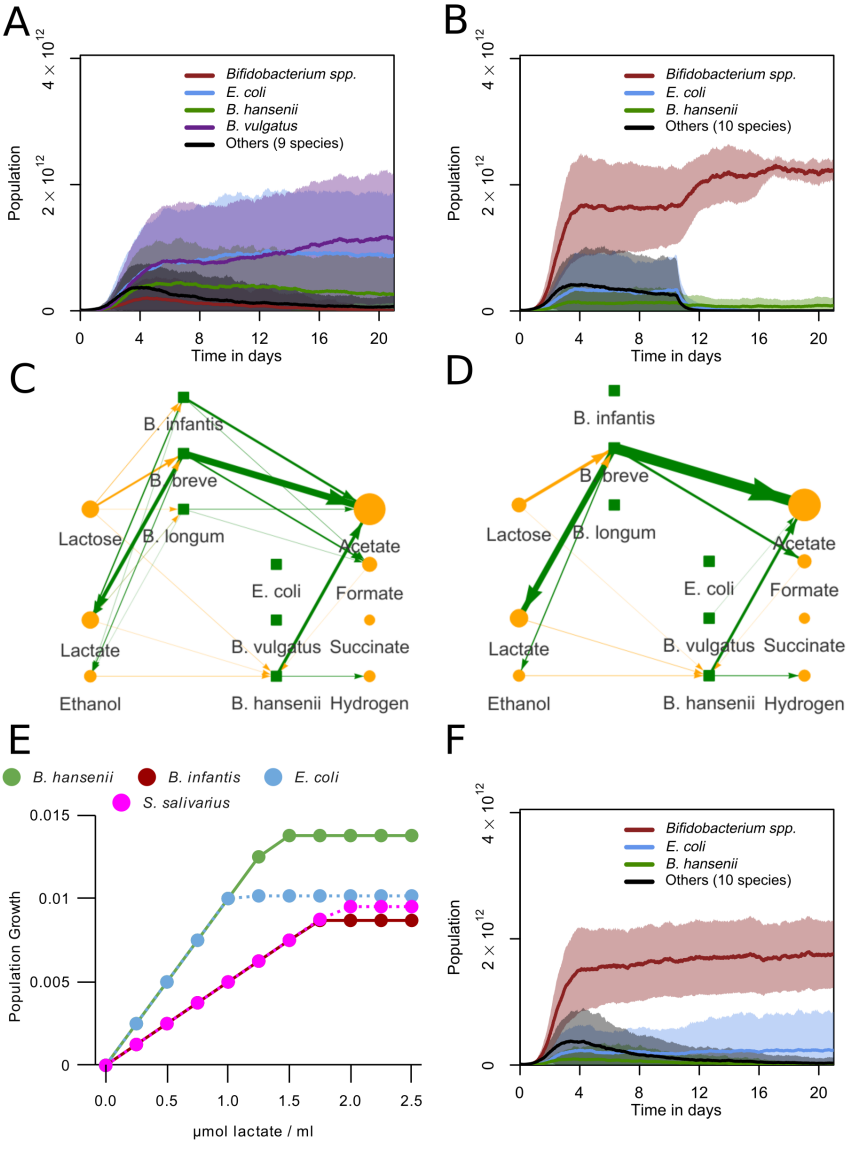


Figure 2.4: *Bifidobacterium* and *E. coli* require lactose, but *B. hansenii* requires lactate

(A) Abundances for grouped species over 21 days. The bifid shunt was disabled for all *Bifidobacterium* species. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (B) Abundances for grouped species over 21 days, but with lactose disabled for non-*Bifidobacterium* species from 5040 timesteps (10.5 days) onwards. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (C,D) Visualisation of metabolic interactions in a sample run. Green lines represent production, yellow lines represent consumption. Line width and intensity are proportional to the amount exchanged with the environment, with a threshold of 0.5 µmol, with no normalization. Metabolite circle size is relative to the most abundant metabolite, with a minimum displayed size of 26 pixels. Data from hour 90 to 93 (step 1800 to 1860) for C and last 3 hours (step 10020 to 10080) for D. (E) Growth per timestep by lactate concentration for populations $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) of some major bacterial species (F) Abundances for grouped species, with lactate consumption disabled for all species. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations.

wondered how these non-*Bifidobacterium* species could persist in the model. At reduced lactose concentrations, *E. coli* produces more ATP per timestep than *Bifidobacterium* spp. (Fig. 2.2B). We hypothesized that non-*Bifidobacterium* species outcompete *Bifidobacterium* at reduced lactose concentrations in the model. To test this hypothesis, we blocked lactose consumption by non-*Bifidobacterium* species after 10.5 days, the midway point of the simulation, and ran 30 simulations. In this way, we ensured that sufficient secondary resources were produced by *Bifidobacterium* in case non-*Bifidobacterium* species made use of them. We observed a sharp population decrease of non-*Bifidobacterium* species starting directly after blocking lactose uptake (Fig. 2.4B). *E. coli* had a near-zero abundance at the end of these simulations. *B. hansenii* also had a significantly lower abundance compared to the baseline ($p<0.01$, two-sample t-test), but still had a presence (S2 Video). Thus, in our simulations, some amount of primary consumption was essential for *E. coli*, but not for *B. hansenii*.

2.2.7 Non-*Bifidobacterium* species benefit from lactate consumption

To analyse how *B. hansenii* could sustain a population as secondary consumer, we investigated the substrates used by non-*Bifidobacterium* species in the model. Analysis of the flow of metabolites between species in a new run showed that the uptake of lactose by *B. hansenii* decreases over time, while the uptake of lactate increased (Fig. 2.4C&D). *E. coli* and *B. hansenii* also produce more ATP from lactate than *Bifidobacterium* at any concentration (Fig. 2.4E). Combined with our earlier observations of cross-feeding on lactate (Fig. 2.3D&H) we hypothesized that *Bifidobacterium* species cannot compete with *E. coli* and *B. hansenii* on pure lactate. Consumption of lactate by *B. hansenii* could be essential to its ability to sustain populations without lactose uptake.

To investigate if lactate is essential in feeding secondary consumer populations, we blocked the lactate uptake reaction for all species and ran 30 simulations. (Fig. 2.4F, S2 Video). At 21 days *B. hansenii* populations were much smaller than in simulations with functional lactate uptake, not significantly larger than the initial populations ($p>0.05$, two-sample t-test). *E. coli* populations were also significantly smaller than those in simulations with lactate uptake, but also significantly larger than their initial population at the start of the run ($p<0.05$, two-sample t-test). *Bifidobacterium*

populations were similar to those in the baseline simulations (Fig. 2.3A). Together with the results in Fig. 2.4B this indicates that *E. coli* does not require lactate as a cross-feeding substrate like *B. hansenii* does.

2.2.8 *E. coli* outcompetes *Bifidobacterium* by taking up early oxygen

During the first days Enterobacteriaceae or Bacilli such as *Streptococcus* are dominant in the *in vivo* infant gut microbiota, after which *Bifidobacterium* spp. are often dominant [16, 3]. This pattern may be explained by the presence of oxygen in the gut shortly after birth [14, 8], e.g. via diffusion through the gut wall [127] or accumulation in the newborn infant gut *in utero* in the absence of microbes [1]. Intracolonic oxygen diffusion depends on host tissue oxygenation and colonic blood flow [152, 153]. To mimic the early presence of oxygen we introduced 0.1 to 10 µmol initial oxygen per lattice site. These values were chosen arbitrarily in absence of precise data, but cover a wide range of outcomes in our model. The oxygen diffused, but did not advect in the model.

0.1 µmol initial oxygen sufficed to reproduce the pattern observed *in vivo*: *E. coli* initially dominated and *Bifidobacterium* became dominant after a few days or weeks of simulated time (Fig. 2.5A, S3 Video). However, we did not see dominance of *Streptococcus* in any of the simulations, in contrast to observations *in vivo* [33]. The initial presence of oxygen strongly affected the other species in the model: *B. hansenii* remained nearly entirely absent, while *B. vulgatus* had a higher abundance than it did without initial oxygen. In the network analysis of a sample run a similar pattern occurs as in the networks of Fig. 2.3C, but with different early species (Fig. 2.5B). Here the Enterobacteriaceae *E. coli* and *Enterobacter cloacea* consume more of the early lactose, along with oxygen. By the end of the run *Bifidobacterium* spp. again took up the role of primary lactose consumer and lactate producer, whereas *E. coli* had fed largely on lactate (Fig. 2.5C), thus assuming a role of secondary consumer.

At 1 or 10 µmol initial oxygen per lattice site *E. coli* and *E. cloacea* remained dominant and the main primary consumers for longer (Fig. 2.5D, S4 Video), and succession to *Bifidobacterium* took much longer, or did not occur at all within the simulated timeframe (Fig. 2.5E). At 10 µmol oxygen per lattice site *Bifidobacterium* either dominated the population or remained much smaller, with few intermediate cases, leading to a bimodal distribution (Fig. 2.5F). This prediction matches *in vivo*

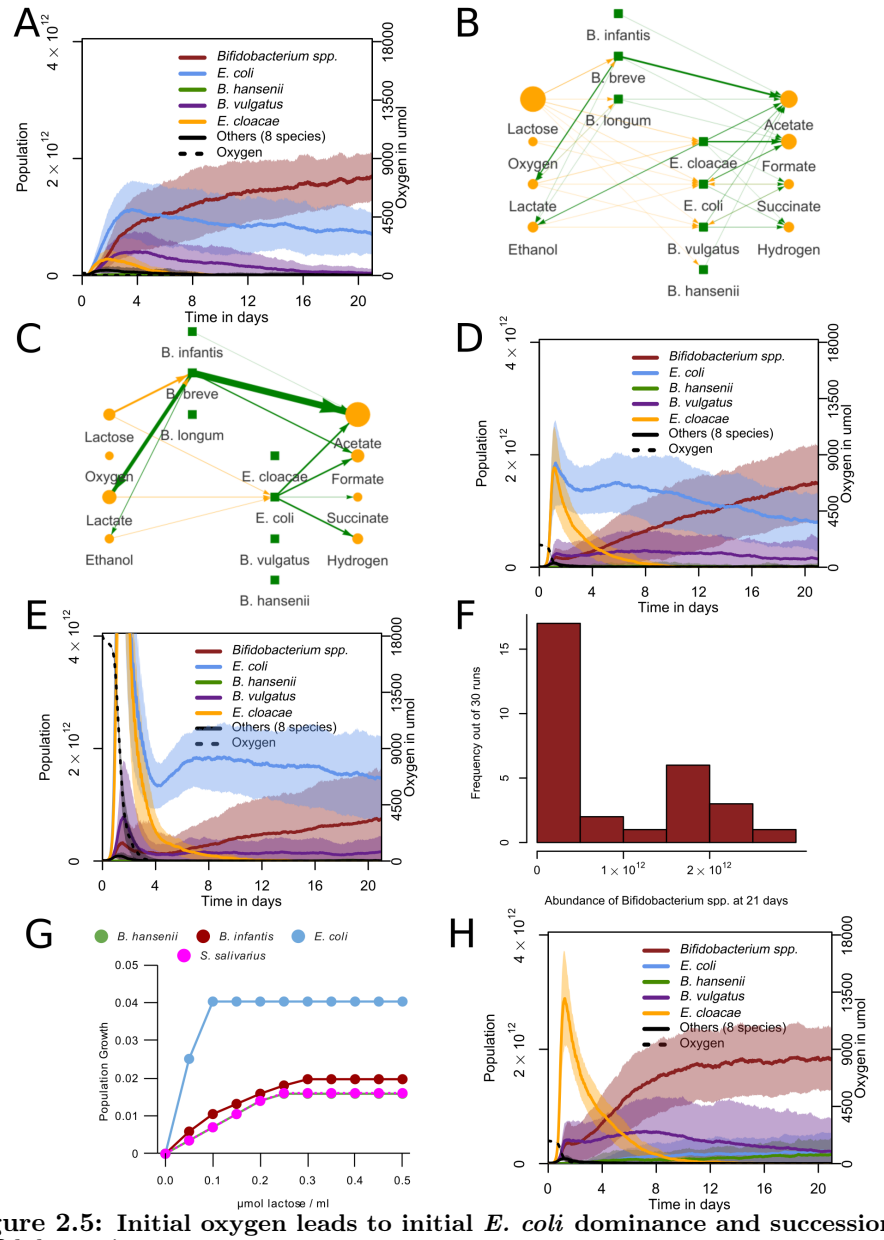


Figure 2.5: Initial oxygen leads to initial *E. coli* dominance and succession by *Bifidobacterium*

(A,D,E) Abundances for grouped species. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations. Initial oxygen per lattice site in µmol is (A) 0.1 (D) 1 (E) 10 (B,C) Visualisation of metabolic interactions in a sample run. Green lines represent production, yellow lines represent consumption. Line width and intensity are proportional to the amount exchanged with the environment, with a threshold of 0.5 µmol, with no normalization. Metabolite circle size is relative to the most abundant metabolite, with a minimum displayed size of 26 pixels. Data from hour 30 to 33 (step 600 to 660) for B and last 3 hours (step 10020 to 10080) for C (F) Distribution of total *Bifidobacterium* abundance at 21 days (10080 timesteps) performed with 10 µmol initial oxygen per lattice site n=30. (G) Growth per timestep by lactose concentration in the presence of abundant oxygen for populations of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) of some major bacterial species (H) Abundances for grouped species with oxygen uptake disabled for *E. coli*. Curve shows mean value, and shaded area one standard deviation, over n=30 simulations.

observations [16, 53]. The proportion of infants dominated by Enterobacteriaceae at 21 days of age varies depending on the study population, but can be 22% [33] or 25% [32] after 21 days, an intermediate value between the 10% (Fig. 2.5D) and the 60% (Fig. 2.5E) that we observe with different oxygen levels. We also initialised four sets of simulations in which oxygen was released from the upper and lower boundaries of the system every timestep (S4 Figure A-D). This led to an increasingly stronger stimulation of *E. coli* with higher levels of oxygen, but did not lead to any of the temporal effects observed *in vivo*. In a separate set of simulations we stopped oxygen release at the midpoint of our simulations. We observed an increase of *Bifidobacterium* spp. and a decrease of *E. coli* after that point (S4 Figure E), leading to a bimodal outcome similar to the simulations with initial oxygen (S4 Figure F).

We further examined the causes of *E. coli* dominance over *Bifidobacterium* in the presence of oxygen in our model. Fig. 2.5G shows that when oxygen is available *E. coli* has a much higher growth per timestep in our model than *Bifidobacterium* spp. for all concentrations of lactose, even though *Bifidobacterium* spp. also produce more ATP than under anaerobic conditions. *B. hansenii* and *Streptococcus salivarius*, in contrast, grew slower than *Bifidobacterium* for all concentrations in the presence of oxygen. This effect depends on the local oxygen concentration, but *E. coli* grows faster than other species even at concentrations much lower than our initial value (S4 Figure G). *E. coli* does not use overflow metabolism in our model, though *E. cloacea* does, and *E. coli* does when in the presence of glucose (S4 Figure H,I,J).

To test if the direct uptake of oxygen was indeed responsible for *E. coli*'s dominance of the microbiota, we disabled the oxygen uptake reaction of *E. coli* for 30 simulations with 1 µmol initial oxygen per lattice site. Under these conditions, *E. coli* failed to dominate the population (Fig. 2.5H). Other species, primarily *E. cloacae*, became dominant early in the simulations as a primary consumer (S4 Video). These species are replaced in all simulations by a population composition similar to that from the simulations without oxygen. In some studies, *Klebsiella* species were the dominant early members of the Enterobacteriaceae instead of *E. coli* [154]. To examine whether our model could also reproduce early dominance by a *Klebsiella* species we initialized the simulations with a set of species containing *K. pneumoniae* instead of *E. coli*. At 0.1 µmol initial oxygen per lattice site, these simulations behaved the same as the original simulations initialized with *E. coli* (S4 Figure K): *Bifidobacterium* became the most abundant group after 21 days in 26 out of 30 simulations.

In simulations initialized with both *K. pneumoniae* and *E. coli*, *K. pneumoniae* became the most abundant in 15 of 30 simulations, and *Bifidobacterium* in the other 15 (S4 Figure L). Altogether, the model predicts that *E. coli* relies on direct consumption of oxygen to dominate the microbiota under initially oxygenated conditions in the model. It also indicates that *E. cloacea* does not have an anaerobic metabolism competitive enough to sometimes become dominant over *Bifidobacterium* spp. as *E. coli* does, and that other Enterobacteriaceae such as *K. pneumoniae* can perform similarly to *E. coli*.

2.2.9 Sensitivity analysis

A number of key parameters in our model were based on reasonable estimates. To test the effects of these parameter values on the simulation results, we performed a sensitivity analysis for these parameters under anaerobic conditions, which best represents the situation at the end of the simulations (Fig. 2.6). As shown in Fig. 2.2B and S2 Figure D, after relaxing the enzymatic constraint *Bifidobacterium* spp. loses its competitive advantage at high lactose concentrations (S4 Video). This leads to reduced dominance or extinction in whole-gut simulations (Fig. 2.6A). Tightening the enzymatic constraint led to *Bifidobacterium* spp. remaining dominant, but with a much smaller population size. Thus the presence of an enzymatic constraint sufficiently large for metabolic shifts, but not its exact level, is crucial for the prediction of *Bifidobacterium* dominance. We set the enzyme constraint at a level strong enough to predict the metabolic shifts observed *in vitro*, but not so strong as to prevent reasonable growth. The ATP required for a unit of population growth had a linear effect on population size, as expected, but had little effect on the relative abundance of *Bifidobacterium* (Fig. 2.6B). The large number of populations with a low ATP requirement also made the model computationally unfeasible. The rate of colonization had little effect on the simulation outcomes: at increased colonization rates, *Bifidobacterium* spp. still dominated the microbiota and its relative abundance remained unchanged (Fig. 2.6C). Even eliminating colonization altogether hardly affected *Bifidobacterium* spp. dominance in most simulations, but led to complete extinction of all *Bifidobacterium* species in 5/30 simulations and of all 12 non-*Bifidobacterium* species in 9/30 simulations. Thus we kept a moderate amount of colonization in the model. We also examined the effect of placing new populations after initialization only in the first column of the lattice, with a probability increased to match the lower

number of eligible squares (S5 Figure A). This also led to a slightly larger but still comparatively small population for the non-*Bifidobacterium* species. To determine the optimal spatial discretization we ran the model on a more coarse and on a more refined lattice, scaling the diffusion, advection, and initialization and colonization parameters accordingly (Fig. 2.6D; see methods section 'Population dynamics'). Lattice refinement did not change the dominance of *Bifidobacterium* nor the spatial distribution of metabolites and populations (S5 Figure B-E), but on a more coarse lattice the smaller number of lattice sites seemed to increase the rate of extinctions and variability of metabolite distributions (S5 Figure F-I). The diffusion of metabolites had a larger effect on the simulation results (Fig. 2.6E). At low diffusion rates, the populations tended to die out, because they had less opportunity to take up lactose as they were exposed only as it 'passed by' due to advection. At high diffusion rates cross-feeders and *Bifidobacterium* consumed all lactate (S5 Figure J); in infant feces substantial amounts of lactate are found [147, 91]. We therefore decided to choose a value for which the system produced the observed concentration of lactate. Finally the amount of mixing of bacteria was varied (Fig. 2.6F). *Bifidobacterium* remained dominant regardless of the diffusion parameter, even when the bacteria were fully mixed (i.e. placed at random locations) each timestep. Without mixing the populations could not spread, for lack of empty space. The bacterial diffusion constant is 1.3 times higher than the metabolic diffusion constant. As shown in Fig. 2.6F the exact setting of the bacterial diffusion has little effect on the outcomes of the simulations, so we have maintained bacterial diffusion at this baseline.

2.3 Discussion

To develop new hypotheses for possible external effects on the initial phases of the infant gut microbiota, we have developed a dynamic multiscale model. Our simulations reproduce the initial dominance of Enterobacteriaceae, particularly *E. coli*, over *Bifidobacterium* species, and the subsequent dominance of *Bifidobacterium* species, out of a broad consortium of infant species. Moreover, our simulations suggest show the consistency of the classical hypothesis that this succession could be due to an initial presence of oxygen, followed by microbial depletion [125, 14]. These predictions on oxygen suggest that higher levels of initial oxygen may explain the absence or strong delay of *Bifidobacterium* development that is observed in some infants [33]. A continuous input of oxygen did not explain the pattern observed *in vivo* (S4 Figure

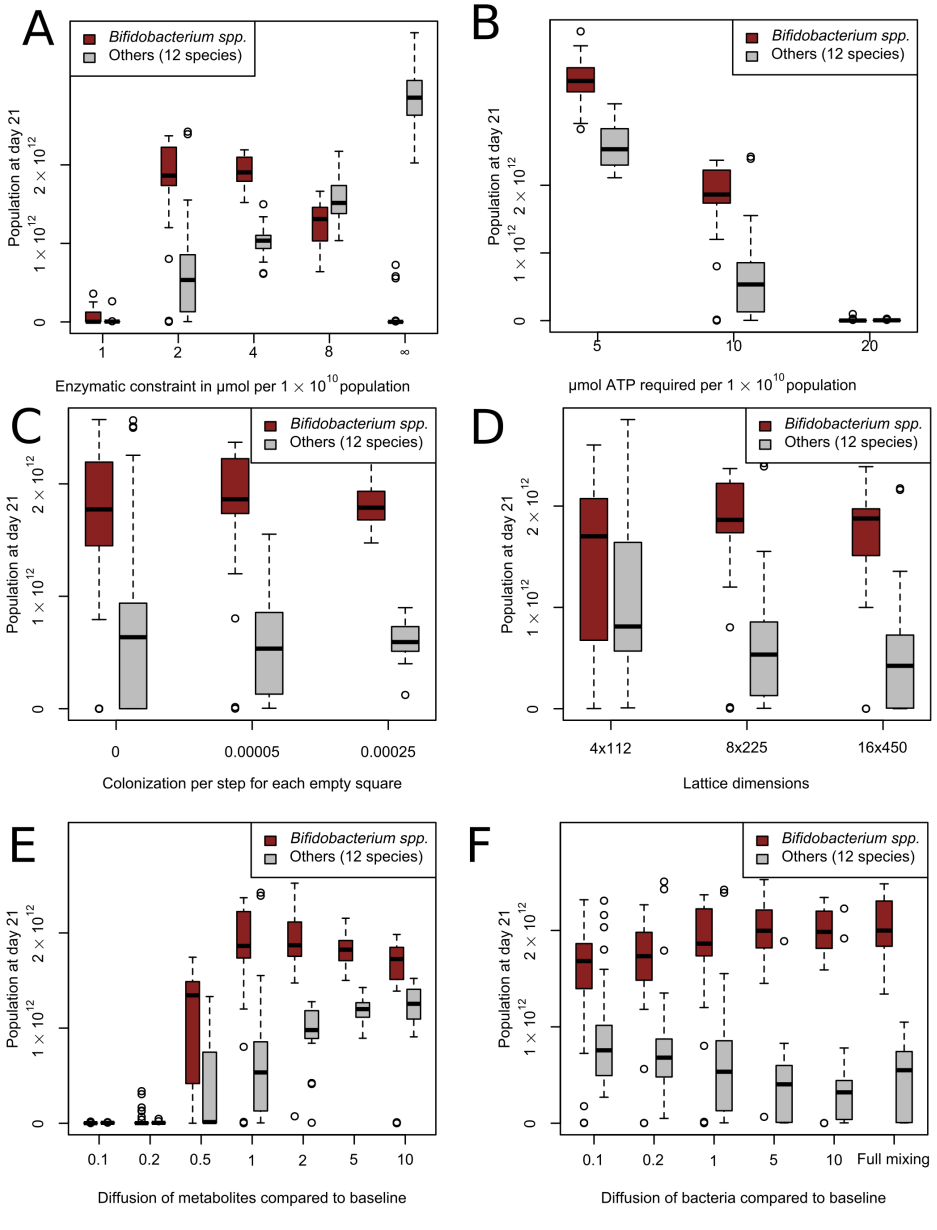


Figure 2.6: Sensitivity analysis for estimated parameters
Abundances for *Bifidobacterium* spp. and other species at the end of day 21 with different parameters. n=30 for each condition. (A) Effect of enzymatic constraint around the default value of 2 µmol flux per timestep per $1 \cdot 10^{10}$ population (B) Effect of the amount of ATP required for growth of $1 \cdot 10^{10}$ bacteria around the default value of 10 µmol (C) Effect of variation in continued influx of populations per timestep per empty square around the default of 0.00005 (D) Effect of variation in width and height of lattice, including adjustment of associated parameters, around the default of 8x225 lattice sites (E) Effect of variation in diffusion for metabolites away from the baseline of $4.4 \cdot 10^{-5}$ cm^2/s (F) Effect of variation in diffusion of bacterial populations away from the baseline of $5.7 \cdot 10^{-5}$ cm^2/s

A-D), but stopping the input of oxygen at the midway point of the simulation did lead to a similar pattern (S4 Figure E-F).

While it is known that there is variation in colonic oxygenation [152, 153], the range of this variation in newborns is not known. In premature infants a positive association exists between the number of days of supplemental oxygen and *Klebsiella* sp. abundance [154]. *Klebsiella* was the primary representative of the Enterobacteriaceae in this study population. However, only very few of these infants developed a *Bifidobacterium* dominated microbiota, regardless of the duration of supplemental oxygen. Nonetheless, it may be desirable to reduce oxygen concentration faster to achieve the positive health effects of *Bifidobacterium* more consistently. As the direct oxidation of lipids and other organic substrates may contribute to creating anaerobic conditions in the colon [127], altered nutrition may allow for a decrease in oxygen in the colon without stimulating facultative anaerobic bacteria such as *E. coli*. There are also indications that short-chain fatty acids produced by *Bifidobacterium* may decrease oxygen levels by stimulating the oxygen use of colonocytes [128]. Future versions of the model could represent these oxygen-depleting mechanisms by localized (e.g., near the intestinal wall) and distributed oxygen consumption terms. At present the model predicts that distributed oxygen usage by such additional processes will speed up the succession to an anaerobic, *Bifidobacterium*-dominated microbiota. The limit case, i.e. extremely fast oxygen consumption, is represented by the anaerobic case (Fig. 2.3). The effect of localized oxygen consumption processes, e.g., near the intestinal wall is a topic of our ongoing research.

Our simulations predict that the dominance of *Bifidobacterium* spp. in the anaerobic infant system could be explained from its consumption of lactose through its bifid shunt metabolism. Our model can correctly represent bifid shunt metabolism due to the implementation of an enzymatic constraint on the maximum total metabolic flux in the bacteria (Fig. 2.2). Both the low and high yield metabolism, and the switch between them based on nutrient availability, have been extensively described *in vitro* [90]. The FBA with enzymatic constraint method we use is similar to flux balance analysis with molecular crowding (FBAwMC) [133]. As in FBAwMC we consider the enzymatic constraint to represent the limited availability of space and production capacity for enzymes within a bacterial cell. However, in the absence of good reaction-specific thermodynamic information for our current model, we do not utilize the reaction-specific technique of FBAwMC. We instead base an enzymatic constraint only on the size of the local population, which is sufficient for modeling

a metabolic switch. This is equivalent to using FBAwMC and placing all crowding coefficients at the same positive number, which has been done previously to model metabolic switches [133]. More broadly, our method is similar to other existing methods that model metabolic switches by including limits on fluxes that represent some physical limit on metabolism, such as proteome constraints or membrane occupancy [155, 144, 156, 118]. It has been shown that metabolic switches in microbes can be modelled in many ways, provided that two simultaneous constraints are in place [112]. In our case, these are the concentration of substrate and the enzymatic constraint.

The metabolic switch of *Bifidobacterium* in our model determines the production of its metabolites. The two most abundant, acetate and lactate, are present in the simulated feces at a median ratio of 3.5:1 (Fig. 2.3B), similar to that found in formula-fed infant feces [147, 91]. We can derive a concentration in mM if we assume each of the eight lattice sites whose metabolites are advected out of the system each step contains 0.05ml of water. This is based on an estimated total volume of 90ml, divided over 1800 lattice sites (Table 2.3). This results in an average concentration of 29 mM acetate and 8 mM lactate in the simulated fecal output analyzed in Fig. 2.3B. These values are close to, and well within the variability, of the average values for acetate and lactate of 31 and 12 mM respectively in the feces of two-week old infants in [91]. The median ethanol:acetate ratio of 1:21 is similar to the ratios of 1:18 and 1:10 reported in term infants [149] and the 1:50 reported in pre-term infants [157]. We use ratios here because the measures per gram dry weight are not available in our model. We observe a lower quantity of ethanol in the simulated fecal output of the simulations where lactate uptake by *Bifidobacterium* was disabled (S3 Figure E).

Bifidobacterium takes up lactate in our model and converts it into pyruvate, catalyzed by lactate dehydrogenase. In our model this pyruvate is then further converted to formate, ethanol, and acetate through the high-yield pathway also used in lactose metabolism. However, strong lactate uptake by *Bifidobacterium* spp. has not been observed *in vitro* (in e.g. [158]). Thus, while the enzymes used for lactate consumption exist in *Bifidobacterium*, it is not known if the pathway is used *in vivo* to the extent predicted by the model. Lactate uptake by *Bifidobacterium* has only little impact on the simulation results (S3 Figure D). Similarly, though growth on lactate by *E. coli* in the presence of oxygen is well documented, this is not the case for the anaerobic condition [105]. Lactate uptake is not essential for *E. coli*, as it can exist in a more purely primary consumer role (Fig. 2.4F). The third consumer of lactate, *B. hansenii*, is present only marginally *in vivo* [3]. This matches with our

results for initially oxygenated conditions, in which *B. hansenii* is also largely absent. However, other species that could fill the same lactate-consuming secondary consumer niche as *B. hansenii* in our model are common *in vivo* [74]. Our selection of species may be further improved to allow the correct species to emerge in this secondary consumer niche. For example, it might be appropriate to add *Eubacterium hallii*, a common infant gut bacterium known to consume *Bifidobacterium*-produced lactate and acetate *in vitro* [82].

More generally, our model predicts a lower diversity overall in both bacterial species and metabolites compared to what is observed *in vivo* [3]. While the abundant species in our model largely match with the abundant species *in vivo*, many other species such as those of the genera *Streptococcus* and *Lactobacillus* have a lower abundance in our model compared to the *in vivo* conditions[3]. We have decided to keep all less abundant species in our simulations, to show how in most cases the correct species out of a broad consortium become dominant. There were no large differences when we ran the simulation without these smaller species (S3 Figure A). Particularly the lack of representation of the Bacilli species in our model outcomes is notable. While some sources report a portion (e.g. 12-14% [33]) of their subjects to be dominated by Bacilli, mainly *Streptococcus*, the *Streptococcus* species in our model never became dominant. We showed that *S. salivarius* has an inferior metabolism on lactose and lactate compared to other species in our model (Fig. 2.2A, Fig. 2.4E, Fig. 2.5G).

The discrepancy between our model and the *in vivo* data in the abundance of less common species such as *Streptococcus* may partially be explained through the focus of our model on carbon dissimilation. Though it would be preferable to represent the whole metabolism of the infant microbiota, including factors such as consumption of amino acids, oligosaccharides and intestinal mucus, the initial version presented here only considers carbon metabolism from lactose. While a more extensive metabolism might have allowed more mechanisms and niches to be discerned, it would also introduce additional free parameters, as there is no clear data on the concentration and uptake of these nutrients. The uptake bounds would have to be set arbitrarily, with uniform values or random sampling. Many substances such as fatty acids or protein residues are not even included as metabolites in the database we use for bacterial metabolism [10], further complicating their introduction to the model. By focusing on carbon metabolism, using ATP production as a proxy for growth rate, and only using lactose as an input nutrient, we can circumvent these problems. ATP production

has been shown to be a good proxy for biomass production in *E. coli* [140]. Supplementation of the *in vitro* and model organism infant gut microbiota with prebiotic carbohydrates led to a larger microbiota, primarily due to larger *Bifidobacterium* populations[62, 134]. This indicates that carbon metabolism is a limiting factor, especially in the absence of prebiotics, as in our model. However, as the increased population in these *in vitro* and *in vivo* studies consisted largely of *Bifidobacterium*, other species were not necessarily also carbon limited. In addition, ATP production may not be a good proxy for biomass production in some species. Species-specific information on the relation between ATP and biomass production may aid future modelling. Finally, gaps in the GEMs we use may cause ATP production itself to be underestimated compared to what occurs *in vivo*.

Besides details on the modelling of metabolism, the way that diffusion and advection of bacteria and metabolites is handled in the simulations may impact the predictions of our model. The sensitivity analysis demonstrated an unrealistic metabolic output when the metabolic diffusion coefficient was raised to $2.2 \cdot 10^{-4}$ cm^2/s (S5 Figure J). Although this value is much lower than diffusion coefficients measured in the adult intestinal lumen ($1 \cdot 10^{-2}$ cm^2/s) [159], in the infant colon mixing may be reduced compared to the adult colon: motor activity that mixes the colon contents [160] is rarer in infants than in older children or adults [161, 162]. Exact measurements of mixing in the infant colon will be hard to obtain. Related to this, while we did not examine the effect of bacterial advection in our current work, our previous work [122] suggested that increased bacterial advection greatly reduces diversity and spatial patterning in the model. In fact, compared to the adult gut, in the infant gut there is increased motor activity driving advection [162] and infants display faster colonic transit than adults [141, 142]. In light of these data, which seem to indicate much faster bacterial advection and stronger luminal mixing than what we have assumed in our model, a potential interpretation of the present set-up of our model is that it represents the dynamics of bacterial populations adhering to the intestinal wall in interaction with metabolites advecting through the lumen. Our future work will explore in more detail how a balance between adhesion of microbiota to the mucus [163], versus advection of bacteria and metabolites in the lumen affects the colonization of the infant colon.

Our modeling approach relates to alternative simulation frameworks, such as Steadycom, BacArena, and COMETS [131, 124, 123, 121]. These frameworks, in particular the new COMETS 2 framework [132] would certainly be suitable for an-

swering questions similar to those asked in the present work. In absence of suitable frameworks at the initial phases of this project, and given the flexibility that comes with using an in-house code base, we preferred to continue developing our own line of gut models [122]. A future implementation of our model in one of the available simulation frameworks would be useful exercise, facilitating future development and comparison of models.

We have used our modeling approach to generate testable hypotheses on the causes and mechanics of succession in the infant gut microbiota, potentially laying the foundation for nutritional interventions that could improve the health of infants. It should be emphasized that the current and future work on this model represent a tool for generating hypotheses, and for testing potential mechanical explanations. We cannot fully represent the complexities of the infant gut microbiota, and any generated hypotheses must necessarily be validated *in vitro* and *in vivo*. There is work to be done to further study the relevant factors, and to bring these results into a more realistic model context. We aim to do so by integrating the prebiotic oligosaccharide and protein content of nutrition into the model. This may lead to the creation of more niches in the model, and so a diversity closer to that of the *in vivo* system. We will also continue to improve the selection and curation of metabolic models, such as by disabling lactate uptake in certain species or including a wider diversity of GEMs. This may lead to novel future insights on the interactions between differences in infant nutrition and succession in the infant microbiota.

2.4 Methods

We use a spatially explicit model to represent the newborn infant microbiota (Fig. 2.1A). Our model is based on an earlier model of the adult gut microbiota [122]. The model consists of a regular square lattice of 225×8 lattice sites, where each lattice site represents 2 $mm \times 2$ mm of space, resulting in an infant colon of 450mm by 16mm. Each lattice site can contain any number of metabolites of the 723 types represented in the model in any concentration and a single simulated bacterial population. The metabolism of these bacterial populations is based on genome-scale metabolic models (GEMs) from the AGORA collection of species-specific GEMs [10]. From this set, we have chosen 15 GEMs based on a consortium of known infant microbiota species [139, 3]. Their metabolic inputs and outputs are calculated using dynamic flux balance analysis (FBA) [113] with an enzymatic constraint functioning as a limit on the

total flux through the network [100]. The effects of the FBA solution are applied to the spatial environment and then recalculated each timestep, creating a spatial dynamic FBA.

We identify the two narrow ends of the rectangle with the proximal and distal ends of the colon. In each timestep metabolites both mix and flow from the proximal to the distal end. Bacterial populations are mixed, but do not flow distally as metabolites do.

At the start of each run, we initialize the system with a large number of small populations. We let these perform their metabolism each timestep. They take metabolites from the environment, and deposit the resulting products. We let the populations grow and divide according to their energy output. Both the initial placement locations and movement of the populations is random, introducing stochasticity in the model. The system develops differently depending on initial conditions, into a diverse and complex ecosystem.

2.4.1 Species composition

Each population is represented by a GEM of a species. 15 different metabolic models are used in our spatially explicit model (Table 2.1), selected based on previous research [139]. From this list of genera, the most prevalent species within that genus in the vaginally born newborn data set from Bäckhed *et al.* was used [3]. One group from the list could only be determined at the family level - Lachnospiraceae. *Ruminococcus gnavus* was chosen to represent this group, based on its high prevalence among this group and the prior inclusion of species from the *Blautia* and *Dorea* genera [43]. Because the genus *Bifidobacterium* is known to be particularly diverse [164], we represented it with models of three different strains. All models are based on the GEMs created by Magnúsdóttir et al. in the AGORA collection [10].

2.4.2 Changes from AGORA

We use updated versions of the AGORA GEMs [165], to which we have applied checks and modifications. Firstly, the objective function was changed from the biomass reaction included in the models to a reaction only requiring ATP production. As this reaction yields only ADP and P_i, it is mass-neutral. This allows us to focus on carbon dissimilation within the GEMs, and the differences in underlying metabolism. ATP yield has been a good proxy for biomass production in previous studies [166].

Focusing on carbon dissimilation means we can leave all unknown uptake bounds at 0, instead of using an arbitrary or randomized level, as in some other studies [123, 131].

We also checked the metabolic networks for reactions that allow for the occurrence of unrealistic FBA solutions, and we have added additional reactions. In the *Bifidobacterium* models the ATP synthase reaction was made reversible. This allows it to function as a proton pump, which *Bifidobacterium* species use to maintain their internal pH [167]. Lactose permease reactions were added to all *Bifidobacterium* species in the model [168] based on those available in other models in the AGORA collection [10]. Combined with the existing reactions in the model and the metabolic constraints this leads to a set of reactions simulating a realistic bifid shunt yield, when substrate is abundant, of 5 mol ATP, 2 mol lactate and 3 mol acetate per mol of lactose[90, 169]. Lactose permease was also added to *Streptococcus salivarius* and *Ruminococcus gnavus* to bring them in line with existing literature on their *in vitro* behavior [170, 171]. *Veillonella dispar* is the only species in the model that does not have any lactose uptake [95]. A complete list of changes is presented in S1 Table.

2.4.3 Checking the validity of the GEMs

After the changes in S1 Table were applied all GEMs used in the model were tested individually to ensure that they could grow on a substrate of lactose. Only *Veillonella* did not pass this test, which is consistent with *in vitro* observations [95]. *Veillonella* did pass when lactate, a common infant gut metabolite, was used instead. We also tested all GEMs individually for spurious growth in absence of substrates. To this end, all uptake bounds except water were set to 0. None of the GEMs grew under these conditions. All models were tested for having a net neutral exchange of hydrogen, carbon, oxygen and nitrogen at each timestep of the model during the simulations. There should be no net change in the number of atoms in the medium due to the calculated fluxes, because the reaction we set as an objective function does not remove any atoms, and there are no other sinks in the simulations. In the present simulation, only water, oxygen, and lactose are introduced into the simulation, so no other atoms than these three are considered. The tests revealed two errors in glycogen metabolism in several GEMs in AGORA that resulted in a energetically-favorable removal of intracellular protons from the system. We corrected these GEMs by replacing the reactions responsible for this erroneous energy source (S1 Table).

With the corrections we applied to the model, these tests always passed - allowing for rounding errors of less than $1 \cdot 10^{-8}$ µmol per FBA solution.

The thermodynamic correctness of all reactions was checked by calculating the net difference in Gibbs free energy between input and output metabolites, using a pre-existing dataset of thermodynamic source data [172, 173]. The conditions assumed for the calculation of the Gibbs free energy were a pH of 7 and an ionic strength of 0.1M [172]. This was recorded for each population at each timestep over the course of a full run of 10080 timesteps. 99.999% of all FBA solutions in the simulations of Fig. 2.3 passed the test by containing less free energy than the associated inputs. The remaining 0.0001% all had an amount of energy in the outputs equal to that of the inputs. All of these solutions had very low growth rates, less than 0.001% of the average growth rate. The sum of these growth rates over 30 simulations was less than 0.05 of an initial population, totalling $2 \cdot 10^6$ bacteria.

2.4.4 FBA with enzymatic constraint

Each timestep of the model, a modified version of flux balance analysis with an enzymatic constraint is used by each population to determine what reactions it should use to achieve the most biomass production from the metabolites available to it [13, 100]. First, each GEM is converted to a stoichiometric matrix S. All reversible reactions are converted to two irreversible reactions, so that flux is always greater than or equal to 0. All reactions identified as exchange, sink, or demand in the metabolic reconstruction are marked as exchange in the matrix. These reactions exchange nutrients or metabolites with the environment. Each timestep, all reactions are assumed to be in internal steady state [13]:

$$S \cdot \vec{f} = 0 \tag{2.1}$$

Where $\vec{f}$ is a vector of the metabolic fluxes through each reaction in the network (in mol per time unit per population unit).

Each exchange reaction that takes up metabolites from the environment ($F_{in}(m)$) is constrained by an upper bound $F_{ub}(m)$ which represents the limited availability of most nutrients from the environment:

$$\vec{F_{in}} \leq \vec{F_{ub}} \tag{2.2}$$

Where $\vec{F_{in}}$ is a vector of fluxes between the environment and the bacterial population (in mol per time unit per population unit). $\vec{F_{ub}}$ is set dynamically at each timestep t by the spatial environment (discussed in detail below) at each lattice site x:

$$\vec{F_{ub}}(\vec{x},t) = \frac{\vec{c}(\vec{x},t)}{B(\vec{x},t)} \tag{2.3}$$

Where $\vec{c}$ is a vector of all metabolite concentrations in mol per lattice site, $\vec{x}$ is a vector of all lattice sites and $B(\vec{x})$ is the size of the local bacterial population in population units. Population units are continuous, the size of B can range from $5 \cdot 10^7$ to $2 \cdot 10^{10}$, or $4 \cdot 10^{10}$ when division is not possible.

There is an additional constraint on the total flux. This constraint represents the limited amount of metabolism that can be performed per cell in each timestep. Each cell can only contain a limited number of enzymes, and each enzyme can only perform a limited number of reactions in limited time. Not limiting flux would lead to an unrealistic situation. The enzymatic constraint uses the sum of all fluxes $\vec{f}$, not just those that exchange with the environment:

$$\sum \vec{f} \leq a \tag{2.4}$$

The enzymatic constraint a is in mol per time unit per population unit (Table 2.3). As both $\vec{f}$ and a are per population unit, this limit scales with population size, allowing each bacterium to contribute equally and independently to the metabolic flux attained in a lattice site. The enzymatic constraint is included as a constraint on each local FBA solution separately at each timestep. The enzymatic constraint was first proposed in a metabolic modelling context in 1990 [100], adapted from the study of other capacitated flow networks. In our context the enzymatic constraint represents the limited availability of physical space for enzymes. Our method is similar to FBA with molecular crowding in that we use an additional constraint to model metabolic capacity [133]. However, we do not utilize the necessarily reaction-specific crowding coefficients.

FBA then identifies the solution space that adheres to these constraints and from this spaces identifies the solution that optimizes the objective function. The set of solutions consists of a set of exchange fluxes $\vec{F}(\vec{x},t)$, and a growth rate $\vec{g}(\vec{x},t)$. These exchange fluxes are taken as the derivatives of a set of partial-differential equations to model the transport of intermediary metabolites (see below). The size of the population increases in proportion to the growth rate produced by the solution.

2.4.5 Environmental metabolites

We modeled a set of 723 extracellular metabolites - the union of all metabolites that can be exchanged with the environment by at least one GEM in the model. In practice, only a fraction of the metabolites occur in any exchange reaction that has flux over it, and only 17 metabolites are ever present in the medium in more than micromolar amounts in our simulations outside of the sensitivity analysis (Table 2.2). Though the model distinguishes between L-lactate and D-lactate, we display them together in our figures. Nearly all lactate produced and consumed in our model is L-lactate.

Table 2.2: Extracellular metabolites present in the model in more than micromolar amounts. BiGG IDs from [10]

Common name	BiGG ID
Acetate	ac
Alpha-Ketoglutaric acid	akg
Carbon dioxide	co2
Ethanol	etoh
Formate	for
Hexadecanoate	hdca
Hydrogen	h2
L-Lactate	lac_L
D-Lactate	lac_D
Lactose	lcts
Octadecanoate	ocdca
Oxygen	o2
Propionate	ppa
Proton	h
Succinate	succ
Tetradecanoate	ttdca
Water	h2o

To mimic advection, the complete metabolic contents of the lattice except oxygen are moved towards the distal end by 2mm (one lattice site with the default parameters) per timestep, i.e. once every 3 simulated minutes. This leads to an average transit time of approximately 11 hours, in agreement with the observed cecum to rectum transit time in newborns [141, 142]. Metabolites moving out of the distal end are removed entirely and analysed separately (see Section ‘ Analysis’ for detail).

Every timestep all metabolites diffuse to the four nearest neighbours $NB(\vec{x})$ at an equal rate for all metabolites (Table 2.3). We used a baseline of $4.4 \cdot 10^{-5}$ cm^2/s to represent mixing of metabolites, which is an order of magnitude higher than normal

diffusion for common metabolites [174] to represent active mixing due to colonic contractions. Metabolites are also added and removed by bacterial populations as a result of the FBA solutions, yielding

$$\frac{d\vec{c}(\vec{x},t)}{dt} = \vec{F_{out}}(\vec{x},t)B(\vec{x},t) - \vec{F_{in}}(\vec{x},t)B(\vec{x},t) + \frac{D}{L^2}\sum_{\vec{i}\in NB(\vec{x})}\left(\vec{c}(\vec{i},t) - \vec{c}(\vec{x},t)\right). \quad (2.5)$$

Where F_{out} is a vector of fluxes from the bacterial populations to the environment, in mol per time unit per population unit.

All lattice sites initially contain water as their only metabolite, except for the conditions where oxygen is added. Metabolites representing the food intake are inserted into the first six columns of lattice sites every three hours (60 timesteps) to approximate a realistic interval for neonates [175]. In the current model this food intake consists solely of lactose, in a concentration in line with human milk [66], assuming 98% host uptake of carbohydrates before reaching the colon, a commonly used assumption [131]. Water is provided as a metabolite in unlimited quantities. Oxygen is placed evenly distributed or at the upper and lower boundaries in some simulations. No other metabolites are available, other than those produced as a result of bacterial metabolism within the model.

2.4.6 Population dynamics

Each local population solves the FBA problem based on its own GEM, an enzymatic constraint a, its current population size $B(\vec{x},t)$ and the local concentrations of metabolites $\vec{c}(t)$ at each timestep, and applies the outcome to the environment (see above) and its own population size, as follows:

$$\frac{d\vec{B}(\vec{x},t)}{dt} = \vec{B}(\vec{x},t)\cdot\vec{g}(\vec{x},t) \quad (2.6)$$

Populations at least 200 times the initial size (Table 2.3) will create a new population in one empty adjacent lattice site, if possible. Half of the old population size is transferred to the new population, in such a way that the total size is preserved. To mimic colonization events new populations are introduced at random into empty lattice sites during the simulation, representing both dormant bacteria from intestinal crypts [143] and small bacterial populations that are formed from ingested bacteria,

which may only become active after having diffused far into the gut. Each empty lattice has a probability of 0.00005 (Table 2.3) each step to acquire a new population of a randomly selected species, as follows:

for *site* ∈ *latticesites* **do**
  if *site* == *empty* **then**
    if *randomnumber* ≤ *placement probability* **then**
      Place metabolic network of random species at *site*

The probability is scaled by $\frac{1}{n^2}$, with n the scaling factor of Δx, the side length of a lattice site. There is an equal probability for any species in the model to be selected. As we consider these new populations to be new colonizers we initialize them at the same population size B as the initial populations in the model (Table 2.3). Each population dies out at a probability of 0.0075 per timestep, creating a turnover within the range of estimated microbial turnover rates in the mouse microbiota [72].

To mimic the mixing of bacterial populations, the lattice sites swap population contents each timestep. We use the following algorithm, inspired by Kawasaki dynamics [176], as also used previously for bacterial mixing [122, 177]: In random order, the content of each site, i.e., the bacterial population represented by its size and the GEM but not the metabolites, are swapped with a site randomly selected from the set consisting of the site itself and the first and second order neighbours. This swap only occurs if both the origin and destination site have not already swapped. Bacterial populations at the most distal column, i.e. at the exit of the colon, are deleted from the system. With this mixing method the diffusion constant of the bacterial populations becomes $5.7 \cdot 10^{-5}$ cm^2/s (Table 2.3). In the simulations that use a finer or a coarser grid (Fig. 2.6D) the number of swaps is scaled as $\frac{1}{n^2}$, with n the scaling factor of Δx, the side length of a lattice site, thus maintaining the same diffusion rate. For $n < 1$ all sites are marked as unswapped once all sites have attempted a swap. This allows for sites to swap multiple times. The same approach was used to change the bacterial diffusion rate in the sensitivity analysis (Fig. 2.6F). To achieve full mixing all bacteria were assigned random non-overlapping locations at every timestep. Note that the diffusion rate of bacterial populations with the default parameters (Table 2.3) is 1.3 times higher than that of the metabolites. As shown by the sensitivity analysis (Fig. 2.6F) this has little effect on the results.

2.4.7 Initial conditions

The simulation is initialized by placing a number of very small populations (B) of the various species randomly across different lattice sites of the environment (Table 2.3). There is a probability of 0.3 for each lattice site to acquire a population, an average of 540 for our lattice. The size of initial populations is scaled to be roughly equivalent to a plausible initial total load of approximately $3 \cdot 10^{10}$[5], assuming a total colon volume of approximately 90 ml. As there is little information on the relative abundance of species in the very early infant gut, we place all species with equal probability. In initially oxygenated conditions, oxygen is also placed as a metabolite now. Water is always considered to be present everywhere. No other metabolites are initially present except for the first feeding.

The ratio of population to bacterial abundance was chosen as a balance between a high bacterial resolution and computational complexity, as the FBA solution of each population is calculated independently per timestep. This resolution splits the infant microbiota into approximately 1-600 units (depending on conditions), allowing for the presence of species at relatively low abundances, and spatial variation. There is no hard limit on the number of populations in the model - it is limited only by the balance between growth (through consumption) and death.

2.4.8 Analysis

At each timestep we record the location and exchange fluxes $\vec{F}(\vec{x}, t)$ of major metabolites as well as the size $B(\vec{x}, t)$ and species of all populations. This is used to analyze both population composition and metabolic fluxes over time and space. In addition, each timestep we record the location and quantity $\vec{c}(\vec{x}, t)$ of all metabolites present in more than micromolar quantities. We also record this metabolite data for all metabolites exiting the system at the distal side. This is considered the closest equivalent to a fecal composition in the model, and these results are compared to data from *in vivo* fecal samples.

To detect any irregularities, we also record the net carbon, hydrogen, and oxygen flux of every population and the system as a whole. The difference in Gibbs free energy per timestep is also recorded per for each population per FBA solution, and seperately over the whole system. Estimated Gibbs free energy is derived from the Equillibrator database [173]. Energy loss l in joules per timestep per population unit is recorded as follows, where m are metabolites, F is the exchange fluxes in mol per

population unit, B is the population size and E contains the Gibbs free energy in joules per mol for each metabolite,

$$l = \sum_{m} F(m) \cdot E(m) \tag{2.7}$$

For specific simulations, reactions are removed from the models. This is performed by deleting some reactions from the GEM before the conversion to the stoichiometric matrix. To remove Fructose-6-phosphate phosphoketolase, we removed the reactions R_PKL and R_F6PE4PL. For other simulations, the uptake of certain metabolites is disabled. This is done by placing the upper flux bound $\vec{F}_{ub}$ of the relevant exchange reaction at 0 for the relevant populations.

Table 2.3: Default parameters of the model

Parameter	Value	Unit
Timestep	180	seconds
Lattice site side length (Δx)	2	millimetre
Width of the lattice	225	lattice sites
Height of the lattice	8	lattice sites
Colon transit time	11	hours
Average number of initial populations	540	-
Initial size per population	$5 \cdot 10^7$	number of bacteria
Population size required for division	$1 \cdot 10^{10}$	number of bacteria
Death probability	0.0075	per timestep per population
Enzymatic constraint	2	µmol flux per timestep per $1 \cdot 10^{10}$ bacteria
New species placement probability	0.00005	per timestep per empty lattice site
Diffusion of metabolites	$4.4 \cdot 10^{-5}$	cm^2/s
Diffusion of populations	$5.7 \cdot 10^{-5}$	cm^2/s
Growth per µmol ATP	$1 \cdot 10^9$	number of bacteria
Lactose input	211	µmol per three hours

2.4.9 Parameters

Relevant parameters are listed in Table 2.3. Based on measurements of the typical length and diameter of the infant colon [135, 136] we estimated a volume of 90 ml. Combined with the average abundance per ml of around 10^{10} after the first days [5], this leads to a very rough estimate of 10^{12} bacteria in the young infant colon. To remain computationally feasible, while still modelling at a high resolution, we divide this population into units of at most $1 \cdot 10^{10}$. Local populations at or above this limit will divide into two equally sized populations when space is available. This prevents

the local population from becoming unrealistically high. Local populations of more than $2 \cdot 10^{10}$, which can only form if no space is available to divide for a longer time period, cease metabolism.

The lactose input is estimated from the known intake of milk, its lactose concentration, and an estimate of pre-colonic lactose absorption of 98% [178, 131]. Little data is available on the growth rate of bacteria within the human colon. Growth rates are expected to be much lower than those found in *in vitro* cultures of individual species [179]. In absence of precise data for infants, here we use a death probability that places the replacement rate within the range of estimated doubling times of the whole gut microbiota in mice [72]. The colonic transit time is based on data for total transit time gathered with carmine red dye [141], adjusted for the mouth to cecum transit time [142]. The timestep interval was set at three minutes, to be able to capture individual feedings at a high resolution. Other parameters were selected to reach plausible outcomes on a metabolic and species level while maintaining computational feasibility. We do not expect the parameters to match precisely with *in vivo* values, if these were measured.

2.4.10 Implementation

The model was implemented in C++11. Our code was based on code used earlier to model the gut microbiota [122]. Random numbers are generated with Knuth's subtractive random number generator algorithm. Diffusion of metabolites was implemented using the Forward Euler method. The GEMs are loaded using libSBML 5.18.0 for C++. We used the GNU Linear Programming Kit 4.65 (GLPK) as a linear programming tool to solve each FBA with enzymatic constraint problem. We used the May 2019 update of AGORA, the latest at time of writing, from the Virtual Metabolic Human Project website (vmh.life). Python 3.6 was used to extract thermodynamic data from the eQuilibrator API (December 2018 update) [173] and determine mean square displacement of our bacterial diffusion. Model screenshots are made using the libpng16 and pngwriter libraries. Other visualisations and statistical analysis were performed with R 4.1.2 and Google Sheets.

2.4.11 Data availability

The code used for the model is available from GitHub at `https://github.com/DMvers/IGMOST`. The data files used for the model are available

from GitHub at `https://github.com/DMvers/IGMOSTdatafiles`.

2.5 Supplemental material

Available from `https://journals.asm.org/doi/full/10.1128/msystems.00446-22`

S1 Table

Table of changed or deleted reactions and annotations.csv
A table of changes made to the AGORA models as a .csv file.

S2 Table

Table of reactions used by *Bifidobacterium longum infantis* in model depending on lactose concentration
A table containing the flux through each reaction with more than nanomolar flux for *Bifidobacterium longum infantis* populations of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml). Water was unlimited, no other metabolites were present. Lactose was present at 0.05, 0.35, or 0.5 µmol per ml.

S1 Video

Video of a run with no initial oxygen, consisting of a visualisation of the distribution of bacterial species and major metabolites, and a visualisation displaying fluxes between population and metabolite pools. Line width and intensity are proportional to the amount exchanged with the environment over the last 60 timesteps, with a threshold of 0.5 µmol, with no normalization. Metabolite circle size is relative to the most abundant metabolite, with a minimum displayed size of 26 pixels.

S2 Video

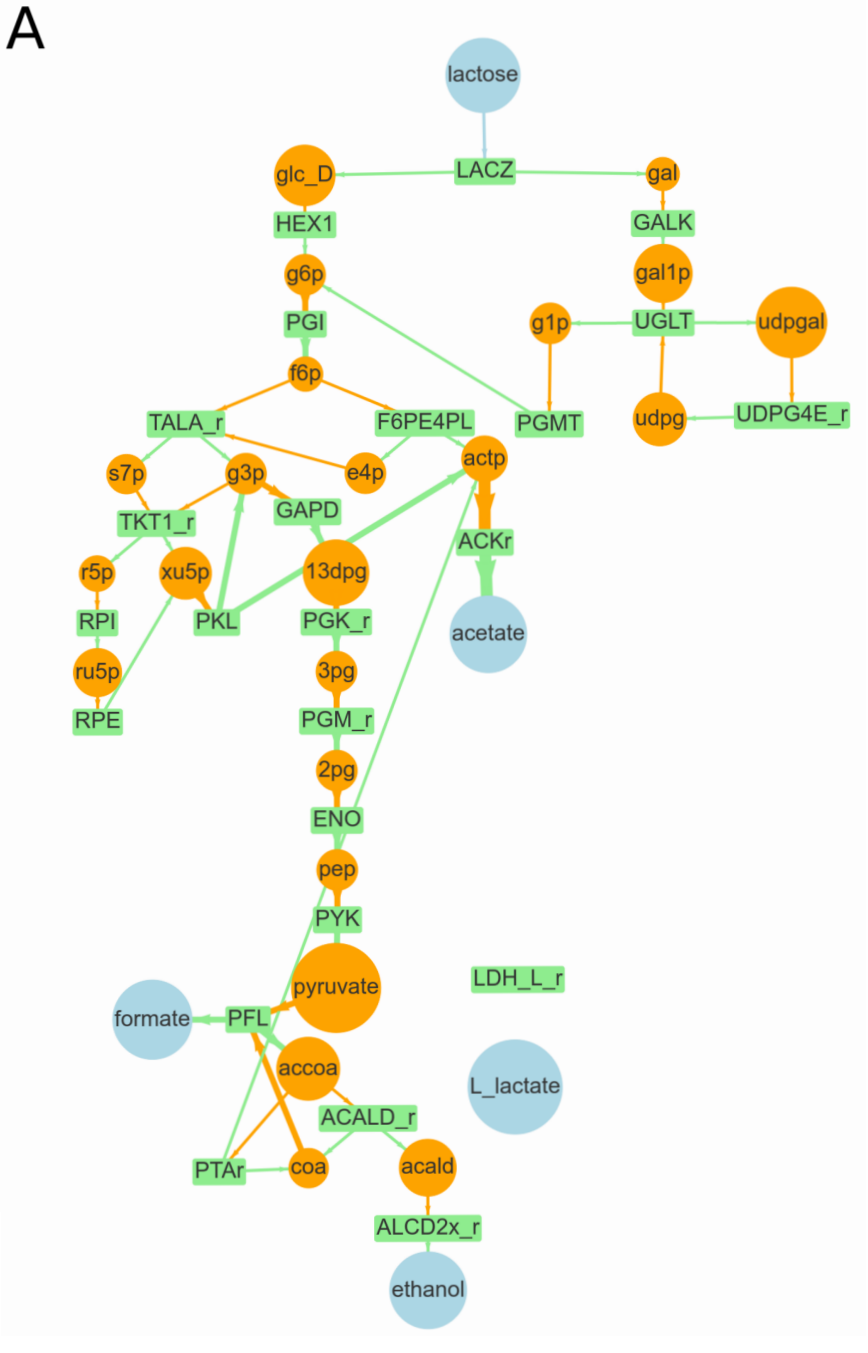
Visualisations, from left to right, of a run from Fig. 2.4A, a run from Fig. 2.4B, and a run from Fig. 2.4F, displaying population and metabolite pool sizes, and fluxes between populations and metabolites. Line width and intensity are proportional to the amount exchanged with the environment over the last 60 timesteps, with a threshold of 0.5 µmol, with no normalization. Metabolite circle size is relative to the most abundant metabolite, with a minimum displayed size of 26 pixels.

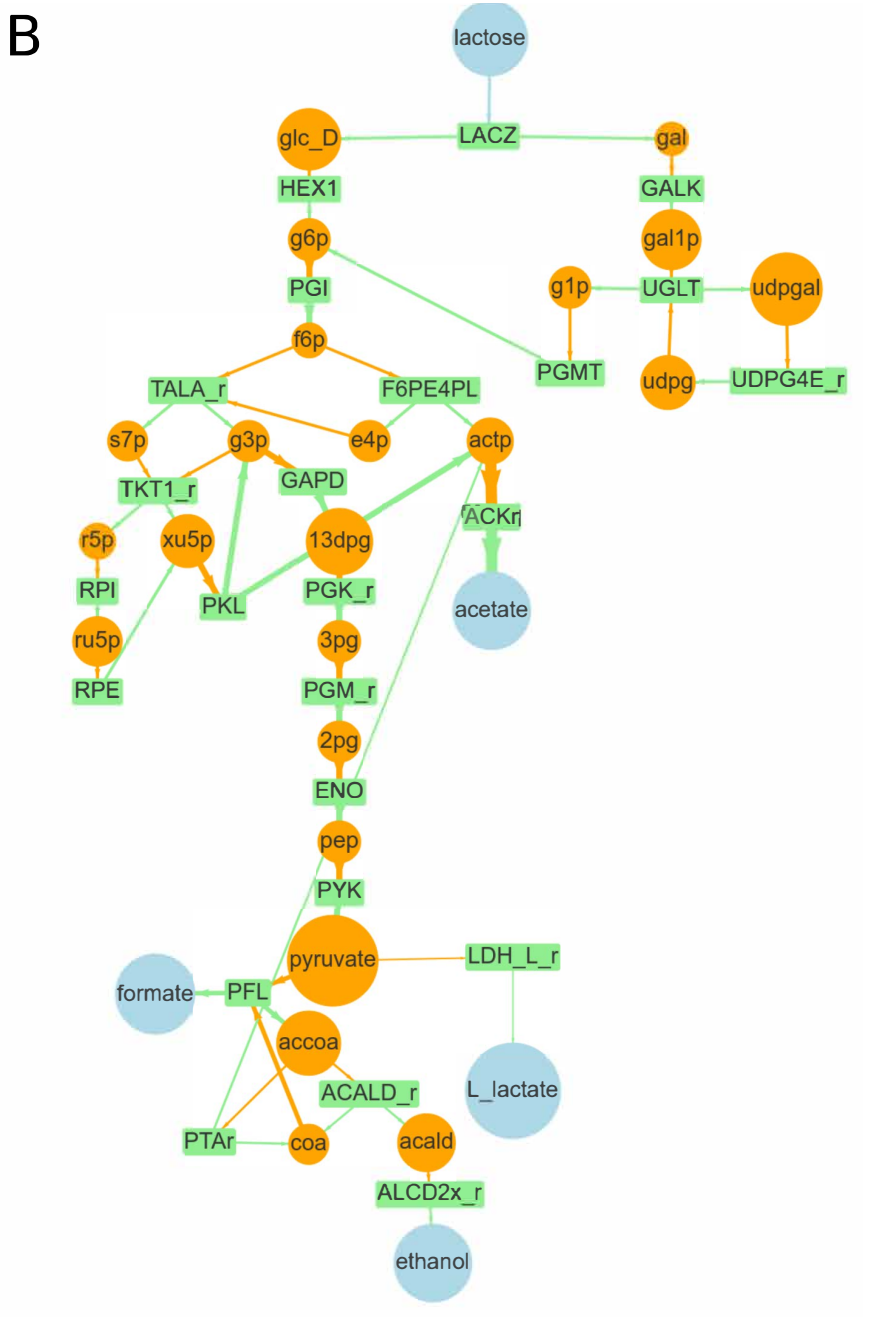
S3 Video

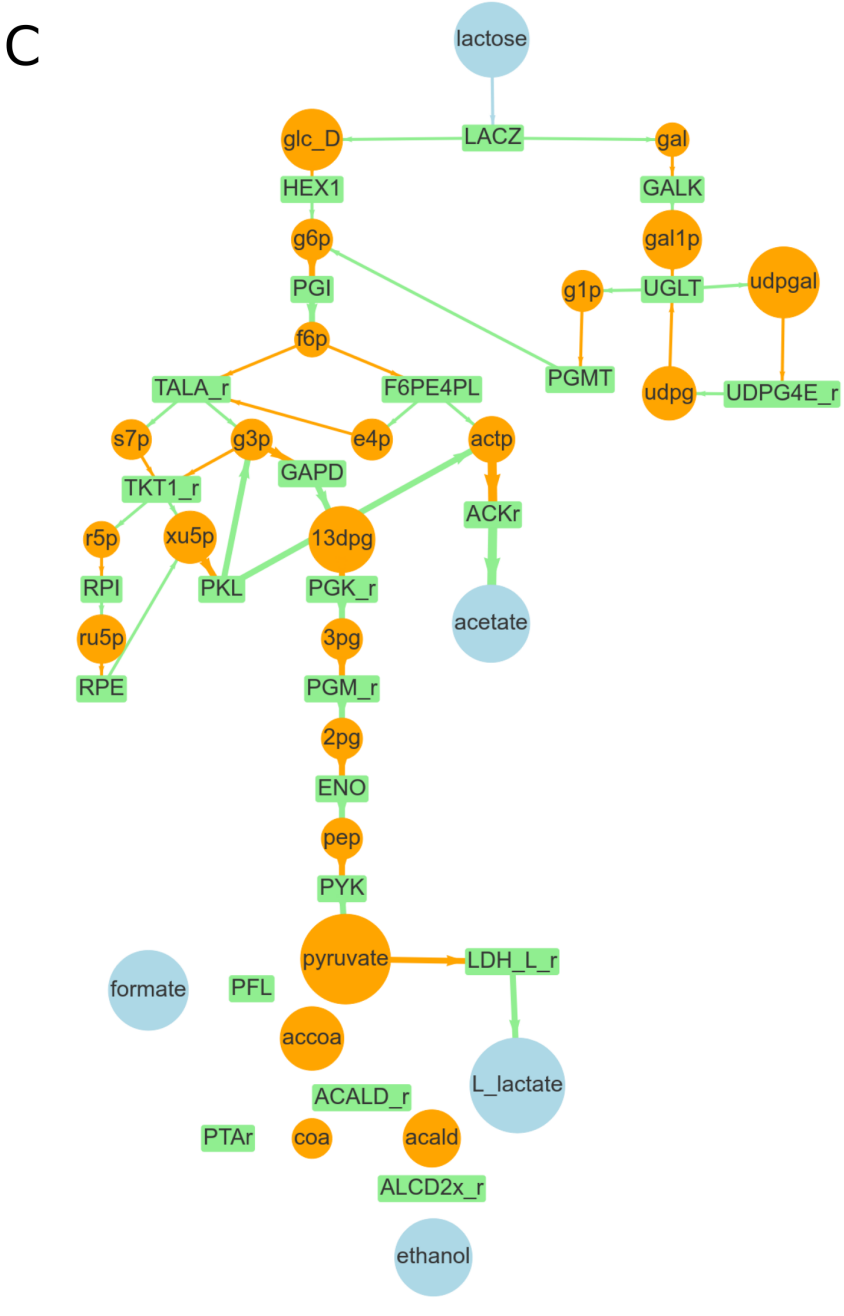
Video of a run with an initial 0.1 µmol of oxygen per lattice site, consisting of a visualisation of the distribution of bacterial species and major metabolites, and a visualisation displaying fluxes between population and metabolite pools. Line width and intensity are proportional to the amount exchanged with the environment over the last 60 timesteps, with a threshold of 0.5 µmol, with no normalization. Metabolite circle size is relative to the most abundant metabolite, with a minimum displayed size of 26 pixels.

S4 Video

Visualisations, from left to right, of a run from Fig. 2.5 D, a run from Fig. 2.5 H, and a run from Fig. 2.6 A without the enzymatic constraint, displaying population and metabolite pool sizes, and fluxes between populations and metabolites. Line width and intensity are proportional to the amount exchanged with the environment over the last 60 timesteps, with a threshold of 0.5 µmol, with no normalization. Metabolite circle size is relative to the most abundant metabolite, with a minimum displayed size of 26 pixels.

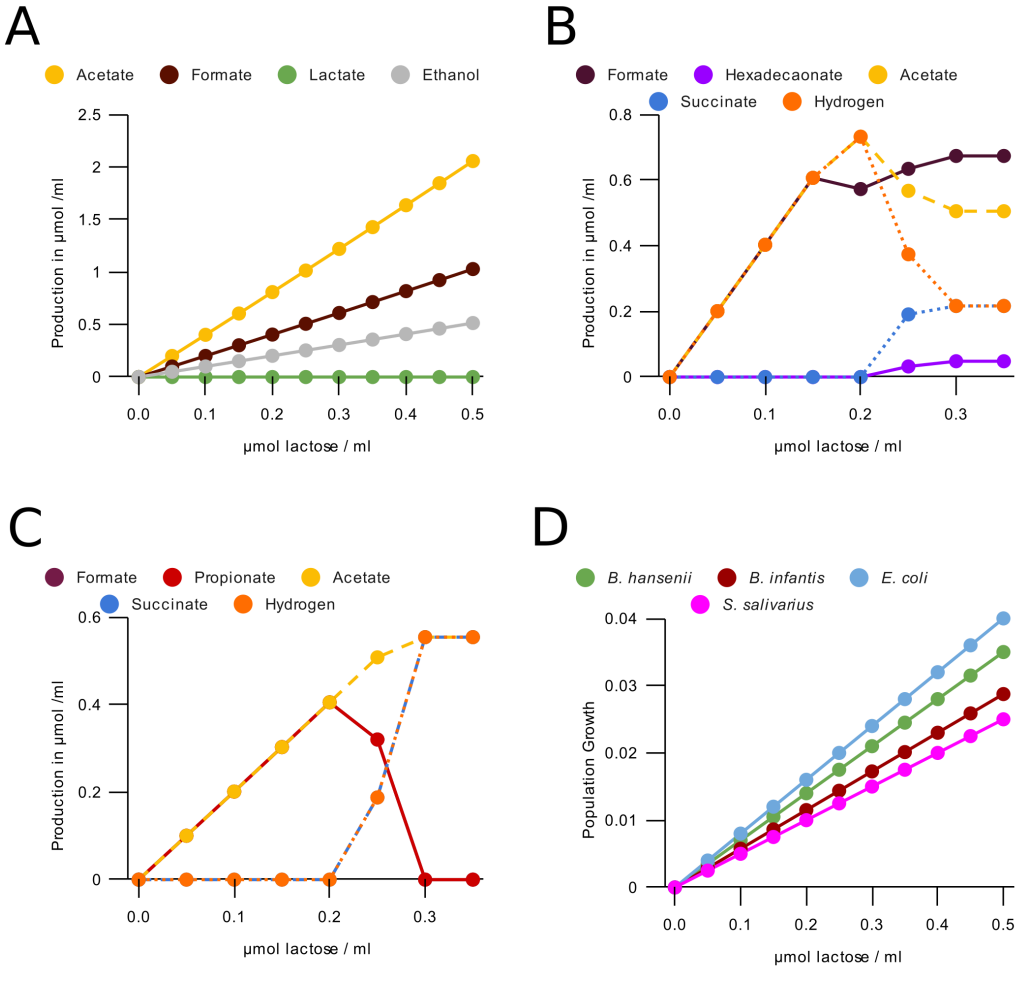

A
lactose
LACZ
glc_D
gal
HEX1
GALK
g6p
gal1p
PGI
g1p
UGLT
udpgal
f6p
udpg
UDPG4E_r
TALA_r
F6PE4PL
PGMT
actp
s7p
g3p
e4p
GAPD
TKT1_r
ACKr
r5p
xu5p
13dpg
RPI
PKL
PGK_r
acetate
ru5p
3pg
RPE
PGM_r
2pg
ENO
pep
PYK
pyruvate
LDH_L_r
formate
PFL
accoa
L_lactate
ACALD_r
PTAr
coa
acald
ALCD2x_r
ethanol


B
lactose
LACZ
glc_D
gal
HEX1
GALK
g6p
gal1p
PGI
g1p
UGLT
udpgal
f6p
PGMT
udpg
UDPG4E_r
TALA_r
F6PE4PL
s7p
g3p
e4p
actp
GAPD
TKT1_r
ACKr
r5p
xu5p
13dpg
RPI
PKL
PGK_r
acetate
ru5p
3pg
RPE
PGM_r
2pg
ENO
pep
PYK
pyruvate
LDH_L_r
formate
PFL
accoa
L_lactate
ACALD_r
PTAr
coa
acald
ALCD2x_r
ethanol


C
lactose
LACZ
glc_D
gal
HEX1
GALK
g6p
gal1p
PGI
g1p
UGLT
udpgal
f6p
PGMT
udpg
UDPG4E_r
TALA_r
F6PE4PL
s7p
g3p
e4p
actp
GAPD
TKT1_r
ACKr
r5p
xu5p
13dpg
RPI
PKL
PGK_r
acetate
ru5p
3pg
RPE
PGM_r
2pg
ENO
pep
PYK
pyruvate
LDH_L_r
formate
PFL
accoa
L_lactate
ACALD_r
PTAr
coa
acald
ALCD2x_r
ethanol

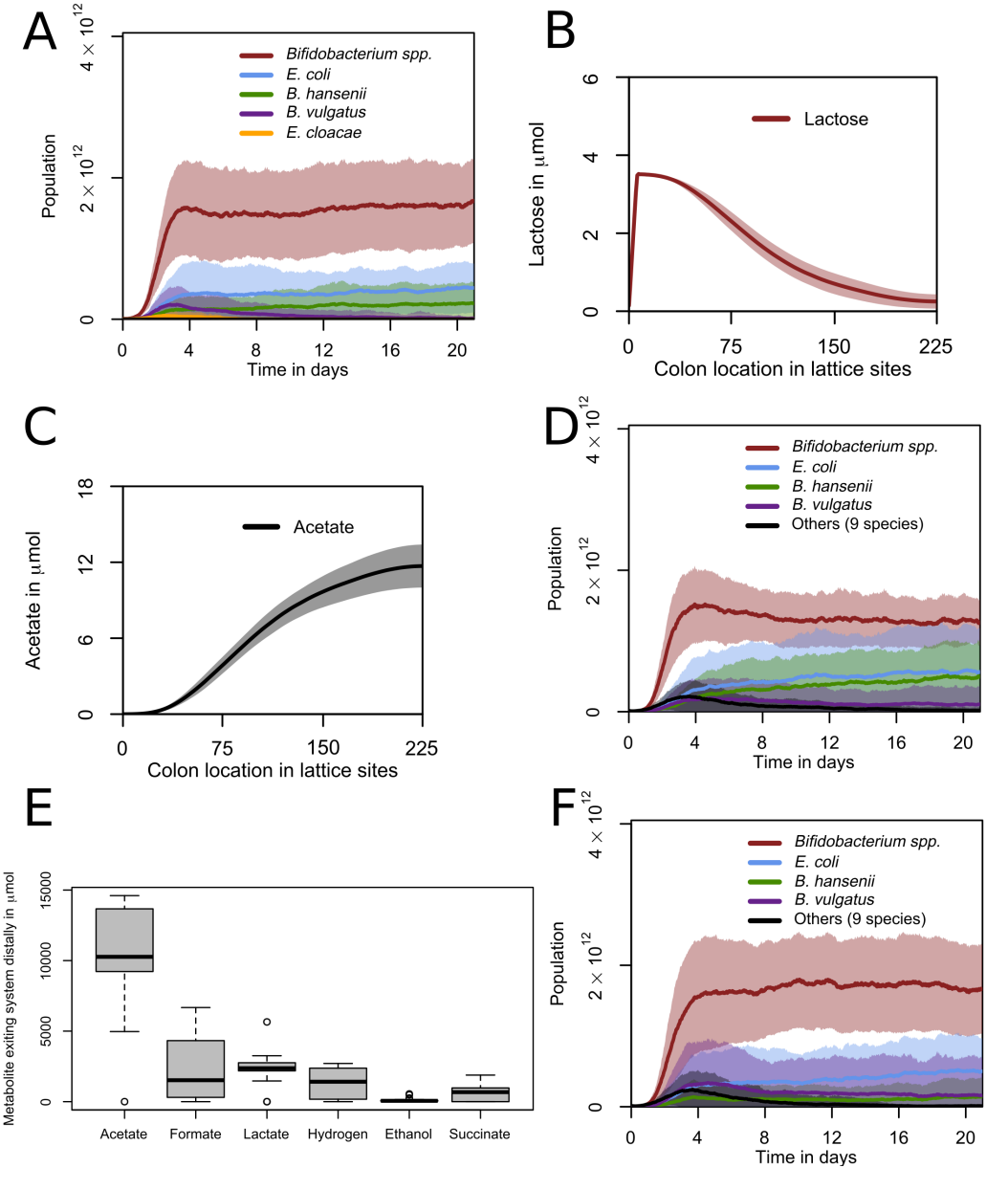
S1 Figure

(A to C) Flux per timestep through the core carbon metabolism of a *B. longum* subsp. *infantis* population of 5×10^9 bacteria with access to one lattice site (0.05 ml) using our FBA with enzymatic constraint method. Line width indicates magnitude of flux, normalised for lactose uptake. Blue circles indicate extracellular metabolites, orange circles intracellular metabolites, and green rectangles indicate reactions. Transport reactions and cofactors are not depicted. All abbreviations are from [10]. Lactose concentration was (A) 0.05 umol per ml (B) 0.35 umol per ml (C) 0.5 umol per ml.



S2 Figure

(A) Growth per timestep by lactose concentration for populations of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) of some major bacterial species with the enzymatic constraint disabled (B) Production of metabolites per timestep by an *E. coli* population of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) (C) Production of metabolites by a *B. vulgatus* population of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml). Formate and propionate formation are always equal, and so are hydrogen and succinate. (D) Growth per timestep by lactose concentration for populations of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) of some major bacterial species with the enzymatic constraint disabled

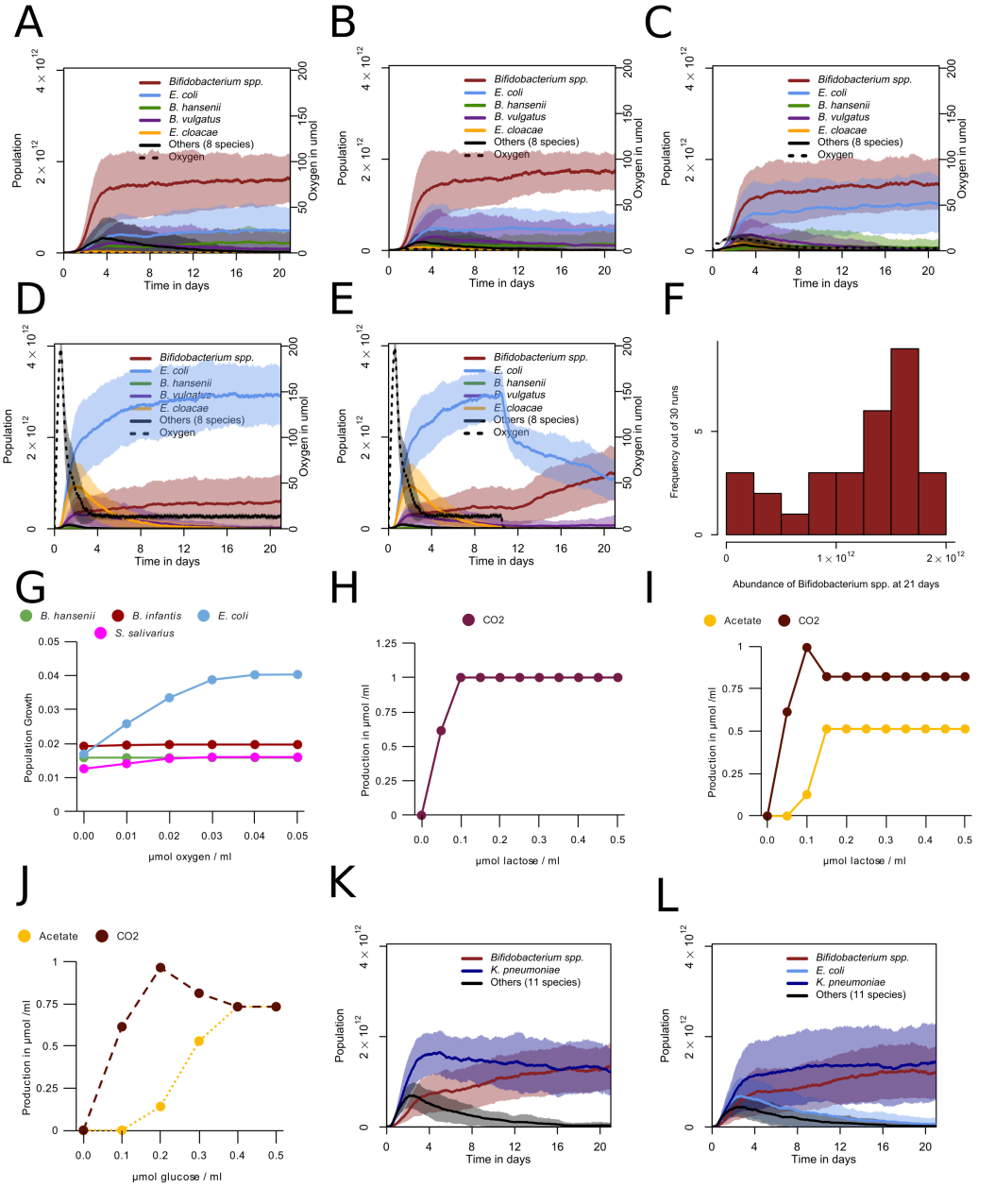


S3 Figure

(A) Populations over time with bacteria in the "others" category of Fig. 2.5 left out. Abundances for grouped species over 21 days. One day is 480 timesteps. Curve

shows mean value and shaded area one standard deviation over n=30 simulations. (B,C) Spatial distribution of (B) lactose and (C) acetate over the last two days (960 timesteps) of the simulation. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (D) Populations over time with lactate uptake disabled for *Bifidobacterium* species. Abundances for grouped species over 21 days. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (E) Distribution of metabolites exiting the system distally over the last two days (960 timesteps) of the simulation in D. n=30. (F) Populations over time with lactate production disabled for *Bifidobacterium* species. Abundances for grouped species over 21 days. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations.

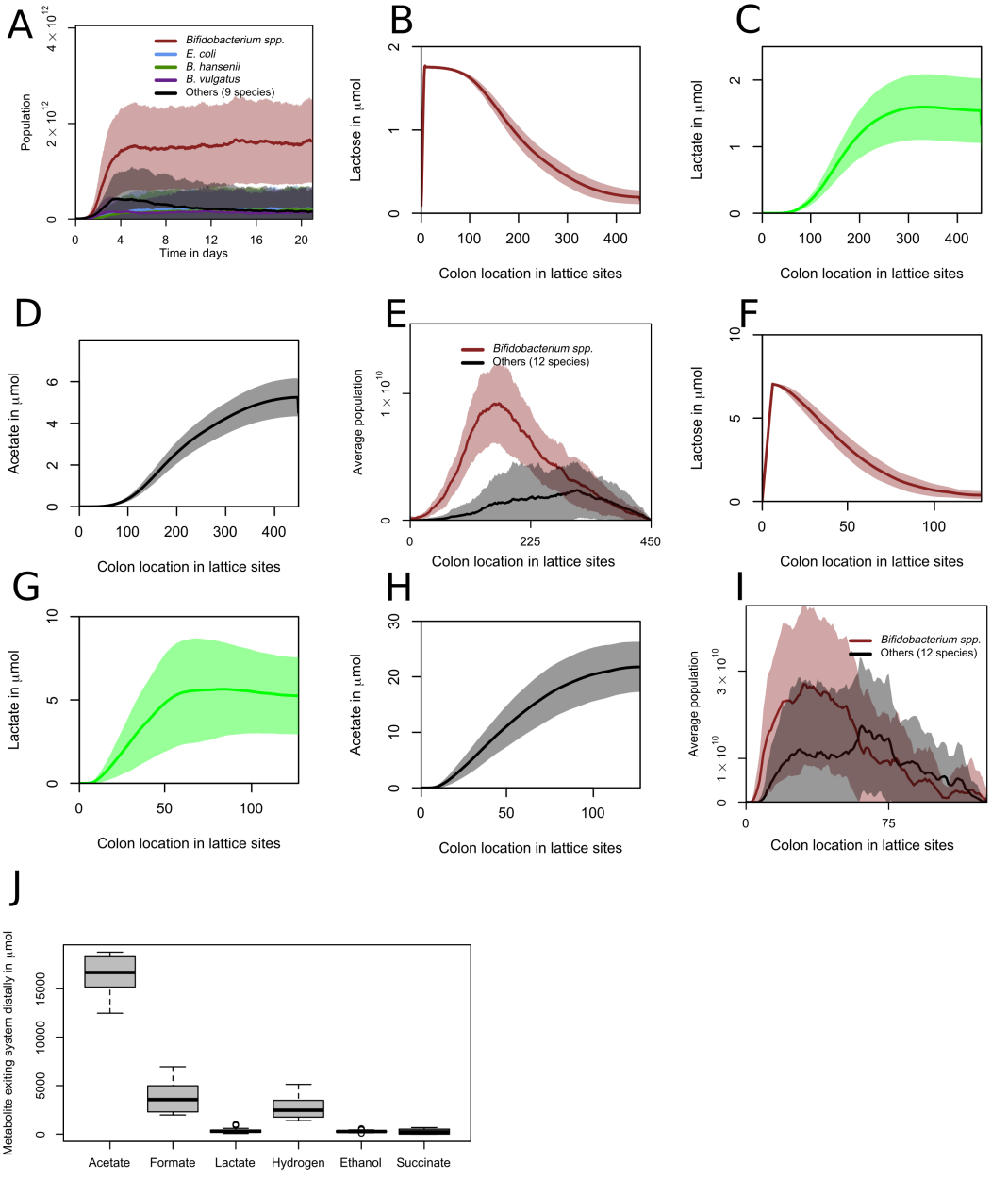
2.5. Supplemental material



S4 Figure

(A,B,C,D) Populations over time with (A) 0.001 (B) 0.01 (C) 0.1 (D) 1 µmol oxygen released in total per step divided over the top and bottom rows. Abundances for grouped species over 21 days. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations per condition. (E) Populations over time with 1 µmol oxygen released per step initially (as in D), decreasing to 0 at step 5040. Abundances for grouped species over 21 days. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations per condition. (F) Distribution of total *Bifidobacterium* abundance at 21 days (10080 timesteps) corresponding to the simulations of E (G) Growth on unlimited lactose per timestep by oxygen concentration for populations of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) of some major bacterial species (H) Production of metabolites by lactose concentration in the presence of abundant oxygen for an *E. coli* population of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) (I) Production of metabolites by lactose concentration in the presence of abundant oxygen for an *E. cloacea* population of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) (J) Production of metabolites by glucose concentration in the presence of abundant oxygen for an *E. coli* population of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml) (K) Populations over time with *K. pneumoniae* instead of *E. coli* and 0.1 µmol initial oxygen per lattice site. Abundances for grouped species. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (L) Populations over time with *K. pneumoniae* in addition to the 15 species from table 2.1, and 0.1 µmol initial oxygen per lattice site. Abundances for grouped species. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations.

2.5. Supplemental material



S5 Figure

(A) Populations over time with new populations after the initial populations only placed in the first column, with the placement probability (Table 2.3) increased to 0.01125. Abundances for grouped species over 21 days. One day is 480 timesteps. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (B,C,D) Spatial distribution of lactose, lactate, and acetate, respectively, over the last two days (960 timesteps) of the simulation with a larger lattice. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (E) Abundance per location for grouped species over the last two days (960 timesteps) of the simulations with a larger lattice. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (F,G,H) Spatial distribution of lactose, lactate, and acetate, respectively, over the last two days (960 timesteps) of the simulation with a smaller lattice. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (I) Abundance per location for grouped species over the last two days (960 timesteps) of the simulations with a smaller lattice. Curve shows mean value and shaded area one standard deviation over n=30 simulations. (J) Distribution of metabolites exiting the system distally over the last two days (960 timesteps) of the simulation with diffusion of metabolites increased by a factor of 5. n=30.

2.6 Contributions

P.V.J, M.P., J.M.W.G., and R.M.H.M acquired funding. D.M.V., P.V.J, M.P., J.M.W.G., and R.M.H.M. conceived and planned the simulations. D.M.V. and D.M. wrote software used for the simulations. D.M.V. performed the simulations and analyzed the data. R.S, E.L., J.M.W.G., and R.M.H.M contributed to the interpretation of the results. J.M.W.G., and R.M.H.M. supervised the project. D.M.V. drafted the manuscript. D.M.V., R.S., E.L., J.M.W.G. and R.M.H.M. revised and edited the manuscript.

2.7 Acknowledgments

This study was financially supported by FrieslandCampina. R.S, E.L., P.V.J., M.P. and J.M.W.G. are currently or were previously employed by FrieslandCampina. The work was carried out in part on the Dutch national e-infrastructure with the support of SURF Cooperative. This work was performed in part using the ALICE compute resources provided by Leiden University.

Chapter 3

2'-Fucosyllactose helps butyrate producers outgrow competitors in infant gut microbiota simulations

Authors
David M. Versluis[a], Ruud Schoemaker[b], Ellen Looijesteijn[b], Jan M. W. Geurts[b], Roeland M. H. Merks[a,c]
Affiliations
[a] Leiden University, Institute of Biology, Leiden, The Netherlands
[b] FrieslandCampina, Amersfoort, the Netherlands
[c] Leiden University, Mathematical Institute, Leiden, The Netherlands

Published in *iScience* Vol. 27 (2024)
doi.org/10.1016/j.isci.2024.109085

Abstract

A reduced capacity for butyrate production by the early infant gut microbiota is associated with negative health effects, such as inflammation and the development of allergies. Here we develop new hypotheses on the effect of the prebiotic galacto-oligosaccharides (GOS) or 2'-fucosyllactose (2'-FL) on butyrate production by the infant gut microbiota using a multiscale, spatiotemporal mathematical model of the infant gut. The model simulates a community of cross-feeding gut bacteria at metabolic detail. It represents the gut microbiome as a grid of bacterial populations that exchange intermediary metabolites, using 20 different subspecies-specific metabolic networks taken from the AGORA database. The simulations predict that both GOS and 2'-FL promote the growth of *Bifidobacterium*, whereas butyrate producing bacteria are only consistently abundant in the presence of propane-1,2-diol, a product of 2'-FL metabolism. The results suggest that in absence of prebiotics or in presence of only GOS, bacterial species, including *Cutibacterium acnes* and *Bacteroides vulgatus*, outcompete butyrate producers by feeding on intermediary metabolites. In presence of 2'-FL, however, production of propane-1,2-diol specifically supports butyrate producers.

3.1 Introduction

Infants develop a complex microbiota shortly after birth, which is important for healthy growth and development [180]. Here we focus on butyrate, a short-chain fatty acid (SCFA) that is produced in significant amounts by the gut bacteria [25] and is absorbed by the gut colonocytes. Production of butyrate by the microbiota has been suggested to improve the health of infants in a number of ways. Firstly, butyrate in the gut is a key energy source for the gut epithelium, making it important for maintaining the gut barrier function [181]. A breakdown of the gut barrier function due to a lack of butyrate is associated with diseases such as inflammatory bowel disease and rectal cancer [181, 182]. Butyrate production in young infants specifically is associated with a reduced risk of allergies and allergy-associated atopic eczema [21, 22, 23]. Infant butyrate producing bacteria provide protection against food allergies when transplanted into a mouse model [183], suggesting causality. Butyrate production is also associated with a reduced risk of colic in infants [24]. Butyrate also modulates the immune system throughout the body, inhibiting inflammation and carcinogenesis [184]. These data suggest it may be desirable to stimulate butyrate production in the infant gut. Using mechanistic computational modeling, here we investigate how stimulation of butyrate producing bacteria may be achieved in the early infant gut microbiota through supplementation with prebiotics.

Microbiota composition and metabolism are influenced by endogenous factors, e.g., gut maturity and inflammation, and exogenous factors, e.g., nutrition, probiotics, and antibiotics. Here we focus on nutrition, which is the primary exogenous factor. Human milk and many infant formulas contain prebiotics such as galacto-oligosaccharides (GOS) and 2'-fucosyllactose (2'-FL), which influence the composition of the gut microbiota and are associated with beneficial health effects for the infant, such as a decreased risk to require antibiotics [53] and reduced manifestation of allergies [185, 186, 187]. It has been hypothesized that some of the health effects associated with prebiotics may be linked to indirect stimulation of butyrate producing bacteria [188, 23]. Thus, both the capacity for butyrate production [21, 23], and prebiotics in nutrition by itself, particularly 2'-FL, have been linked to reduced manifestations of allergies [185, 186, 187].

Butyrate producing bacteria such as *Anaerobutyricum hallii* cannot directly consume GOS or 2'-FL, but they can consume metabolites of GOS or 2'-FL digestion [82]. The primary consumers of GOS and 2'-FL in the infant gut are *Bifidobac-*

terium spp.[64, 63]. Metabolites produced by *Bifidobacterium* spp., in turn, become important food sources for butyrate producing bacteria. For example, *in vitro* it has been found that the butyrate producing bacterium *A. hallii* (formerly *Eubacterium hallii* [189]) can feed on lactate and propane-1,2-diol (1,2-PD), which are metabolites of *Bifidobacterium* spp. [82]. *A. hallii* can also coexist with *Bifidobacterium longum* ssp. *infantis in vitro* on a substrate of glucose or 2'-FL [82].

Despite these in vitro findings that demonstrate potential coexistence of *Bifidobacterium* spp. and butyrate producing bacteria, *in vivo*, i.e. in the infant gut microbiota, butyrate producing bacteria often only have a low abundance and butyrate is found in the feces of only 35% of infants [20]. It is unclear why butyrate producing bacteria and butyrate are not commonly abundant *in vivo*, given that *in vitro* cross-feeding on lactate occurs readily [82], and that lactate-producing *Bifidobacterium* species are abundant in the gut of most infants [3, 31]. Using computational modeling we explore the conditions that may stimulate butyrate producing bacteria *in vivo* in the infant gut. To this end we will compare simulations of simple microbial communities, such as those studied *in vitro*, with simulations of more complex communities that may more closely resemble the *in vivo* situation.

Briefly, the computational model suggests that in simple microbial communities, populations of butyrate producing bacteria can cross-feed on *Bifidobacterium* metabolites. However, in more complex communities the intermediary metabolites are consumed by competitors instead of butyrate producing bacteria. In the presence of 2'-FL, populations of butyrate producing bacteria are nevertheless supported. The mechanism suggested by our simulations is that *Bifidobacterium* produces 1,2-PD from 2'-FL, which specifically feeds butyrate producing species, allowing these to outgrow competing cross-feeders. We provide predictions for interactions in *in vivo* and *in vitro* systems and suggestions for *in vitro* verification of these predictions.

3.2 Results

3.2.1 Model outline

To develop new hypotheses on how oligosaccharides can stimulate the production of butyrate, we further develop a multiscale metabolic model (Fig. 3.1A & B) of the carbon metabolism of the infant gut microbiota [103]. The computational model is based upon our earlier models of the adult and infant microbiota [122, 103]. In

comparison with these previous models, the present model simulates a larger number of small bacterial populations, using a larger, more diverse, and further curated set of metabolic models of gut bacteria from the AGORA database [101]. In particular, we have included the butyrate producers *A. hallii*, *Roseburia inulinivorans* and *Clostridium butyricum* and the digestion of the prebiotic oligosaccharides GOS and 2'-FL by *Bifidobacterium longum* ssp. *infantis*. The complete community model integrates these predictions of metabolism over space and time to create a multiscale model that covers the development and variation of the infant gut microbiota over the first three weeks of life. Other multiscale metabolic modelling techniques have been used previously to model the adult human microbiota in frameworks such as SteadyCom and Comets [131, 132]. The model presented here distinguishes itself from these frameworks by its focus on the infant gut microbiota, by including factors such as prebiotics and the initial presence of oxygen at birth.

Briefly, the spatial model simulates the ecology of an intestinal microbial ecosystem, and features genome-scale metabolic models (GEMs) of intestinal bacteria, spatial structuring, exchange of extracellular metabolites, and population dynamics. The system is simulated on a regular square lattice of 225×8 boxes of 2×2 mm, representing a typical infant colon of 45×1.6 cm. Each box contains a simulated metapopulation of one of 20 bacterial species commonly present in the infant gut [3] (Table 3.1), and concentrations of simulated nutrients and metabolites such as extracellular oligosaccharides and short-chain fatty acids. Based on the concentrations of metabolites, the systems predicts the growth rate for each metapopulation as well as the uptake and excretion rates of metabolites using a GEM taken from AGORA [10], a database of metabolic networks of intestinal bacteria. The system is initialised by distributing, on average, 540 populations over the system at random. Oxygen is introduced during initialisation, and water is always available.

After initialisation, the model is simulated in timesteps representing three minutes of real time. Each timestep of the simulation proceeds as follows. Every 3 hours (i.e., 60 timesteps), a mixture of simulated lactose and/or oligosaccharides is added to the leftmost six columns of lattice sites. Then, each step, the model predicts the metabolism of each local population using flux balance analysis (FBA) based on the metabolites present in the local lattice site, the GEM of the species, and the enzymatic constraint. The enzymatic constraint limits the total amount of metabolism that can be performed by each local population per timestep by limiting the maximum summed flux for each FBA solution. The enzymatic constraint is determined

by the local population size. This approach allows us to model metabolic switches and trade-offs [100, 103]. The FBA solution includes a set of influx rates and efflux rates of metabolites that are used to update the environmental metabolite concentrations. The local populations are assumed to grow at a rate linearly proportional to the rate of ATP production[140], which is predicted by FBA by optimizing for ATP production rates. Populations may create a new population in a neighbouring lattice site if the local population is 200 times the initial size (Fig. 3.1A-1). Populations of more than 400 times the local size, which can only form when density if so high new populations cannot be created, stop metabolism to represent quiescence. Populations spread at random into adjacent lattice sites (Fig. 3.1A-2); metabolites diffuse and advect towards the back of the tube (Fig. 3.1A-3&4). To mimic excretion, metabolites and populations are deleted from the most distal column each timestep. To represent bacterial colonisation, new populations of randomly selected species are introduced into empty lattice sites at a small probability. All parameters are given in table 3.2. Details of the model are given in section Methods.

Table 3.1: Species and subspecies included in the model. color indicates color used in figures.

Name	Phylum	Anaerobic status per [101]	Butyrate producing
Bifidobacterium longum ssp. *infantis*	Actinomycetota	Obligate anaerobe	no
Bifidobacterium longum ssp. *longum*	Actinomycetota	Obligate anaerobe	no
Collinsella aerofaciens	Actinomycetota	Obligate anaerobe	no
Cutibacterium acnes	Actinomycetota	Facultative anaerobe	no
Rothia mucilaginosa	Actinomycetota	Microaerophile	no
Eggerthella sp. *YY7918*	Actinomycetota	Nanaerobe	no
Streptococcus oralis	Bacillota	Facultative anaerobe	no
Staphylococcus epidermidis	Bacillota	Facultative anaerobe	no
Gemella morbillorum	Bacillota	Facultative anaerobe	no
Enterococcus faecalis	Bacillota	Facultative anaerobe	no
Lactobacillus gasseri	Bacillota	Facultative anaerobe	no
Ruminococcus gnavus	Bacillota	Obligate anaerobe	no
Veillonella dispar	Bacillota	Obligate anaerobe	no
Anaerobutyricum hallii	Bacillota	Obligate anaerobe	yes
Roseburia inulinivorans	Bacillota	Obligate anaerobe	yes
Clostridium butyricum	Bacillota	Obligate anaerobe	yes
Parabacteroides distasonis	Bacteroidota	Nanaerobe	no
Bacteroides vulgatus	Bacteroidota	Nanaerobe	no
Haemophilus parainfluenzae	Pseudomonadota	Aerobe	no
Escherichia coli SE11	Pseudomonadota	Facultative anaerobe	no

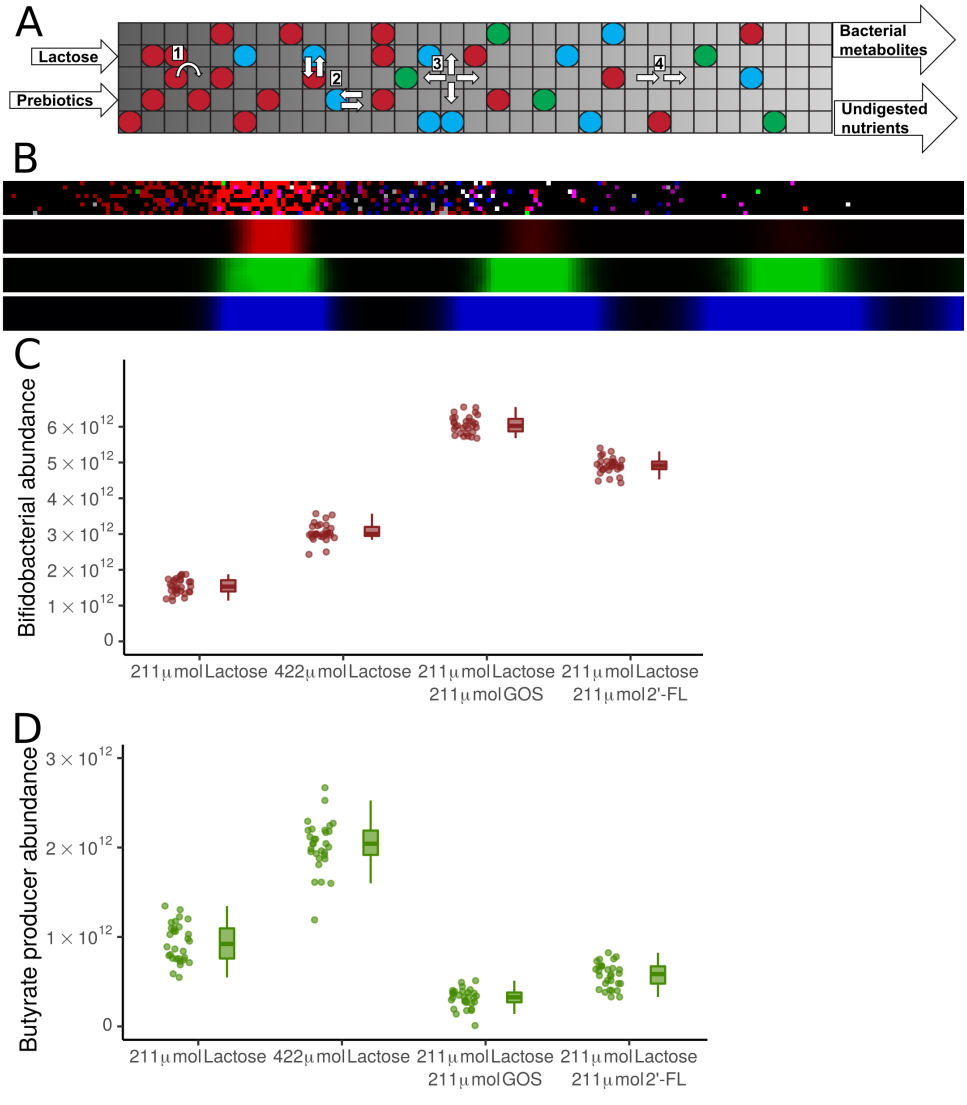


Figure 3.1: Model predicts coexistence of *Bifidobacterium* and butyrate producing bacteria in absence of competition
(A) Schematic of the model. Circles represent bacterial populations, color represents species. Flow through the tube is from left (proximal) to right (distal). Nutrients entered the system proximally. All metabolites leave the system distally. Lattice dimensions are schematic.
(B) Screenshots of the model at a single time point, showing, from top to bottom, the bacterial layer, lactose, lactate, and acetate. Brightness indicates growth in the bacterial layer, and concentration in the metabolic layers.
(C,D) Abundance of (C) *Bifidobacterium* spp., (D) butyrate producing bacteria, at the end of 21 days for 30 sets of simulations with no prebiotics, no prebiotics and additional lactose, with GOS, or with 2'-FL at the end of 21 days. n=30 for each condition, each simulation is represented by one dot.

3.2.2 Model with simplified consortium of species predicts coexistence of butyrate producing bacteria and *Bifidobacterium*

We first simulated the model using a simplified consortium of species, the two *Bifidobacterium longum* subspecies (table 3.1) and three butyrate producing species: *Anaerobutyricum hallii*, *Clostridium butyricum*, and *Roseburia inulinivorans*. We performed 30 simulations for each of four conditions, in which the following sugars were introduced every three simulated hours: (1) 211 µmol lactose and no prebiotics, (2) 422 µmol lactose and no prebiotics, (3) 211 µmol lactose plus 211 µmol GOS, and (4) 211 µmol lactose plus 211 µmol 2'-FL. We estimated 211 µmol lactose to be a realistic amount of lactose to reach the infant colon, given infant intake and small intestinal uptake [66, 131]. As there is little absorption by the small intestine of prebiotics [190], the amount of prebiotics in the nutrition consumed by the infant would be much smaller than the amount of lactose. We also include the 422 µmol lactose condition to control for the possibility that effects in the conditions with prebiotics are due the larger amount of sugar present conditions, instead of their type. The condition with 422 µmol lactose does not correspond to an *in vivo* condition. We analyzed the abundance of each species at the end of 10080 timesteps, representing 21 simulated days. In each of the four conditions *Bifidobacterium* bacteria (Fig. 3.1C) and butyrate producing bacteria coexisted (Fig. 3.1D), and, paradoxically, butyrate

Table 3.2: Parameters of the model

Parameter	Value	Unit
Lattice side length	2	mm
Width of lattice	225	lattice sites
Height of lattice	8	lattice sites
Timestep	180	seconds
Average number of initial populations	540	-
New population placement probability	0.00005	per timestep per empty lattice site
Population death probability	0.0075	per timestep per population
Initial size per population	$5 \cdot 10^7$	no. of bacteria
Population size to create a new population	$1 \cdot 10^{10}$	no. of bacteria
Maximum population size	$2 \cdot 10^{10}$	no. of bacteria
ATP to grow one cell	$1 \cdot 10^{-15}$	mol
Enzymatic constraint	2	µmol flux per timestep per $1 \cdot 10^{10}$ bacteria
Nutrient input	211	µmol per nutrient per 60 timesteps
Initial oxygen	0.1	µmol per lattice site
Metabolic advection	2	mm per timestep
Diffusion (metabolites and bacteria)	$6.3 \cdot 10^5$	square cm per second

producing bacteria were reduced in presence of prebiotics.

3.2.3 In presence of competitors, model predicts coexistence of butyrate producing bacteria and *Bifidobacterium* in the presence of 2'FL but not in presence of GOS

We next examined the behaviour of the system in the presence of a more complex consortium, consisting of all 20 species and subspecies listed in Table 3.1, simulating the same four conditions. The selection of these species is described in methods section 'Species Composition'. In absence of prebiotics, regardless of the quantity of lactose, the model predicted that *Bifidobacterium*, *Bacteroides* and *Escherichia* became the most abundant genera after three weeks (Fig. 3.2A, S1 Video), consistent with *in vivo* observation [3, 31]. We also observed some abundance of Bacilli in accordance with *in vivo* observations [3, 33, 31]. The higher quantity of lactose resulted in a higher average abundance for all major groups. In absence of prebiotics, butyrate producing bacteria achieved a combined abundance over $1 \cdot 10^{10}$ in only 4 of the 30 simulations with 211 µmol of lactose per 3 hours, and 6 of the 30 with 422 µmol of lactose (Fig. 3.2B). In the remaining simulations, the butyrate producing bacteria remained almost absent, staying below $1 \cdot 10^{10}$ bacteria. In the simulations with GOS, *Bifidobacterium* was more abundant than in the condition without prebiotics ($p<0.001$,Fig. 3.2A) whereas the butyrate producing bacteria were not affected ($p=0.18$) (Fig. 3.2B). With GOS, butyrate producing bacteria also had a combined abundance of over $1 \cdot 10^{10}$ bacteria at the end of 13 of the 30 simulations (Fig. 3.2B). Interestingly, in the condition with 2'-FL the abundance of butyrate producing bacteria was over $1 \cdot 10^{10}$ bacteria at the end of 19 of 30 simulations (Fig. 3.2B), and the butyrate producing species were more abundant (Fig. 3.2A, S2 Video) than in the other conditions. Thus 2'-FL but not GOS stimulated butyrate producing bacteria in the complex community. To test for any concentration-dependence or cross-talk between 2'-FL and GOS we next performed sets of 30 simulations in presence of 211 µmol lactose and levels of 2'-FL and GOS varying between 21.1 µmol to 211 µmol per three hours and combinations thereof (Fig. S1). The amount of 2'-FL ($p=0.017$, Kruskal-Wallis rank sum test) but not that of GOS ($p=0.658$, Kruskal-Wallis rank sum test) affected the abundance of butyrate producing bacteria, further supporting the prediction that 2'-FL but not GOS stimulates butyrate producing bacteria in the complex community.

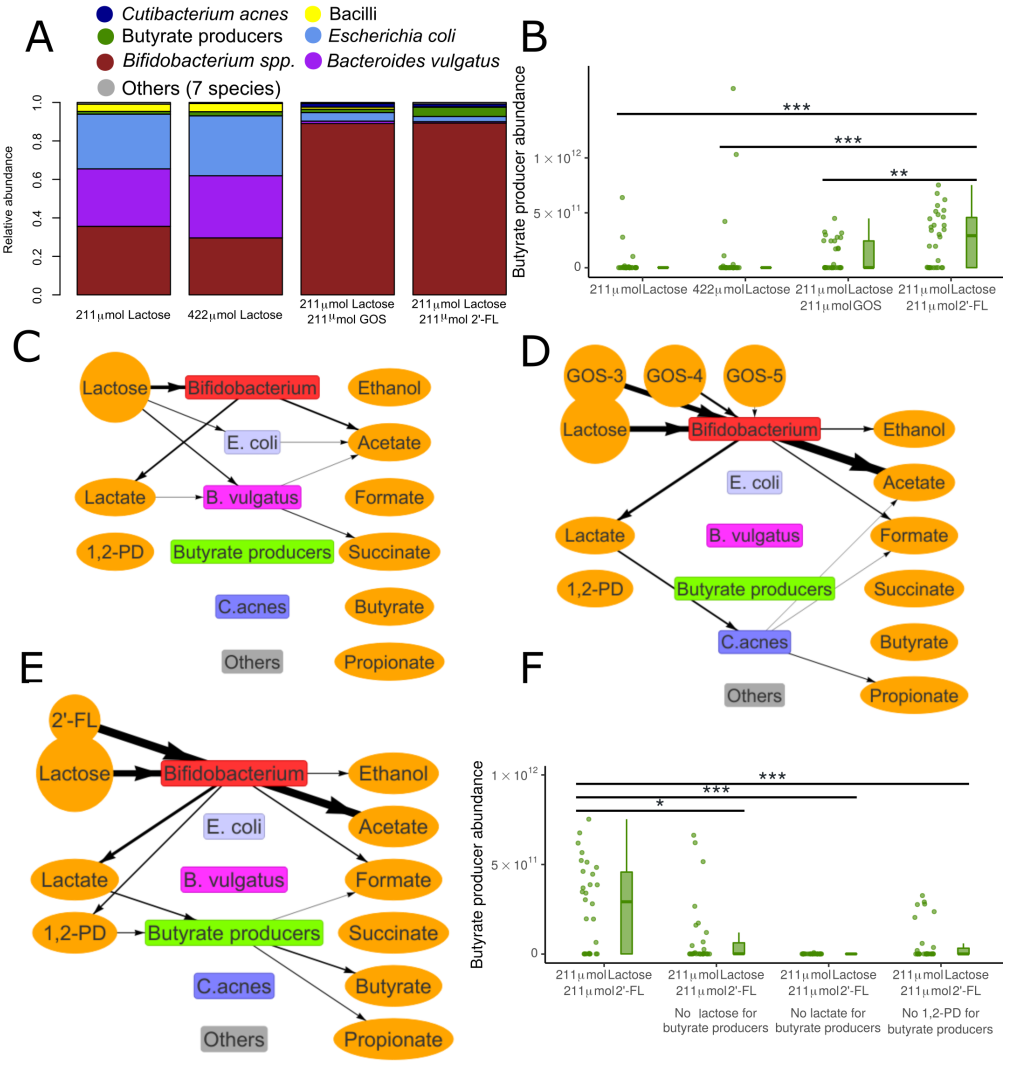


Figure 3.2: Unlike GOS, 2'-FL leads to stimulation of butyrate producing bacteria through 1,2-PD in the full simulated microbiota

(A) Relative abundance of bacterial species in the condition with no prebiotics, no prebiotics and additional lactose, with GOS, or with 2'-FL at the end of 21 days. n=30 for each condition, each simulation is weighed equally. The key to the species in each group is in table 3.1.

(B) Abundance of butyrate producing bacteria at the end of 21 days for the four conditions of A. n=30 for each condition. Each simulation is represented by one dot. $p<0.001$ for 2'-FL compared to no prebiotics and no prebiotics with additional lactose. $p=0.004$ for 2'-FL compared to GOS.

(C,D,E) Visualisation of metabolic interactions in a sample simulation (C) without prebiotics (211 µmol lactose per three hours) (D) with GOS (DP3,DP4, and DP5 displayed separately) (E) with 2'-FL. Line width is scaled with the flux per metabolite over the last 60 timesteps, multiplied by the carbon content of the molecule, with a minimum threshold of 100 µmol atomic carbon. Data from last 3 hours, step 10020 to 10080. Circles indicate nutrients.

(F) Abundance of butyrate producing bacteria with 2'FL at the end of 21 days. Uptake of lactose, lactate, or 1,2-PD by butyrate producing bacteria is disabled in the 'no lactose', 'no lactate', and 'no 1,2-PD' conditions respectively. $p=0.010$,$p<0.001$,$p<0.001$ for each disabled uptake compared to the baseline, respectively n=30 for each condition. Each simulation is represented by one dot. NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$

In order to investigate why 2'-FL led to a more consistent abundance of butyrate producing bacteria we analysed the metabolic interactions between bacterial species. We visualised the network of metabolic fluxes between the bacteria using arrows between species and metabolite pools in Fig. 3.2C-E. The resulting diagrams show both primary consumption, i.e., uptake of nutrients such as lactose, GOS, and 2'-FL, and cross-feeding, i.e., uptake of metabolites produced by other species. Sample visualisations of the condition without prebiotics (211 µmol lactose) (Fig. 3.2C, S3 Video) and the condition with GOS (Fig. 3.2D) revealed co-occurrence of species and cross-feeding, but no butyrate production. In these simulations the cross-feeding metabolite lactate, which is a known substrate for butyrate producing bacteria [82], was consumed by *Bacteroides vulgatus* and *Cutibacterium acnes*, respectively. Butyrate formation only occurred in the sample simulation with 2'-FL (Fig. 3.2E). Only in presence of 2'-FL and not in the other conditions, was a flux of 1,2-PD directed towards the butyrate producing species (Fig. 3.2E and S4 Video). We therefore hypothesised that butyrate producing species may be more abundant in the model simulations with 2'-FL, because 2'-FL digestion by *Bifidobacterium* produces 1,2-PD as a cross-feeding substrate. 1,2-PD is a known *Bifidobacterium* metabolite from 2'-FL *in vitro* [82]. To test this hypothesis, we performed new sets of simulations with 2'-FL in which we blocked the uptake by butyrate producing bacteria of either lactose, lactate, or 1,2-PD, i.e, the uptake of metabolites most consumed by butyrate producing bacteria was disabled. Indeed, blocking the uptake of any of these metabolites led to a reduction of butyrate producing bacteria (Fig. 3.2F). Thus a flux of lactose, lactate, but also 1,2-PD that is only produced in presence of 2'-FL, was required for sustaining butyrate producing bacteria in our simulations.

We next turned to the model with the simplified consortium of species, the two *Bifidobacterium* subspecies and three butyrate producing species, to test if uptake of lactose, lactate and 1,2-PD was also required for butyrate producing bacteria to become abundant with this consortium. After blocking the uptake of lactose, lactate, or 1,2-PD by butyrate producing bacteria, the abundance of butyrate producing bacteria was reduced at the end of the simulations compared to the control (Fig. 3.3A). Surprisingly, however, and in contrast to the complete system (Fig. 3.2F), butyrate producing populations retained an abundance over $1 \cdot 10^{10}$ bacteria in respectively 27 and 30 of 30 simulations when lactose or 1,2-PD uptake was disabled. Thus neither lactose nor 1,2-PD were essential for butyrate producing bacteria. Altogether, 1,2-PD, and thus 2'-FL, was required for butyrate producing bacteria in the com-

plex system, but not in the simplified system. Thus these model simulations suggest that supplementation with 2'-FL introduces a flux of 1,2-PD from *Bifidobacterium* spp. to butyrate producing bacteria that prevents competitive exclusion of butyrate producers by competitors such as *B. vulgatus* (Fig. 3.2C) or *C. acnes* (Fig. 3.2D).

3.2.4 *Bacteroides vulgatus* and *Cutibacterium acnes* are effective competitors on different substrates

In the 2'-FL condition butyrate producing bacteria fed on lactate and 1,2-PD (Fig. 3.2E). In the conditions without 2'-FL no 1,2-PD was produced and lactate was consumed by *B. vulgatus* or *C. acnes* (Fig. 3.2C&D). This suggests that, in absence of 1,2-PD, *B. vulgatus* and *C. acnes* outcompete the butyrate producing bacteria for lactate. To investigate whether these species could indeed be responsible for outcompeting butyrate producing bacteria we again turned to the model with the simplified consortium and added the potential competitors *B. vulgatus* and *C. acnes* to the consortium one by one. First we studied the simplified consortium in absence of prebiotics in the conditions with 211 µmol and 422 µmol lactose per three hours. The abundance of butyrate producing bacteria was reduced in presence of *B. vulgatus* but not in presence of *C. acnes* (Fig. 3.3B, 422 µmol visualized in Fig. S3). After blocking lactose or lactate uptake by *B. vulgatus* in the condition with 211 µmol lactose, the abundance of butyrate producing bacteria was restored (Fig. 3.3B), indicating that *B. vulgatus* required both lactose and lactate to effectively outcompete the butyrate producing bacteria.

In the conditions with GOS, the situation was reversed: *C. acnes* but not *B. vulgatus* outcompeted butyrate producing bacteria (Fig. 3.3C). After blocking uptake of lactate by *C. acnes* the abundance of butyrate producing bacteria was restored (Fig. 3.3C). *C. acnes* does not use lactose in the model. Taken together, these simulations suggest that lactate is required for competitive exclusion of butyrate producing bacteria by *C. acnes*. In the condition with 2'-FL *B. vulgatus* did not outcompete butyrate producing bacteria (Fig. 3.3D). *C. acnes* ($p=0.001$) moderately suppressed butyrate producing bacteria, with 29 of 30 simulations still predicting an abundance of butyrate producing bacteria of over $1 \cdot 10^{10}$ bacteria. This agrees with the simulations using the full consortium (Fig. 3.2B), which also displayed a robust abundance of butyrate producing bacteria in the 2'-FL condition.

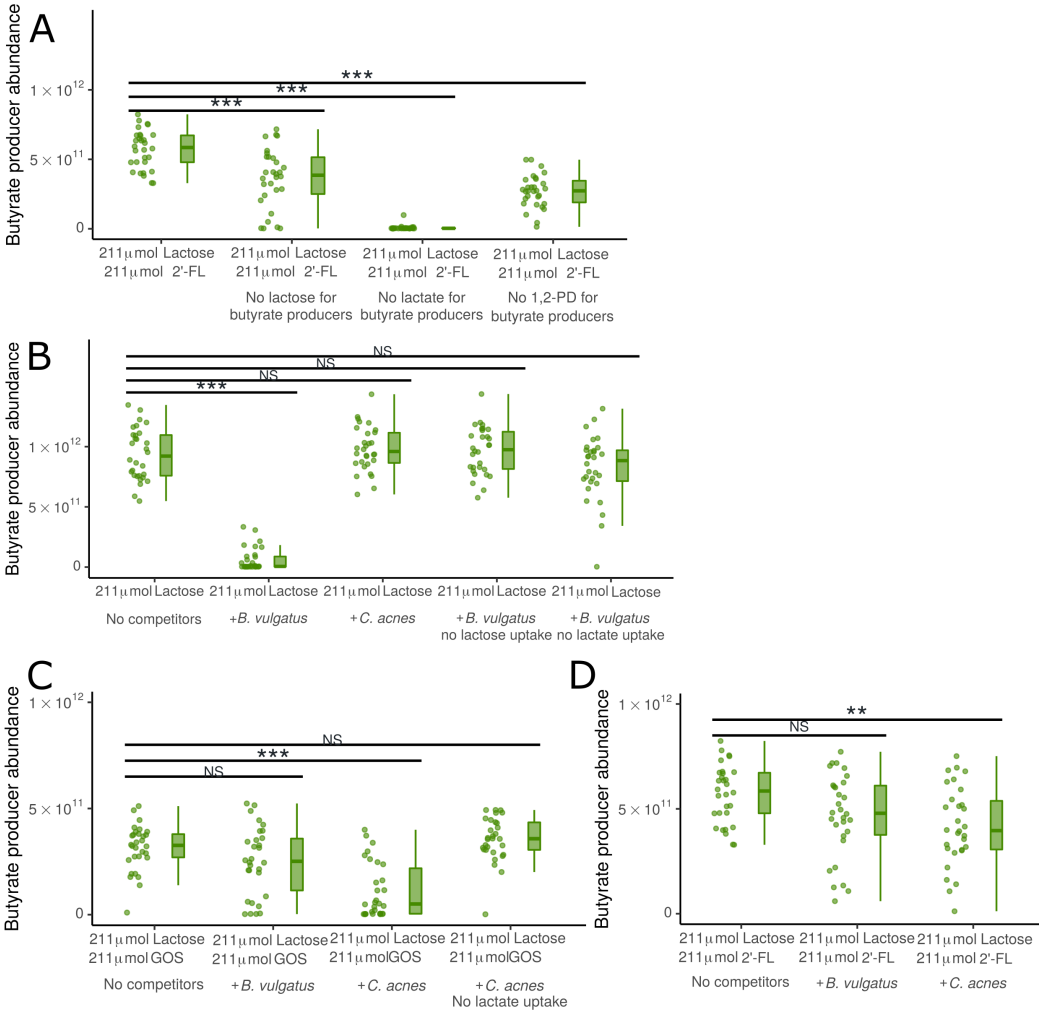


Figure 3.3: 2'-FL makes butyrate producing bacteria resistant to competition by other infant gut bacteria

(A) Abundance of butyrate producers with 2'-FL and without competitors (only *Bifidobacterium* and butyrate producers) at the end of 21 days. Uptake of lactose, lactate, or 1,2-PD is disabled for butyrate producers in the 'no lactose', 'no lactate', and 'no 1,2-PD' conditions respectively. n=30 for each condition. Each simulation is represented by one dot. ($p<0.001$ for each disabled uptake compared to the baseline)

(B,C,D) Abundance of butyrate producers at the end of 21 days (B) without prebiotics, either without competitors (only *Bifidobacterium* and butyrate producers), with addition of *B. vulgatus*, with addition of *B. vulgatus* unable to take up either lactose or lactate, or with addition of *C. acnes*. n=30 for each condition. Each simulation is represented by one dot. $p<0.001$ for abundance of butyrate producers with *B. vulgatus* compared to no competitors

(C) with GOS, either without competitors (only *Bifidobacterium* and butyrate producers), with addition of *C. acnes*, with addition of *C. acnes* unable to take up lactate, or with addition of *B. vulgatus*. n=30 for each condition. Each simulation is represented by one dot. $p<0.001$ for abundance of butyrate producers with *C. acnes* compared to no competitors

(D) with 2'-FL, either without competitors (only *Bifidobacterium* and butyrate producers), with addition of *C. acnes*, or with addition of *B. vulgatus*. n=30 for each condition. Each simulation is represented by one dot. $p=0.001$ for abundance of butyrate producers with *C. acnes* compared to no competitors. NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$

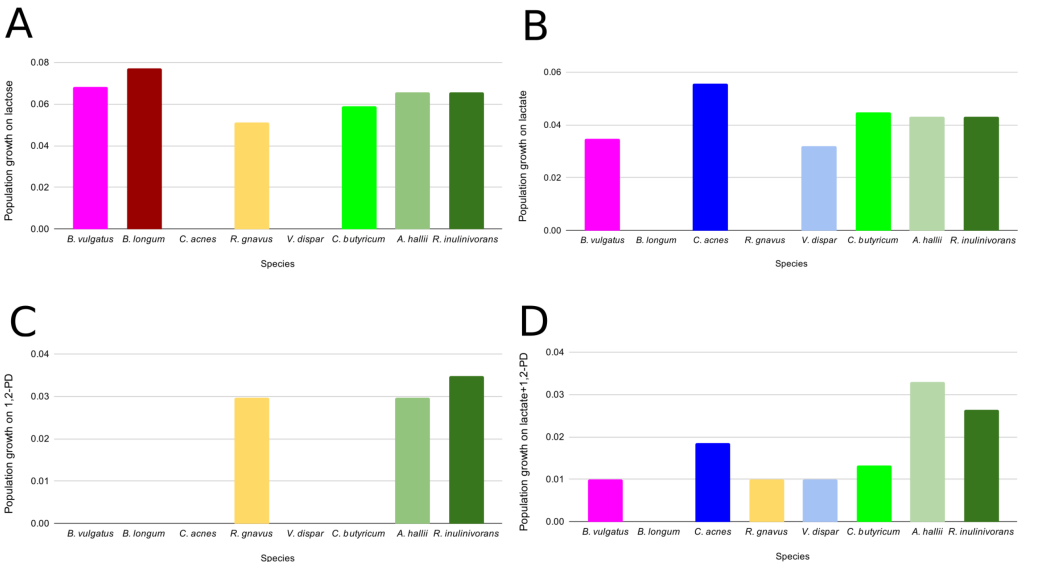


Figure 3.4: Populations of butyrate producing bacteria only grow much faster than their competitors on a mixed substrate of 1,2-PD and lactate
(A) Growth on unlimited lactose and water over a single timestep for butyrate producing bacteria (three rightmost bars, in green) compared to other lactose-fermenting bacteria in the model.
(B) Growth on unlimited lactate and water over a single timestep for butyrate producing bacteria (three rightmost bars, in green) compared to other lactate-fermenting bacteria in the model.
(C) Growth on unlimited 1,2-PD, acetate, and water over a single timestep for butyrate producing bacteria (two rightmost bars, in green) compared to another 1,2-PD-fermenting bacterial species in the model.
(D) Growth on 1 µmol per ml of 1,2-PD and lactate, and unlimited acetate and water, over a single timestep for butyrate producing bacteria (three rightmost bars, in green) compared to other bacteria in the model for populations of $5 \cdot 10^9$ bacteria with access to one lattice site (0.05ml)

3.2.5 Butyrate producing bacteria can use a mixture of lactate and 1,2-PD as substrates in the 2'-FL condition to grow faster than their competitors

To analyse how butyrate producing bacteria can outcompete other species only in the presence of 2'-FL but not in the presence of GOS or without prebiotics, we next examined the growth rates per timestep on unlimited quantities of the three key substrates for butyrate producing bacteria indicated above: lactose, lactate, and 1,2-

PD. With unlimited availability of lactose, the growth of the three butyrate producing species was reduced relative to the growth of most other species (Fig. 3.4A). With unlimited lactate, growth for butyrate producing species was superior to the other species, but not to *C. acnes* (Fig. 3.4B). In presence of unlimited 1,2-PD and acetate the butyrate producing species *A. hallii* and *Roseburia inulinivorans* grew faster than the other species (Fig. 3.4C). On a mixture of limited lactate and 1,2-PD, with acetate available, two of the three butyrate producing species also grew faster compared to all other species (Fig. 3.4D). Thus the unique ability of butyrate producing bacteria to grow on 1,2-PD and acetate in the model allowed them to outcompete other lactate-consuming species in environments with 1,2-PD, such as those where *Bifidobacterium* consumes 2'-FL. However, they would be unable to outcompete the same species in conditions without 1,2-PD.

3.2.6 Sensitivity analysis

Finally, to test the generality of our observations we performed a sensitivity analysis on the system. The enzymatic constraint (Fig. S2A&B), the death rate and growth rate (through the ATP required per population unit) (Fig. S2C&D), the placement of new populations of random species in empty lattice sites (colonization) (Fig. S2E&F), the diffusion of metabolites and populations (Fig. S2G&H),the amount of initial oxygen (Fig. S2I&J), and quiescence for large populations (Fig. S2K) were varied. We used three conditions for most changed parameters: 211 µmol lactose, 211 µmol lactose plus 211 µmol GOS, and 211 µmol lactose plus 211 µmol 2'-FL per three hours. We only used the latter two for disabling quiescence, as no populations entered quiescence during our initial simulations with 211 µmol lactose. We found minor sensitivity for most parameter changes (Fig. S2). We found the most notable effects when we disabled colonization or initial oxygen. When we disabled colonization the abundance of butyrate producing bacteria was lower in all three conditions ($p<0.001$ for all). The absence of initial oxygen increased the abundance of butyrate producing bacteria in the condition without prebiotics and with 2'-FL ($p=0.002$,$p=0.035$). This indicates that the presence of initial oxygen and sustained colonization are particularly important in the simulated system.

3.3 Discussion

This paper describes a computational study of the effects of the prebiotics GOS and 2'-FL on butyrate producing bacteria in the infant gut microbiota. We have used the model to generate novel hypotheses to explain the — sometimes counter-intuitive — mechanisms at the biochemical and population level that underlie the effects of prebiotics. The model predicts that butyrate producing bacteria can coexist with *Bifidobacterium* in the infant gut with or without GOS or 2'-FL as long as no other bacterial species are present. As soon as other bacterial species are introduced into the model, we found that they can act as competitors, thus reducing the abundance of butyrate producing bacteria. Specifically, the model predicts that *B. vulgatus* outcompetes butyrate producing bacteria in absence of prebiotics. The predicted mechanism is that *B. vulgatus* consumes lactose and lactate, important food sources of the butyrate producing species. In presence of GOS, the model predicts that *C. acnes* becomes the key competitor of the butyrate producing bacteria due to its lactate consumption. In presence of 2'-FL, however, the butyrate producing species are no longer outcompeted. The mechanism as predicted by the model is as follows. The breakdown of 2'-FL by *Bifidobacterium* produces 1,2-PD. 1,2-PD becomes an additional food source for the butyrate producing bacteria, helping them to outgrow competitors. Thus, our modeling results predict that only 2'-FL, but not GOS supports populations of butyrate producing bacteria in their competition against species such as *B. vulgatus* and *C. acnes*.

The following *in vitro* and *in vivo* observations agree with these model predictions. Firstly, the model predicts co-existence and crossfeeding between *Bifidobacterium* and butyrate producing species on 2'-FL. In agreement with the model predictions, co-existence of and cross-feeding between *Bifidobacterium* and butyrate producing bacteria occurs *in vitro* within simplified, synthetic communities on glucose, fucose, and 2'-FL in the absence of competitors [82]. Secondly, the model predicts that in presence of the competitors such as *B. vulgatus* and *C. acnes*, *B. vulgatus* will become abundant in absence of prebiotics and outcompete butyrate producing species. In agreement with this model prediction, *B. vulgatus* is often abundant in the *in vivo* infant gut microbiota [3], and it can consume lactose *in vitro* [191]. No information is available on lactate consumption of *B. vulgatus*, but the related *Bacteroides fragilis* is able to consume lactate *in vitro* [192]. Thirdly, the model predicts that *C. acnes* outcompetes butyrate producing bacteria in presence of GOS by consuming

lactate. In agreement with this prediction, *C. acnes* is found in 22% of infants [3] and *Cutibacterium avidum*, closely related to *C. acnes* [193], reduces the abundance of the butyrate producer *A. hallii* in an *in vitro* lactate-fed microbiota from infant fecal samples [194]. Both *C. acnes* and *C. avidum* consume lactate *in vitro* [195]. Finally, the model predicts that butyrate producing bacteria become competitive through cross-feeding on 1,2-PD, which is produced by *Bifidobacterium longum* from 2'-FL. In agreement with this prediction, the butyrate producer *A. hallii* cross-feeds on 1,2-PD in an *in vitro* synthetic community of *A. hallii* and *B. longum* [82]. Also in line with this prediction, 2'-FL supplementation increased the abundance of butyrate producing bacteria in *in vitro* fecal communities based on infant fecal samples, which likely include key competitors of butyrate producing species [188]. An *in vitro* colonic fermentation model inoculated with infant feces has previously been used to study the effect of introducing specific competitors to a lactate-consuming infant gut microbiota [194]. This approach could also be used to test if *B. vulgatus* and *C. acnes* are viable competitors in the infant gut and if the presence of 1,2-PD allows butyrate producing species to outcompete other bacteria.

More broadly, the model simulations without prebiotics predict that *Escherichia*, *Bacteroides*, and *Bifidobacterium* become the three most abundant genera, which agrees with the most abundant genera found in the infant gut microbiota around the age of three weeks [3, 31]. The relative abundances the model predicts for butyrate producing species range from 1.4% without prebiotics to 4.8% with 2'-FL, both of which are within the broad range of values reported for the butyrate producing community [20]. However, for two less abundant groups, Bacilli and *Veillonella*, the model predictions disagree with *in vivo* data. Firstly, an initially dominant Bacilli phase is sometimes seen *in vivo*, e.g. in 17.6% of subjects in [33], but not in any model outcomes. An initially dominant Bacilli phase is associated in non-premature infants with a shorter gestational period [33], but it is unclear exactly what factors are responsible. A similar initial dominance of Bacilli that often occurs in premature infants has been hypothesised to be related to selection pressures by the immune system, a different composition of initial colonizers [196], or a defective mucin barrier [16]. Secondly, the model predicted a very low *Veillonella dispar* abundance in all conditions. These predictions contradict *in vivo* data [91, 3] in which *V. dispar* is relatively abundant. *V. dispar* likely has a lower abundance in the model due to an incorrectly reduced growth rate relative to the other species in the model on lactate, the primary energy source of *V. dispar* [95], lactate, (Fig. 3.4B). We do not

expect a large influence on the overall model predictions from this discrepancy, as *C. acnes* has a metabolism similar to that of *V. dispar* in the model and *in vitro*: both produce propionate, consume lactate, and cannot consume lactose [195]. However, we cannot exclude that other species in the model, such as *Veillonella* spp., may be more important competitors *in vivo* than the competitors that the model predicts.

Potential sources of the discrepancies between model predictions and experimental data include: (1) errors in the metabolic predictions of the underlying FBA models; (2) computational errors, and (3) incomplete representation of the biology underlying infant digestion. A typical error occurring in FBA models is an incomplete prediction of metabolic shifts, which is in part due to the assumption of FBA models that the growth rate or energy production is optimised [13]. For example, the FBA model does not correctly predict the metabolic shift from high-yield to low-yield metabolism as observed *in vitro* in *Bifidobacterium* growing on increasing concentrations of GOS and 2'-FL [96, 90]. FBA only predicts high-yield metabolism. The model, therefore, likely underestimates total lactate production. The effects of this discrepancy on the results are difficult to predict, but as lactate is a cross-feeding substrate, the underestimation of lactate may cause the model to underestimate the abundance of cross-feeding species such as *C. acnes* or butyrate producing bacteria. The optimality assumption of FBA also ignores any other 'task' that a bacterium has, besides growth. For example, sporulation, toxin production, or metabolic anticipation [197] may limit biomass production. The model does not represent such genetically regulated mechanisms.

Further errors in the model predictions can be due to simplifications in the FBA model. For example, we assume that the total flux through the reaction networks is capped (Eq. 3.4), so as to mimic the maximum volume in a cell that can be filled with enzymes. Here each enzyme is assumed to have equal maximum flux, and the optimization problem then predicts the optimal relative flux distribution. In reality, due to differences in enzyme concentration and enzyme efficiency these maximum fluxes can of course differ, which affects the predictions of FBA [133, 144]. If species-specific data on efficiency and genetic regulation of pathways become available, such weighting could be included in the model. The metabolic predictions from the FBA layer could be further improved in future versions of the model by integrating thermodynamic plausibility and favorability into FBA, which have previously improved metabolic predictions for intracellular metabolism [198, 199]. Additionally, the FBA model includes an extracellular compartment in which long GOS chains are

broken down to shorter GOS chains, but it is not possible for extracellular breakdown products to diffuse during this process. Such extracellular digestion may lead to additional competition effects, because competitors may 'steal' digestion products without investing in the enzymes themselves [88]. Such effects may become important if additional species are introduced in the model that digest prebiotics extracellularly, such as *Bifidobacterium bifidum* [64].

Computational errors in the model (2) include the discretization of time, the discretization of space, and rounding errors in the FBA solver. Firstly, all processes in the model are assumed to be constant within each timestep, which means the model only roughly approximates the continuous temporal dynamics of processes such as metabolism and diffusion. Secondly, we discretize the three-dimensional continuous cylindrical space of the gut into a two-dimensional rectangular grid of lattice sites. We consider each lattice site to be of equal volume and to have equal flow through it. This simplification introduces many errors, as lattice sites must represent different shapes of three-dimensional space, and these shapes are not connected as they would be in three-dimensional space. It is difficult to estimate what impact these discretizations have on the model. Finally, the FBA solver uses floating point arithmetic to generate a deterministic but not exact solution to each FBA problem. These distortions are very small, typically on the order of 10^{-15} µmol per metabolite per FBA solution, so we do not expect a notable effect on the results.

Errors in the model predictions due to an incomplete representation of the biology underlying infant digestion (3) include missing organisms, missing ecological interactions, the simplifications we made to the metabolic input, and missing representation of host interactions. Firstly, the model does not include fungi or archaea in the infant gut. Both groups occur at a lower absolute abundance than the bacterial microbiota, but may still influence it [200]. Secondly, the model does not include interactions between bacteria other than cross-feeding and competition for resources. Missing interactions include acidification of the gut [159], the production of bacteriocins [201] and the effects of phage infections [202], all of which have species-specific effects. Thirdly, the model does not include the input of fats, proteins, or minerals into the gut. This means that the model cannot represent stimulation of growth by digestion of fats or proteins, nor potential limits on growth due to, for example, the lack of iron [203] or essential amino acids [204]. Finally, the model does not represent the interactions of the host with the microbiota, such as the continuous secretion by the gut wall of mucin [30] and oxygen [152], and the uptake of short-chain fatty

acids [205]. Colonic mucins in particular could greatly influence the microbiota, as *B. bifidum* consumes colonic mucins extracellularly, which facilitates cross-feeding by butyrate producing bacteria *in vitro* [27].

Despite the inevitable limitations of the model, we have shown here how the model can be used to produce testable predictions on the effects of prebiotics and competition on butyrate producing bacteria in the infant gut microbiota. Future versions of the model may be a useful help in follow-up studies on the effects of nutrition on bacterial population dynamics in the infant and adult gut microbiota.

3.4 Methods

We used a spatially explicit model to represent the newborn infant gut microbiota. The model is based on our earlier models of a general microbiota [122] and the infant microbiota [103]. Prebiotic digestion is the most important addition in the present version of the model.

The model consists of a regular square lattice of 225×8 lattice sites, with each lattice site representing 2×2 mm of space. Taken together this represents an infant colon of 450×16 mm, in line with *in vivo* estimates [135, 136]. Each lattice site can contain an amount of the 735 metabolites represented in the model, as well as a single bacterial population.

3.4.1 Species Composition

Species were selected based on [3], using sheet 2 of their Table S3. We selected the 20 entries with the highest prevalence in vaginally delivered newborns. After removing two duplicate entries we selected a representative species for each genus from the AGORA database [10]. We then added an additional *Bifidobacterium longum* ssp. *infantis* GEM to serve as prebiotic consumer, and a *Roseburia inulinivorans* GEM. *Roseburia* spp. have been shown to be a prevalent butyrate producing bacterium in infants in other studies [20]. Together, these form the list of species (Table 3.1). This represents an improvement over earlier work [103], where we used a smaller set of species based on a smaller dataset of infant microbiota data. This earlier set did not include butyrate producers, and so did not allow us to examine cross-feeding that leads to butyrate production. This extended set allows us to examine these cross-feeding interactions.

3.4.2 Changes from AGORA

The model uses GEMs generated in the AGORA project [101]. We have applied various changes and additions to these models (Table S1).

We have added digestion of GOS or 2'-FL to the *B. longum* ssp. *infantis* GEM as follows. 2'-FL digestion was implemented by adding reactions representing an ABC-transporter and an intracellular fucosidase that breaks 2'-FL down to lactose and fucose [96]. GOS was represented through separate DP3, DP4, and DP5 fractions [206]. The DP4 and DP5 fractions are broken down extracellularly to DP3 and DP4 fractions respectively, releasing one galactose molecule in the process [207]. The DP3 fraction is taken up with an ABC transporter, and broken down internally to lactose and galactose [207].

We have also further expanded earlier curation of the AGORA GEMs [103]. We disabled anaerobic L-lactate uptake for the *Bifidobacterium* GEMs and for *E. coli* in line with available literature [158, 208]. To have the GEMs correspond with existing literature on lactose uptake we added a lactose symporter to *Anaerobutyricum hallii* [27], both *Bifidobacterium longum* GEMs [209], *Roseburia inulinivorans* [210], *Haemophilus parainfluenzae* [211], and *Rothia mucilaginosa* [212]. We also added galactose metabolism to *R. inulinivorans* [213] and *R. mucilaginosa* [212]. Further changes were made to prevent unrealistic growths and the destruction of atoms within reactions (Table S1).

3.4.3 Validity checks

After applying the changes in Table S1 we tested all GEMs individually for growth on a substrate of lactose and water. In line with literature, this did not lead to growth for *Veillonella disparans* [95], *Cutibacterium acnes* [195], *Eggerthella* sp. YY7918 [214], and *Gemella morbillorum* [215]. All other species grew on this substrate. We also checked for any spurious growth by checking each GEM for growth with only water present.

During each simulation, the model checks the FBA solutions for thermodynamic plausibility. The model uses a database of Gibbs free energy values [173] for all metabolites except 2'-FL and GOS. Values for 2'-FL and GOS were generated from the values for lactose and fucose, and lactose and galactose, respectively. Separate values were generated for the separate fractions of GOS. All values assumed a pH of 7 and an ionic strength of 0.1 M. We found that in the simulations of Fig. 3.2A with the

baseline level of lactose, combined with those with GOS and 2'-FL (n=90) 99.98% of all FBA solutions had a lower or equal amount of Gibbs free energy in the output compared to the input. The remaining 0.02% of FBA solutions was responsible for 0.003% of total bacterial growth.

3.4.4 FBA with enzymatic constraint

Although other aspects of the model were changed, the FBA approach we used is identical to that used in the earlier model [103]. The model uses a modified version of flux balance analysis with an enzymatic constraint to calculate the metabolic inputs and outputs of each population at each timestep [13, 100]. Each GEM is first converted to a stoichiometric matrix S. Reversible reactions are converted to two irreversible reactions, so that flux is always greater than or equal to 0. Reactions identified in the GEM as 'exchange', 'sink', or 'demand' in the GEM are also recorded as 'exchange' reactions. These exchange reactions are allowed to take up or deposit metabolites into the environment. Each timestep, all reactions are assumed to be in internal steady state:

$$S \cdot \vec{f} = 0, \tag{3.1}$$

where $\vec{f}$ is a vector of the metabolic fluxes through each reaction in the network, in mol per time unit per population unit.

Each exchange reaction that takes up metabolites from the environment F_{in} is constrained by an upper bound F_{ub} which represents the availability of metabolites from the environment. It is determined as follows:

$$\vec{F_{in}} \leq \vec{F}_{ub}, \tag{3.2}$$

where $\vec{F_{in}}$ is a vector of fluxes between the environment and the bacterial population. $\vec{F}_{ub}$ is a vector of upper bounds on these fluxes. $\vec{F}_{ub}$ is set dynamically at each timestep t by the spatial environment at each lattice site $\vec{x}$:

$$\vec{F}_{ub}(\vec{x}, t) = \frac{\vec{c}(\vec{x}, t)}{B(\vec{x}, t)}, \tag{3.3}$$

where $\vec{c}$ is a vector of all metabolite concentrations in mol per lattice site, $\vec{x}$ is the location and $B(\vec{x}, t)$ is the size of the local bacterial population. The size of B can range from $5 \cdot 10^7$ to $2 \cdot 10^{10}$ bacterial cells.

Finally the enzymatic constraint constrains the total flux through the network. It represents the maximum, total amount of flux that can be performed per cell in each population:

$$\sum \vec{f} \leq a. \tag{3.4}$$

The enzymatic constraint a is in mol per time unit per population unit. As both $\vec{f}$ and a are per population unit, this limit scales with population size, so each bacterial cell contributes equally to the metabolic flux possible in a lattice site. The enzymatic constraint is included as a constraint on each FBA solution. Given the constraints, FBA identifies the solution that optimises the objective function, ATP production. The solution consists of a set of input and output exchange fluxes $\vec{F_{in}}(\vec{x},t)$ and $\vec{F_{out}}(\vec{x},t)$, and a growth rate $g(\vec{x},t)$. The exchange fluxes are taken as the derivatives of a set of partial-differential equations to model the exchange of metabolites with the environment. The size of the population increases proportionally to the growth rate in the FBA solution.

To mimic quiescence at high densities, populations above the spreading threshold of $2 \cdot 10^{10}$ bacteria do not perform metabolism. In practice this rarely occurs because we maintain sufficient space for populations to spread into empty lattice sites. In the simulations of Fig. 3.2A (n=120) metabolism was not performed in, on average, 0.05% of all populations in a timestep.

3.4.5 Environmental metabolites

We model 735 different extracellular metabolites. This is the union of all metabolites that can be exchanged with the environment by at least one GEM in the model. In the simulations 39 metabolites are present in the medium in more than micromolar amounts at any point. We combine the L-lactate and D-lactate metabolites for Fig. 3.1B, Video S1 and Video S2. Nearly all lactate in the model is L-lactate.

To represent the mixing of metabolites by colonic contractions we apply a diffusion process to the metabolites at each timestep. Metabolic diffusion is applied in two equal steps to the model. In each step, 14.25% of each metabolite diffuses from each lattice site to each of the four nearest neighbours. This causes a net diffusion each timestep of $6.3 \cdot 10^5$ cm^2/s. Metabolites are also added and removed by bacterial populations as a result of the FBA solutions, yielding

$$\frac{d\vec{c}(\vec{x},t)}{dt} = \vec{F_{out}}(\vec{x},t)B(\vec{x},t) - \vec{F_{in}}(\vec{x},t)B(\vec{x},t) + \frac{D}{L^2}\sum_{\vec{i}\in \mathrm{NB}(\vec{x})}\left(\vec{c}(\vec{i},t) - \vec{c}(\vec{x},t)\right), \quad (3.5)$$

where $\vec{F_{out}}(\vec{x},t)$ is a vector of fluxes from the bacterial populations to the environment, in mol per time unit per population unit, D is the diffusion constant, L is the lattice side length, and $NB(\vec{x})$ are the four nearest neighbours.

All metabolites except oxygen are moved distally by one lattice site every timestep to represent advection. The transit time, including diffusion, is approximately 11 hours, corresponding with *in vivo* observations in newborn infants [141, 142]. Metabolites at the most distal column of the lattice, the end of the colon, are removed from the system at each timestep.

Every 60 timesteps (representing three hours) metabolites representing inflow from the small intestine are inserted into the first six columns of lattice sites. Three hours represents a realistic feeding interval for neonates [175]. Food intake contains 211 µmol of lactose by default, a concentration in line with human milk [66], assuming 98% host uptake of carbohydrates before reaching the colon [131]. In some simulations 211 µmol of additional lactose, GOS, or 2'-FL is added. Because there is very little uptake of prebiotics by the infant [190], the oral intake of prebiotics would be much lower than that of lactose. GOS is inserted as separate fractions of DP3, DP4, or DP5 based on analysis of the composition of Vivinal-GOS [206]. 64% is DP3, 28% is DP4 and 8% is DP5. Water is provided in unlimited quantities. Oxygen is placed during initialisation [125] at 0.1 µmol per lattice site. No other metabolites are available, other than those produced as a result of bacterial metabolism within the model.

3.4.6 Population dynamics

During initialization there is a probability of 0.3 for each lattice site to get a population of size $5 \cdot 10^7$ of a random species (Table 3.1). Taken together, this averages around 540 populations, leading to a total initial bacterial load of $2.7 \cdot 10^{10}$, in line with *in vivo* estimates [5] when we assume a uniform bacterial density and a total colon volume of 90 ml. In each timestep each local population solves the FBA problem based on its own GEM, the enzymatic constraint a, its current population size $B(\vec{x},t)$ and the local concentrations of metabolites $\vec{c}(\vec{x},t)$, and applies the outcome to

the environment (see above) and the growth rate $g(\vec{x},t)$ to its own population size, as follows:

$$\frac{dB(\vec{x},t)}{dt} = B(\vec{x},t)g(\vec{x},t). \tag{3.6}$$

Each step, each population of at least $1 \cdot 10^{10}$ bacteria (Table 3.2) will create a new population if an adjacent empty lattice site is available. Half of the old population size is transferred to the new population, so that the total size is preserved. To mimic colonisation new populations are introduced at random into empty lattice sites during the simulation, representing both dormant bacteria from colonic crypts [143] and small bacterial populations formed from ingested bacteria, which may only become active after being moved far into the gut. Each empty lattice site has a probability of 0.00005 (Table 3.2) each step to acquire a new population of a randomly selected species. All species have an equal probability to be selected. We initialise these populations at the same population size B as the initial populations in the model (Table 3.2). Each population dies out at a probability of 0.0075 per timestep, creating a turnover within the range of estimated microbial turnover rates in the mouse microbiota [72].

To mix the bacterial populations the lattice sites swap population contents each timestep. We use an algorithm inspired by Kawasaki dynamics [176], also used previously for bacterial mixing [103, 122]: In random order, the bacterial content of each site, i.e., the bacterial population represented by its size $B(\vec{x},t)$ and the GEM, are swapped with a site randomly selected from the Moore neighbourhood. This swap only occurs if both the origin and destination site have not already swapped in this timestep. With this mixing method the diffusion constant of the bacterial populations is $6.3 \cdot 10^5 cm^2/s$, equal to that of the metabolites. This represents an improvement over our earlier work [103], where the diffusion of metabolites and populations differed. Bacterial populations at the most distal column, i.e. at the exit of the colon, are removed from the system. To increase the bacterial diffusion rate in the sensitivity analysis this process was executed five times, marking all sites as unswapped after each execution. To decrease the bacterial diffusion rate the number of swaps was limited to a fifth of the usual number of swaps.

3.4.7 Analysis

We record the size, species, location, and important exchange fluxes $\vec{F_{in}}(\vec{x}, t)$ and $\vec{F_{out}}(\vec{x}, t)$ for each population at each timestep. To detect irregularities we also record the net flux of carbon, hydrogen, oxygen, and Gibbs free energy for every population at each timestep. Gibbs free energy is estimated using the Equillibrator database [173]. Energy loss l in joules per timestep per population unit is recorded as follows, where i are metabolites, F is the exchange flux rate in mol per timestep per population unit and E contains the Gibbs free energy in joules per mol for each metabolite,

$$l = \sum_i F(i) \cdot E(i). \tag{3.7}$$

3.4.8 Parameters

Parameters of the system are listed in table 3.2. We estimate a total volume of 90ml for the infant colon [135, 136], which leads to a rough estimate on the order of 10^{12} bacteria in the newborn infant colon given an abundance per ml of around 10^{10} [5]. Values for free parameters were estimated and evaluated in the sensitivity analysis.

3.4.9 Implementation

We implemented the model in C++11. We based the model on our own earlier models of the gut microbiota [122, 103]. Random numbers are generated with Knuth's subtractive random number generator algorithm. Diffusion of metabolites was implemented using the Forward Euler method. The GEMs are loaded using libSBML 5.18.0 for C++. We used the GNU Linear Programming Kit 4.65 (GLPK) as a linear programming tool to solve each FBA with enzymatic constraint. We used the May 2019 update of AGORA, the latest at time of writing, from the Virtual Metabolic Human Project website (vmh.life). We used Python 3.6 to extract thermodynamic data from the eQuilibrator API (December 2018 update) [173]. When not noted otherwise p-values were calculated with R 4.2.1 using the Mann-Whitney test from the 'stats' package 3.6.2. Model screenshots were made using the libpng16 and pngwriter libraries. Other visualisations were performed with R 4.2.1 and Google Sheets. Raincloud visualisations used a modified version of the Raincloud plots library for R [216].

3.5 Supplemental material

Available from `cell.com/iscience/fulltext/S2589-0042(24)00306-7`

S1 Table

Table of changed or deleted reactions and annotations.csv
A table of changes made to the AGORA models as a .csv file.

S1 Video

Video of a simulation with no prebiotics, consisting of a visualisation of the distribution of bacterial species and major metabolites. Lines represent, from top to bottom: Bacteria, lactose, 2'-FL, lactate (Both L and D), acetate, 1,2-PD, butyrate, succinate, CO2, H2, propionate

S2 Video

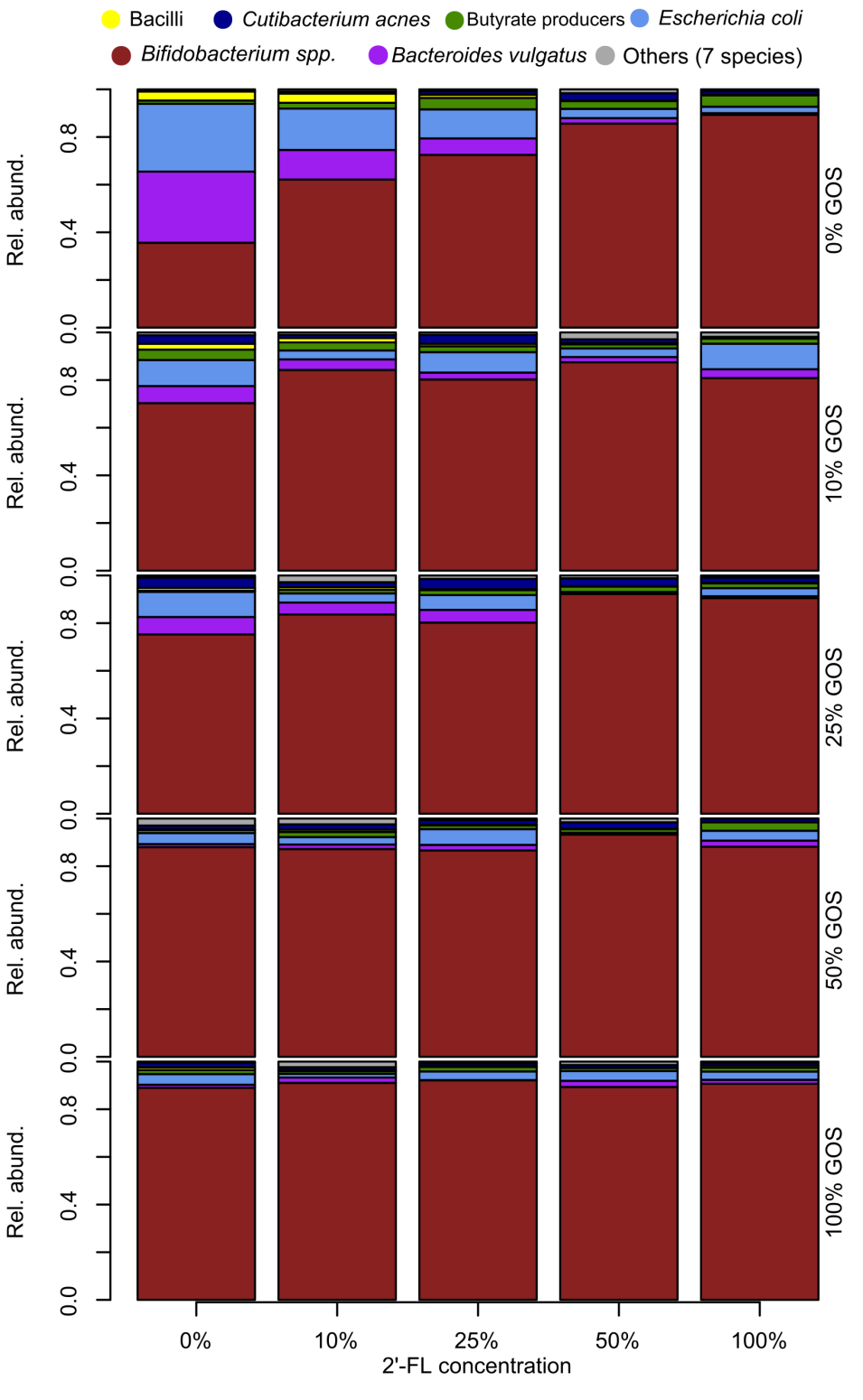
Video of a simulation with 2'-FL, consisting of a visualisation of the distribution of bacterial species and major metabolites. Lines represent, from top to bottom: Bacteria, lactose, 2'-FL, lactate (Both L and D), acetate, 1,2-PD, butyrate, succinate, CO2, H2, propionate

S3 Video

Video of a simulation without prebiotics, displaying fluxes between population and metabolite pools Line width is scaled with the flux per metabolite over 60 timesteps per frame, multiplied by the carbon content of the molecule, with a minimum threshold of 100 µmol atomic carbon.

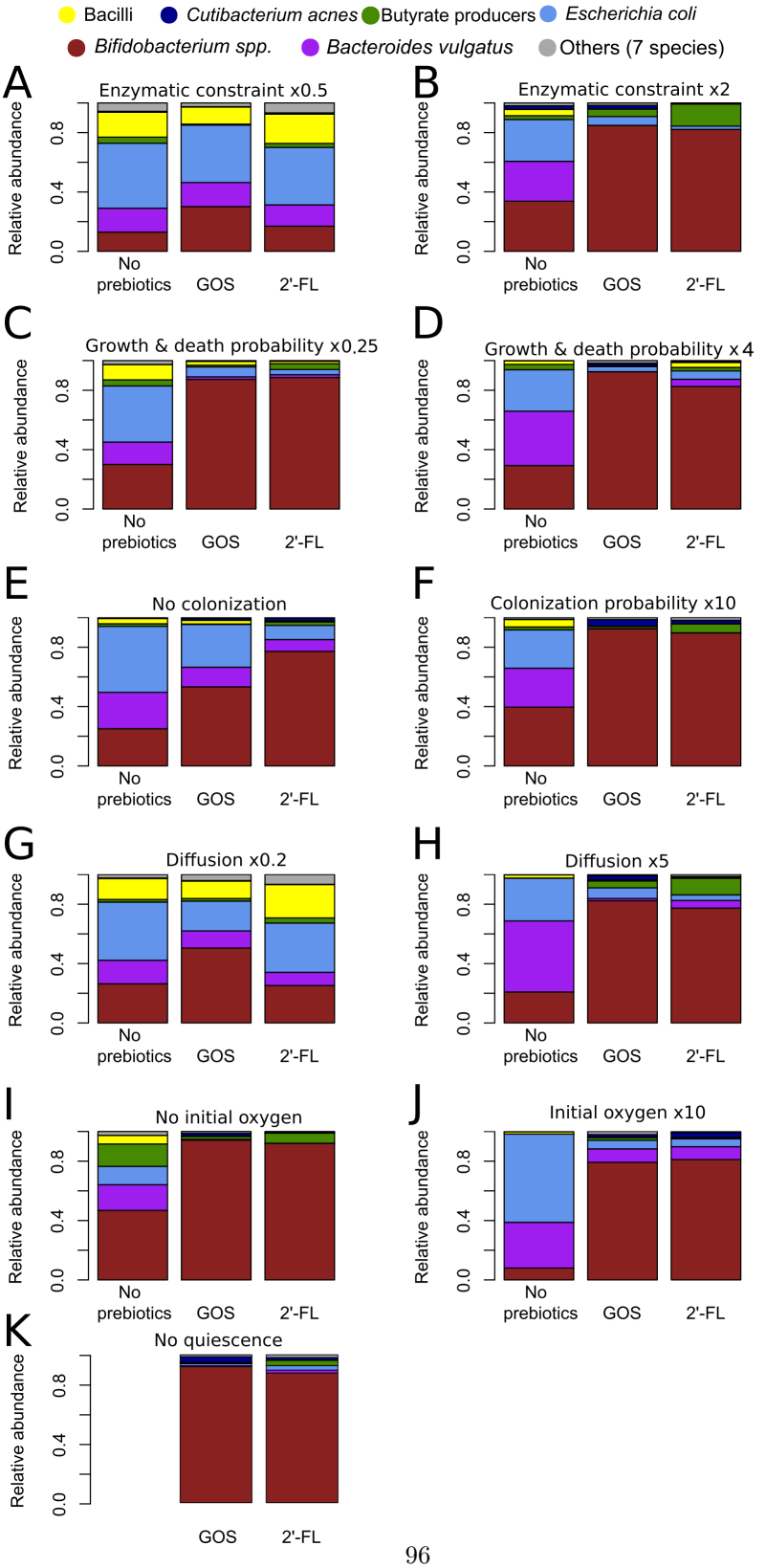
S4 Video

Video of a simulation with 2'-FL, displaying fluxes between population and metabolite pools. Line width is scaled with the flux per metabolite over the 60 timesteps per frame, multiplied by the carbon content of the molecule, with a minimum threshold of 100 µmol atomic carbon.


Bacilli
Cutibacterium acnes
Butyrate producers
Escherichia coli
Bifidobacterium spp.
Bacteroides vulgatus
Others (7 species)
Rel. abund.
0.0
0.4
0.8
0% GOS
10% GOS
25% GOS
50% GOS
100% GOS
0%
10%
25%
50%
100%
2'-FL concentration

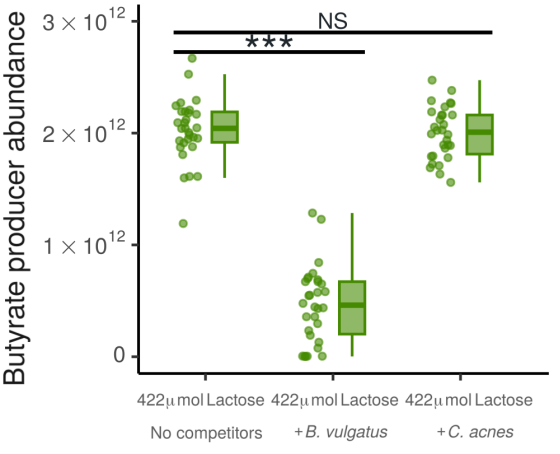
S1 Figure

Relative abundance of bacterial species at the end of 21 days with varying inputs of 2'-FL and GOS compared to the fixed input amount of lactose. n=30 for each condition, each simulation is weighed equally.


Bacilli
Cutibacterium acnes
Butyrate producers
Escherichia coli
Bifidobacterium spp.
Bacteroides vulgatus
Others (7 species)
A
Enzymatic constraint x0.5
B
Enzymatic constraint x2
C
Growth & death probability x0.25
D
Growth & death probability x 4
E
No colonization
F
Colonization probability x10
G
Diffusion x0.2
H
Diffusion x5
I
No initial oxygen
J
Initial oxygen x10
K
No quiescence
Relative abundance
0.0
0.4
0.8
No prebiotics
GOS
2'-FL

S2 Figure

(A to K) Relative abundance of bacterial species in the conditions with no prebiotics, with GOS, or with 2'-FL at the end of 21 days, with the following alteration from the baseline of Fig. 3.2A: (A) Enzymatic constraint loosened by a factor of 2, to 4 µmol flux per timestep per $1 \cdot 10^{10}$ population (B) Enzymatic constrained tightened by a factor of 2, to 1 µmol flux per timestep per $1 \cdot 10^{10}$ population (C) Growth decreased by a factor of 10, by increasing the ATP per bacteria to $1 \cdot 10^{-14}$, with the death probability decreased to 0.00075 per population per timestep. (D) Growth increased by a factor of 10 by decreasing the ATP per bacteria to $1 \cdot 10^{-16}$, with the death probability increased to 0.075 per population per timestep (E) Colonisation removed by setting the probability for new populations to be placed after initialization to 0 (F) Colonisation increased by x10 by setting the probability per empty lattice to acquire a new population to 0.0005 per timestep (G) Diffusion of both metabolites and bacteria decreased by a factor of 5 to $1.26 \cdot 10^{-6}$ cm^2/s (H) Diffusion of both metabolites and bacteria increased by a factor of 5 to $3.15 \cdot 10^{-5}$ cm^2/s (I) No initial presence of oxygen (J) Initial oxygen increased to 1 µmol per lattice site (K) Quiescence disabled
For each figure: n=30 for each condition, each simulation is weighed equally.



S3 Figure

Abundance of butyrate producing bacteria at the end of 21 days with 422 µmol of lactose per three hours and without prebiotics, either without competitors (only *Bifidobacterium* and butyrate producing bacteria), with addition of *B. vulgatus*, or with addition of *C. acnes*. n=30 for each condition. Each simulation is represented by one dot.
NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$

3.6 Contributions

J.M.W.G., and R.M.H.M acquired funding. D.M.V., J.M.W.G., and R.M.H.M. conceived and planned the simulations. D.M.V. wrote software used for the simulations. D.M.V. performed the simulations and analyzed the data. R.S, E.L., J.M.W.G., and R.M.H.M contributed to the interpretation of the results. J.M.W.G., and R.M.H.M. supervised the project. D.M.V. drafted the manuscript. D.M.V., R.S., E.L., J.M.W.G. and R.M.H.M. revised and edited the manuscript.

3.7 Acknowledgments

This study was financially supported by FrieslandCampina. R.S, E.L., and J.M.W.G. are currently employed by FrieslandCampina. This work was performed using the ALICE compute resources provided by Leiden University.

Chapter 4

4

Human milk oligosaccharide decreases microbial mucin consumption by stimulating bacteria that do not share extracellular resources in infant gut simulations

Authors

David M. Versluis[a], Clair Wijtkamp[b], Ellen Looijesteijn[c], Jan M. W. Geurts[c], Roeland M. H. Merks[a,b]

Affiliations

[a] Leiden University, Institute of Biology, Leiden, The Netherlands

[b] Leiden University, Mathematical Institute, Leiden, The Netherlands

[c] FrieslandCampina, Amersfoort, the Netherlands

In preparation

Abstract

Intestinal mucin acts as a barrier protecting the infant gut wall against diseases such as colitis and rotavirus. *In vitro* experiments have shown that the gut microbiota of breastfed infants consumes less mucin than the microbiota of non-breastfed infants, but the mechanisms are incompletely understood. The main difference between human milk and most infant formulas is the presence or absence of human milk oligosaccharides (HMOs). We therefore hypothesize that HMOs protect the mucin layer by favoring non-mucin consuming bacteria in the microbiota. To understand the underlying mechanisms we analysed the potential effect of HMOs on the developing infant gut microbiota using a computational modeling approach that describes the metabolism and ecology of the infant gut microbiota. We show that in the presence of the HMO 2'-fucosyllactose (2'-FL), extracellular digestion of 2'-FL by the mucin-consumer *Bifidobacterium bifidum* makes this species vulnerable to competitors, including *Bacteroides vulgatus*. Our model predicts that the non-mucin consuming *Bifidobacterium longum*, which digests HMO intracellularly, can then become dominant. Paradoxically, in monocultures *B. longum* grows less efficiently on HMOs than *B. bifidum*. In mixed cultures, extracellular HMO degradation products resulting from *B. bifidum* activity can serve as substrate for *B. bifidum*, or for other bacteria instead. These digestion products become 'public goods'. We conclude that the theory of public goods in microbial ecology may be key to understanding the effects of HMOs on the mucin-consuming potential of the developing infant gut microbiota.

4.1 Introduction

The interior of the gastrointestinal tract is covered by a layer of mucin which protects the gut against diseases such as colitis [217] and rotavirus infection [218]. The mucin layer consists largely of mucin glycoproteins. Some species of intestinal bacteria, such as *Bacteroides* spp. and *Bifidobacterium bifidum*, can consume these mucins [63, 87]. As mucin protects the gut against infection, consumption of intestinal mucins by the residing microbiota potentially increases the risk of disease [28]. It is thought that breastfeeding shapes the infant gut microbiota in such a way that the microbiota consumes less mucin [219]. This hypothesis is based on *in vitro* observations showing that microbiota from the feces of breastfed infants consume mucins more slowly than the microbiota from the feces of non-breastfed infants [219].

As the infant formula at the time of this study did not contain any human milk oligosaccharides (HMO), we hypothesize that it is the HMO in human milk that shaped the microbiota of the breastfed infants to consume less mucin. HMOs are the most abundant component of human milk after lactose and lipids [220]. They are not digested by the host, but exclusively by intestinal bacteria, shaping the microbiota [169]. About 200 HMO structures exist, each of which consists of a core of lactose with other sugars attached to it [106]. We will focus in particular on the potential effects of the HMO 2'-fucosyllactose (2'-FL), as it is the most abundant HMO in the milk of most humans [221] and its digestion by the infant gut microbiota is well-characterized [222]. We also examine galacto-oligosaccharides (GOS), which have a similar structure and are consumed by the same bacteria, but are not HMOs [223]. GOS are frequently added to commercial infant formula [185]. HMOs may decrease mucin consumption by stimulating non-mucin consumers at the expense of mucin consumers. Indeed, the non-mucin consumer *Bifidobacterium longum* specializes in HMOs, including 2'-FL [65], and a high abundance of *B. longum* in the infant gut microbiota lowers total mucin consumption in breastfed infants [30]. Paradoxically, in *in vitro* experiments, the mucin consumer *Bifidobacterium bifidum* digests and takes up HMOs such as 2'-FL more efficiently than *B. longum* [224, 222], but *B. bifidum* is much less abundant than *B. longum* in the infant gut [3, 225]. We hypothesize that differences in HMO digestion strategies cause *B. longum* to be more abundant than *B. bifidum in vivo*.

The carbohydrate metabolism of *B. bifidum* and *B. longum* is similar, as both species use bifid shunt metabolism to metabolise sugars [90], but the species cru-

cially differ in how they break down and transport HMOs into the cell [222]. *B. longum* imports most HMOs, including 2'-FL, into the cell using ATP-dependent active transport [169]. In contrast, *B. bifidum* digests most HMOs, including 2'-FL, outside the cell, secreting enzymes that break down HMOs into smaller sugars such as lactose [220]. The import of lactose produced from 2'-FL does not require active transport [226]. Thus, *B. bifidum*'s metabolism of 2'-FL is thought to be more efficient. However, a major disadvantage of extracellular digestion is that other species can take up the digestion products [83]. Thus, we hypothesize that cross-feeding by other species on the digestion products of *B. bifidum* explains the relatively low abundance of *B. bifidum in vivo*.

Extracellular products that can benefit other species instead of the producer are a type of public goods [227]. The role of public goods in inter-species interaction has been studied in detail in many systems, in particular in budding yeast *Saccharomyces cerevisiae* [88]. Yeast can secrete the enzyme invertase that breaks down sucrose into fructose and glucose. With this enzyme it can grow in glucose-limited medium rich in sucrose. However, most of the glucose is lost to the cell that expresses the invertase: only around 1% of glucose is taken up by the cell, the rest is lost to the environment. This way, glucose functions as a public good. Because the production of invertase is costly, it can be advantageous for a yeast cell to not produce invertase, but only take up the glucose produced by other cells. In a yeast community invertase producers and invertase non-producers can co-exist [88]. This co-existence is possible because they are playing a 'snowdrift game', in which the less common strategy (producing or not producing invertase) has an advantage over the more common strategy [88]. The abundance of producers and non-producers can be influenced by both the mechanics of public good production as well as pressure from the surrounding bacterial community [228].

The extracellular digestion of oligosaccharides by *B. bifidum* has also been considered as a form of public goods production [83], and may be influenced by the same mechanisms. In our model of the infant microbiota, *B. bifidum* is a producer of public goods from 2'-FL and every other species is a non-producer. Strong pressure from cross-feeding or 'stealing' by the surrounding bacterial population may then explain why *B. bifidum* is less abundant than other bacteria. The advantage of the less common strategy (producing) over the more common strategy (non-producing) may explain the continued low presence of *B. bifidum* in the infant gut.

We investigated (1) whether 2'-FL in milk can explain the reduced mucin con-

sumption observed in the microbiota of breastfed infants, and (2) whether public goods metabolism of 2'-FL can explain the relatively low abundance of the mucin consumer *B. bifidum* compared to *B. longum* in the infant gut microbiota. As these mechanisms involve an interplay between the molecular, population and physical levels, we used a multiscale mathematical model. Such models have previously been used to model spatial bacterial interactions and the gut microbiota [132, 131, 122, 123]. In particular, a previous computational model found that the inclusion of consumable mucin in the model could be an important factor in explaining the diversity of species found in the adult gut microbiota [123]. We based the model on our previous microbiota models [122, 103, 229], which modelled microbial interactions in the adult and infant gut microbiota, including relevant host factors such as flow rate and the presence of initial oxygen. These models have previously been used to generate predictions on the effects of factors such as diarrhoea, oxygenation, and HMO supplementation on microbiota composition and metabolism [122, 103, 229]. We have extended these models to also include species-specific public goods-producing and non-public goods-producing metabolism of mucin and 2'-FL.

Briefly, we show that our model can reproduce cross-feeding between *B. bifidum* and *Anaerobutyricum hallii* observed *in vitro* and that it predicts:
(1) 2'-FL reduces mucin consumption by causing a higher abundance of *B. longum*.
(2) The low abundance of *B. bifidum* can be explained by other species consuming the public goods it produces from 2'-FL.
(3) *B. longum* is more abundant than mucin consumers because its 2'-FL metabolism is intracellular, and so does not produce public goods.

4.2 Results

4.2.1 Public goods model reproduces *in vitro* cross-feeding between mucin consuming and non-consuming species

We first tested if our model system could reproduce consumption of mucin by a cross-feeding community of two species. *In vitro*, *A. hallii* cannot grow on mucin in single culture, whereas it can grow on mucin in co-culture with *B. bifidum* [27]. *B. bifidum* digests mucin extracellularly, which leads to cross-feeding metabolites becoming available in the medium as public goods[27]. *A. hallii* can then grow on

these cross-feeding metabolites [27]. We first investigated whether our computer model could reproduce this cross-feeding.

We simulated this two-species system using a multiscale model [103, 229]. Briefly, the model integrates simulations of the metabolism of bacterial species to generate predictions for microbial abundances and metabolic interactions over space and time. Bacterial metabolism is predicted using flux balance analysis (FBA) on genome-scale metabolic models (GEMs). Growth is determined by a biomass reaction, which FBA maximizes. The default biomass reaction in the model is the ATP production rate. We have extended this model with species-specific mucin metabolism and extracellular public goods production from mucin and 2'-FL. The extracellular public goods production is separate from the FBA predictions of metabolism (Fig. 4.1A&B). We distributed all nutrients and metabolites equally across all lattice sites at the start of each timestep. We also supplied the system only with water and mucin, and only used GEMs of *B. bifidum* and *A. hallii*. We used a fucosylated core-2 mucin structure to represent all mucin oligosaccharides. For details see the methods sections 'Non-spatial model' and 'Nutrient input'.

We compared two model variants. In the first model variant, called the 'control model', we assumed that the FBA solution selects what digestion products of mucin to create, and these were first available in the local lattice site, such that species that produced the digestion enzymes had exclusive benefits from the digestion products in the timestep when they were created. In a second variant, which we called the 'public goods model', all possible mucin digestion products were produced outside of the FBA, and these diffused before they became available for uptake by the populations. This potentially made these products available to populations at other sites. In simulations of the control model, *A. hallii* could not grow, even when*B. bifidum* produced cross-feeding products (Fig. 4.1C). However, in the 'public goods' model variant, the model predicted that *B. bifidum* also released galactose into the medium. *A. hallii* could then grow in co-culture with *B. bifidum* by consuming the galactose (Fig. 4.1D), in agreement with experimental observations [27].

Summarizing, the multiscale model correctly reproduced the cross-feeding of *A. hallii* and *B. bifidum*. Therefore the model with public goods metabolism of mucin provides a good basis for simulating more complex situations, as we will show in the next sections.

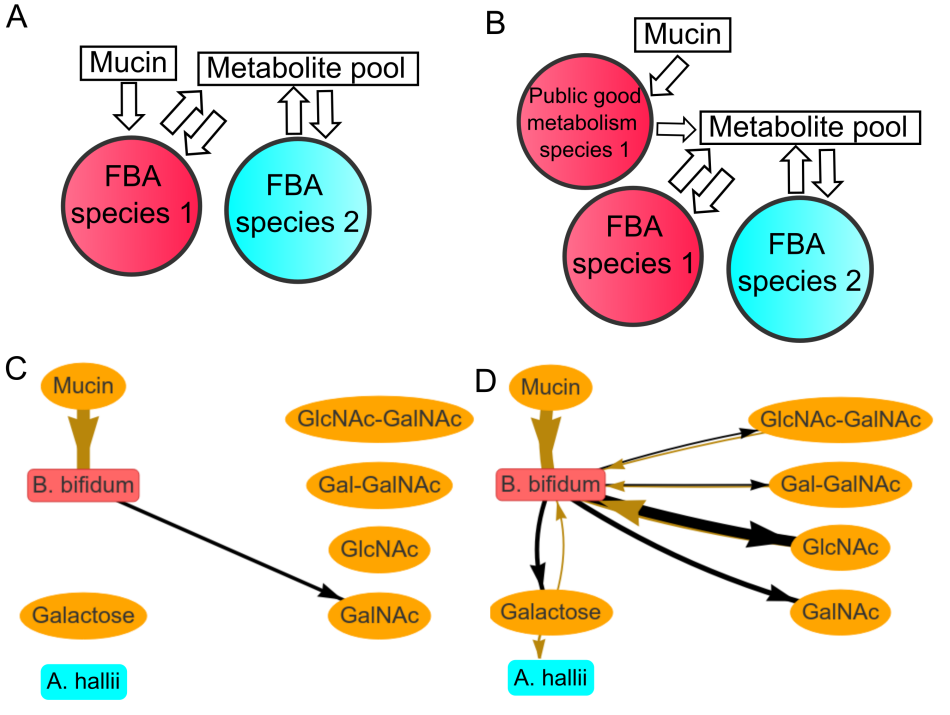


Figure 4.1: Model outline and effect of the public goods model
(A) Schematic of the non-public goods model without spatial separation (see methods section 'Non-spatial model') fed with mucin. Mucin is consumed by specific species, and all bacterial populations share a common nutrient pool.
(B) Schematic of the public goods model without spatial separation fed with mucin. Mucin is consumed separately from other substrates to represent extracellular metabolism, and all bacteria need to retrieve the metabolites of mucin metabolism from the common pool.
(C,D) Metabolic network for *B. bifidum* and *A. hallii* in a single simulation with no spatial separation and mucin as the only nutrient source not using the public goods model (C) and using the public goods model (D). Yellow arrows indicate uptake, black arrows indicate production. Line width is scaled with the flux per metabolite, multiplied by the carbon content of the molecule, with a minimum threshold of 100 µmol atomic carbon.

4.2.2 Public goods model reproduces *in vivo* pattern of microbiota composition and mucin consumption

We next turned our attention to mucin consumption in the full infant microbiota. Two observations suggest that breastfeeding shapes the infant microbiota such that mucin consumption is reduced. Firstly, *in vitro* data suggest that the microbiota from breastfed infants consume less mucin than the microbiota from non-breastfed infants [219]. Secondly, *in vivo*, a breastfed infant microbiota dominated by *B. longum* consumes less mucin [30].

We tested if we could reproduce and analyse these observations in a model of a complex microbiota, as described by the full infant gut microbiota model as introduced previously [103, 229], which we have extended with mucin production and consumption, and extracellular digestion of the HMO 2'-FL (see methods section 'changes to GEMs'). Briefly, the spatial model represents the infant colon with a regular square lattice of 225×8 boxes of 2×2 mm, (Fig. 4.2 A&B). Nutrients are introduced into the system at regular intervals, and populations, nutrients, and metabolites diffuse through the system. Nutrients and metabolites advect through the system, and are removed at the distal end. In this extended version of the model, mucin is produced at one of the boundaries, representing the intestinal wall. The other boundaries represent the center, and the entrance and exit of the intestine.

The model contains 21 bacterial species, each represented by a genome-scale metabolic model (GEM) (Table 4.1). The species were selected based on *in vivo* data [3], as described in methods section 'Species composition'. Of these species, only *B. bifidum* and *B. longum* could digest 2'-FL in the model. Again we simulated a 'public goods' model and non-public goods 'control' model. *B. bifidum* produced public goods from 2'-FL only in the public goods model, but *B. longum* always digested 2'-FL without producing public goods, even in the public goods model of 2'-FL metabolism. Several species could digest mucin, all of which produced public goods.

We simulated the model with four conditions, differing in nutrient input and handling of public goods 2'-FL metabolism: (1) a nutrient input of 211 µmol of lactose per 60 timesteps (2) a nutrient input of 422 µmol of lactose per 60 timesteps (3) a nutrient input of 211 µmol of lactose and 211 µmol of 2'-FL per 60 timesteps, without public goods 2'-FL metabolism, and (4) a nutrient input 211 µmol of lactose and 211 µmol of 2'-FL per 60 timesteps, with public goods 2'-FL metabolism. We did this to examine which approach better reproduced *in vivo* observations. 211 µmol

of lactose per 60 timesteps is our best estimate for lactose input into the infant gut, but we also included an increased amount of lactose in condition 2 to control for the increased amount of sugars present in the system when 2'-FL is included. This allowed us to distinguish between an effect of 2'-FL and an effect of total sugar input.

In conditions 1 and 2, i.e. without 2'-FL, the simulations predicted a high abundance of all *Bifidobacterium* species, *Escherichia coli*, and *Bacteroides vulgatus*, (Fig. 4.2C) and a high mucin consumption (Fig. 4.2D). The model predictions without 2'-FL matched well with the outcomes commonly observed in non-breastfed infants, who generally did not consume HMO. As observed in the formula-fed infant gut microbiota, the model predicts a diverse microbiota, with *Bifidobacterium*, *Bacteroides* and *Escherichia* as major groups [53, 3, 31] and high mucin consumption [219].

In condition 3, with 2'-FL and without public goods 2'-FL metabolism, the model predicted that *B. longum* and *B. bifidum* became highly abundant, leading to a high mucin consumption (Fig. 4.2C & D). In condition 4, with 2'-FL and public goods 2'-FL metabolism, we observed a dominance of *B. longum* and lower mucin con-

Table 4.1: Species and subspecies included in the computational model. Color indicates color used in figures.

Name	Phylum	Mucin consumption	2'-FL consumption
Bifidobacterium longum ssp. *infantis*	Actinomycetota	No	Without public goods
Bifidobacterium breve	Actinomycetota	No	No
Bifidobacterium bifidum	Actinomycetota	With public goods	With public goods
Collinsella aerofaciens	Actinomycetota	No	No
Cutibacterium acnes	Actinomycetota	No	No
Rothia mucilaginosa	Actinomycetota	No	No
Eggerthella sp. *YY7918*	Actinomycetota	No	No
Streptococcus oralis	Bacillota	No	No
Staphylococcus epidermidis	Bacillota	No	No
Gemella morbillorum	Bacillota	No	No
Enterococcus faecalis	Bacillota	No	No
Lactobacillus gasseri	Bacillota	No	No
Ruminococcus gnavus	Bacillota	No	No
Veillonella dispar	Bacillota	No	No
Anaerobutyricum hallii	Bacillota	No	No
Roseburia inulinivorans	Bacillota	No	No
Clostridium butyricum	Bacillota	No	No
Parabacteroides distasonis	Bacteroidota	With public goods	No
Bacteroides vulgatus	Bacteroidota	With public goods	No
Haemophilus parainfluenzae	Pseudomonadota	No	No
Escherichia coli SE11	Pseudomonadota	No	No

sumption (Fig. 4.2C & D). Mucin consumption was not lower in all simulations, as the distribution became bimodal (Fig. 4.2D). Moreover, the public goods model outcomes matched the *in vivo* data on breastfed infants better than the non-public goods model, as *B. longum* is generally more abundant than *B. bifidum* in the infant gut microbiota [3, 225], and *in vitro* mucin consumption by the microbiota of breastfed infants is generally lower [219]. A high *B. longum* abundance is also associated with lower mucin consumption in breastfed infants [30].

We concluded that predictions of the public goods model of 2'-FL metabolism matched observations of the inhibitory effect of breastfeeding and *B. longum* abundance on mucin consumption better than the non-public goods model. We proceeded to focus on the public goods model of 2'-FL metabolism.

4.2.3 Model outcomes robust for different oligosaccharides but not for all parameters or growth assumptions

We next analysed the robustness of our results by performing a parameter sensitivity analysis. We first asked if and how the rate of public goods production could affect the model predictions. We repeated the simulations with 2'-FL (condition 4) for public goods production that was ten times slower, or ten times faster (see methods section 'Extracellular metabolism'). *B. bifidum* was much less abundant than *B. longum*, as it was in the original predictions, regardless of the parameter setting (Fig. S1A). The bimodal distribution of mucin consumption was also similar to the original predictions (Fig. S1B).

We next performed a sensitivity analysis for 6 parameters of the model without 2'-FL (condition 1), and for the model with 2'-FL and public goods 2'-FL metabolism (condition 4). This revealed that the amount of oxygen and mucin in the model were particularly important for the scenario without 2'-FL, but not with 2'-FL (Fig. S2).

We then repeated the simulations of conditions 3 and 4 with galacto-oligosaccharides (GOS) instead of 2'-FL. GOS are oligosaccharides often present in infant formula, and, like 2'-FL, are primarily digested by *Bifidobacterium* spp. [64]. However, GOS are digested differently from 2'-FL. GOS consist of a mixture of oligosaccharides with different chain lengths. All *Bifidobacterium* species in the model break down chains of four or more sugars extracellularly [206, 230, 86, 231]. Chains of three sugars are imported using active transport by *B. longum* and *Bifidobacterium breve*, but broken down extracellularly by *B. bifidum*. A full description of the digestion of GOS in the

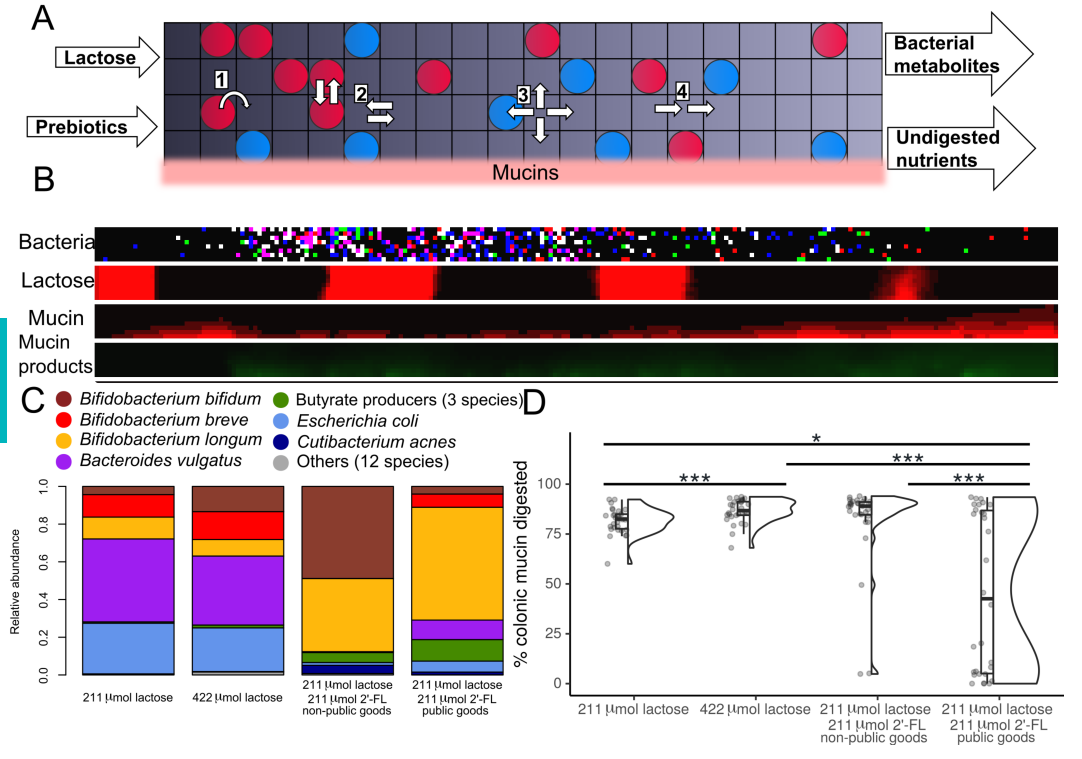


Figure 4.2: 2'-FL metabolism stimulates a microbiota that consumes less mucin and has dominant *B. longum* exclusively in the public goods model
(A) Schematic of the model. Circles represent populations, color represents species. Lactose and 2'-FL enter left (proximally). Mucin enters at the bottom. All nutrients and metabolites move distally (right) and leave the system there. Numbers indicate processes in the model: (1) Metabolism, growth, and division of populations (2) Diffusion of populations (3) Diffusion of metabolites (4) Advection of metabolites
(B) Screenshot of a model simulation, showing from top to bottom the bacterial layer, lactose, mucin, and mucin products. Color indicates species in the bacterial layer. Brightness indicates growth in the bacterial layer, and concentration in the other layers.
(C) Average relative abundance of bacterial species at the end of 21 days with four different conditions: (1) with 211 µmol lactose, (2) 422 µmol lactose, (3) 211 µmol lactose and 211 µmol 2'-FL without public goods 2'-FL metabolism, or 211 µmol lactose and (4) 211 µmol 2'-FL with public goods 2'-FL metabolism. n=30 for each condition.
(D) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final three hours of the model, per condition of C. n=30 for each condition
NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$

model is in methods section 'Changes to GEMs'. Compared to the simulations with 2'-FL, the simulations with GOS had a higher *B. bifidum* and *B. breve* abundance ($p<0.001$) and a lower abundance of butyrate producing bacteria ($p<0.001$), but a similar distribution of mucin consumption (Fig. S3). Mucin consumption with GOS was lower with public goods metabolism than without ($p=0.025$, Fig. S3), as it also was with 2'-FL.

For all simulations so far, we considered a fucosylated core-2 mucin structure, but, due to genetic variation, some humans have intestinal mucins that are not fucosylated [232, 73]. To examine the consistency of our results for the non-fucosylated core-2 mucin we repeated the simulations of all four conditions shown in Fig. 4.2C and Fig. S4 for non-fucosylated core-2 mucin. For details see methods section 'Changes to GEMs'. The alternative mucin structure led to similar results as with the default mucin structure, but significantly more *E. coli* in the conditions with lactose ($p<0.01$) and GOS ($p=0.04$) (Fig. S4). This may be because *E. coli* in the model can completely degrade the non-fucosylated mucin, but not the default mucin (see methods section 'Changes to GEMs' and table S2). The abundance of *E. coli* is still within the ranges typically observed *in vivo* [3, 31].

To assess the effects of the simplifications made in the biomass reactions (ATP production rate biomass), we next repeated conditions 1 and 3 from Fig. 4.2C with biomass reactions different from the assumption that ATP production rate determines biomass production, which we used in simulations so far. To test to what extent the simulation results depend on the exact biomass reactions, here we compare the earlier obtained results with results produced with four alternative biomass reactions: (1) ATP production rate and acetate production rate, (2) only acetate production rate, (3) ATP production rate and pyruvate production rate, or (4) ATP production rate and ketoglutarate production rate. Acetate, pyruvate, and ketoglutarate are all used as building blocks for bacterial biomass in *E. coli in vitro* [233]. Alternative biomass reactions 1,3, and 4 required respectively 1, 3, and 5 mol of ATP to be produced per mol of carbohydrate. This brings the ratio between carbon atoms and ATP in line with those in the default biomass reactions of the GEMs [101], which we used as the basis for our metabolic models. The carbon used in these GEMs is used for amino acids, cell wall and cell membrane components, and regeneration of other components. Alternative biomass reaction 1, ATP production rate and acetate production rate, led to outcomes similar to those with the default biomass reaction (Fig. S5 A&E). However, the populations of *Lactobacillus gasseri* were unable to

grow on lactose with this alternative biomass condition, contrary to *in vitro* data [234]. The other three alternative reactions led to model predictions that did not match *in vivo* data (Fig. S5 B-D). With alternative 2 the model always predicted a microbiota consisting nearly entirely of *B. longum*, which is not typically observed *in vivo* [3, 31] (Fig. S5 B), while with alternative 3 and 4 the model predicted the extinction of *B. vulgatus* in all simulations, which also does not match with *in vivo* observations [3, 31] (Fig. S5 C&D).

We concluded that the model was robust for an alternative prebiotic oligosaccharide, a non-fucosylated mucin, changes to the public goods production rate, and several other parameter changes. However, the assumed biomass reaction greatly impacted the predictions of the model. We decided not to continue with alternative biomass reaction 2-4, as these led to unrealistic outcomes. Alternative biomass reaction 1 led to predictions that were no worse than those with the default biomass reaction, but also did not match the *in vivo* data better. A better match could have been achieved by, for example, predicting a higher abundance of Firmicutes as is sometimes observed *in vivo* [16]. Alternative biomass reaction 1 also led to *L. gasseri* no longer growing on lactose, and possibly further inaccuracies, as requiring acetate will block growth in some circumstances. We proceeded to use the biomass function that required only ATP, as the resulting microbiota composition adequately matched *in vivo* observations of microbiota composition while allowing for growth of all GEMs [3, 31]. We discuss potential future improvements to the biomass reaction in the discussion.

4.2.4 *B. bifidum* produces cross-feeding metabolites that may allow it to be exploited by other bacteria

As we observed a bimodal distribution of mucin consumption in the model simulations with 2'-FL and public goods 2'-FL metabolism (Fig. 4.2B), we hypothesized that both a different metabolism and a different bacterial composition could be associated with the high and low mucin consumption simulation results. We next investigated why the model with public goods metabolism of 2'-FL led to (1) a lower mucin consumption in some simulations, but not in others, (2) a high *B. longum* abundance, and (3) a low *B. bifidum* abundance.

To gain insight into the potential mechanisms driving low versus high mucin consumption, we visualized the metabolic interactions in the final 60 timesteps of a

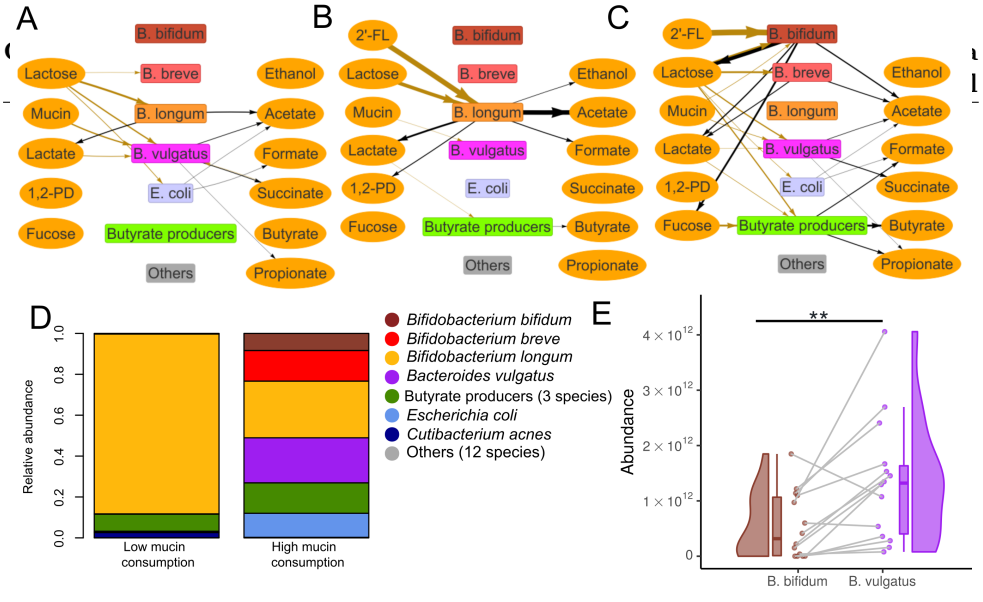


Figure 4.3: High consumption simulations show complex cross-feeding but higher abundance of *B. vulgatus* than *B. bifidum*
(A,B,C) Visualisation of metabolic interactions in timesteps 10020 to 10080 (last 3 hours) of a sample simulation (A) with 211 µmol lactose per 60 timesteps (B) with additional 211 µmol 2'-FL per 60 timesteps, public goods 2'-FL metabolism, and low mucin consumption (C) with additional 211 µmol 2'-FL per 60 timesteps, public goods metabolism for 2'-FL, and high mucin consumption. Line width is scaled with the flux per metabolite over the last 60 timesteps, multiplied by the carbon content of the molecule, with a minimum threshold of 100 µmol atomic carbon.
(D) Average relative abundance of bacterial species in the condition with public goods 2'-FL metabolism, split by low (<50%) or high (>50%) mucin consumption at the end of 21 days. n=16&n=14, respectively.
(E) Absolute abundance of *B. bifidum* and *B. vulgatus* in the simulations with public goods 2'-FL metabolism and high mucin consumption. Data from the same simulation is connected with a line. n=14.
NS: Not significant, *: p<0.05, **:p<0.01, ***:p<0.001

simulation without 2'-FL (Fig. 4.3A), of a simulation with public goods metabolism of 2'-FL and low mucin consumption (Fig. 4.3B) and of a simulation with public goods metabolism of 2'-FL and high mucin consumption (Fig. 4.3C).

The network visualisations revealed notable differences between these three situations. In the control simulation without 2'-FL a complex community formed, with high mucin consumption and extensive cross-feeding between *B. longum* and *B. vulgatus* (Fig. 4.3A). The simulation with 2'-FL and low mucin consumption showed less diversity: it became dominated by 2'-FL consuming *B. longum* while non-mucin consumers were predicted to cross-feed on lactate (Fig. 4.3B.) The high mucin consumption simulation with 2'-FL had a distinct complex community, in which *B.*

bifidum consumed 2'-FL, and many species cross-fed on the lactose produced as a public good by *B. bifidum* (Fig. 4.3C). *B. longum* was only present at a very low abundance in this simulation. From these networks we concluded that different communities arose in the simulations with high mucin consumption compared to the simulations with low mucin consumption.

To investigate the bacterial species associated with these communities in the model we categorized the relative abundances of the simulations according to the rate of mucin consumption (Fig. 4.3D). Low mucin consuming simulations (<50% of mucin consumed, n=16) were consistently dominated by *B. longum*. High mucin consuming simulations (>50% of mucin digested, n=14) instead showed a higher abundance of *B. bifidum*, but also more *E. coli* and *B. vulgatus*. Interestingly, *B. bifidum* was typically not the dominant species in either set of simulations. We hypothesized that other species were taking advantage of the public good 2'-FL metabolism by *B. bifidum*, and proceeded to analyze the relationship between *B. bifidum* and the other species. In support of our hypothesis, *B. vulgatus* was more abundant than *B. bifidum* (paired samples Wilcoxon test, p=0.02) and also positively correlated (Spearman correlation,p<0.01,r=0.68) (Fig. 4.3E). This implies that *B. vulgatus* benefited from a higher abundance of *B. bifidum*. *B. vulgatus* was the only species within the high mucin consumption simulations to be both more abundant and positively correlated (Fig. S6).

Altogether, we observed in our simulations that *B. bifidum* produced public goods that were consumed by other species, but was consistently less abundant than *B. vulgatus*, leading us to hypothesize that *B. vulgatus* exploited *B. bifidum*'s metabolism of 2'-FL.

4.2.5 *B. bifidum* is exploited when digesting 2'-FL but *B. longum* is not

We continued to examine the possible interactions between *B. vulgatus* and *B. bifidum*. In particular, we hypothesized that *B. vulgatus* could exploit *B. bifidum* only in the public good model of 2'-FL metabolism, not without 2'-FL, nor with 2'-FL modelled without public goods metabolism.

To further investigate possible interactions between *B. vulgatus* and *B. bifidum* in the model we studied a simplified model with only *B. vulgatus* and *B. bifidum* and no other species. We then performed new simulations with the four conditions of

Fig. 4.2.

In the simulations with only *B. bifidum* and *B. vulgatus*, both species became abundant in absence of 2'-FL (Fig. 4.4A). In presence of 2'-FL, and with the non-public good variant of 2'-FL metabolism, *B. bifidum* became much more abundant than *B. vulgatus* (Fig. 4.4A) . However, with public goods 2'-FL metabolism *B. vulgatus* became much more abundant than *B. bifidum* (Fig. 4.4A), despite *B. vulgatus* not consuming 2'-FL itself. The cross-feeding network showed that without public goods metabolism of 2'-FL, *B. bifidum* did not produce lactose as a public good (Fig. 4.4B). In contrast, in the model with public goods metabolism of 2'-FL, *B. vulgatus* fed on the lactose produced by *B. bifidum* from 2'-FL (Fig. 4.4C). We concluded that the model predicted that the public goods metabolism of 2'-FL caused *B. vulgatus* to become relatively more abundant than *B. bifidum*, compared to the condition without 2'-FL.

Mucin consumption in the limited consortium of *B. bifidum* and *B. vulgatus* was high in all conditions, regardless of their relative abundances, as both species consume mucin (Fig. 4.4D). To confirm whether *B. longum* reduced mucin consumption in competition with *B. vulgatus* we repeated the conditions with lactose and non-public goods digestion of 2'-FL (conditions 1-3), but with only *B. longum* and *B. vulgatus*. As neither species can use public goods 2'-FL digestion we did not repeat condition 4. In absence of 2'-FL (conditions 1&2), *B. longum* and *B. vulgatus* were approximately equally abundant (Fig. 4.4E). In contrast with *B. bifidum*, in presence of 2'-FL *B. longum* became consistently more abundant than *B. vulgatus* (Fig. 4.4E). *B. longum* also did not create lactose from 2'-FL like *B. bifidum* did (Fig. 4.4F). Mucin consumption with 2'-FL was also lower compared to the simulations without 2'-FL ($p<0.01$,Fig. 4.4G).

In the limited consortium of *B. longum* and *B. vulgatus* the reduction in mucin consumption with 2'-FL was small, and *B. vulgatus* was still consistently present. We observed that *B. vulgatus* cross-fed on lactate produced by *B. longum* (Fig. 4.4F). We hypothesize that the consumption of lactate may have allowed *B. vulgatus* to survive in this limited consortium, but that this does not occur in the full consortium (Fig. 4.3D). We next added *Roseburia inulinivorans*, a lactate consumer from the full consortium, to see if it could out-compete *B. vulgatus* and so lower mucin consumption. *R. inulinivorans* indeed out-competed *B. vulgatus* (Fig. 4.4E) and mucin consumption was indeed lower (Fig. 4.4G, $p<0.01$). To control for a possible effect of *R. inulinivorans* on mucin consumption without 2'-FL or without *B. longum* we

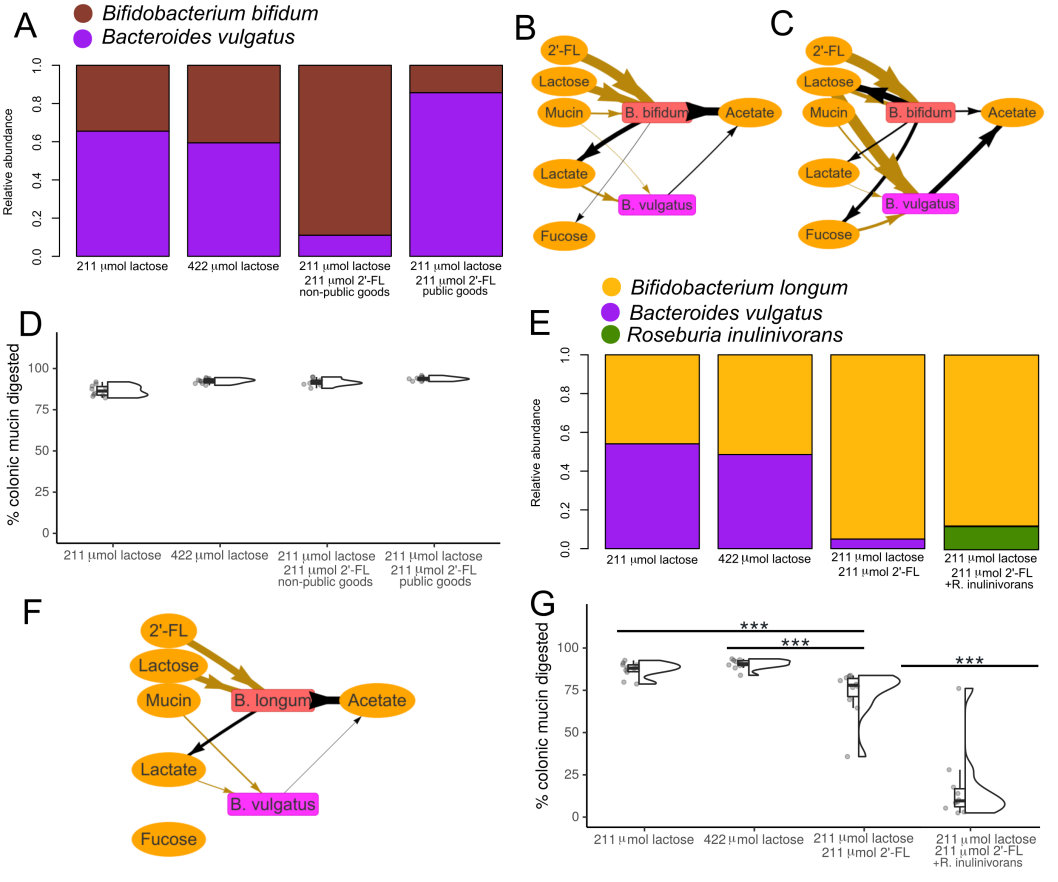


Figure 4.4: *B. bifidum*, but not *B. longum*, is exploited when digesting 2'-FL
(A) Average relative abundance of *B. bifidum* and *B. vulgatus*, with per 60 timesteps: 211 µmol lactose, 422 µmol lactose, 211 µmol lactose and 211 µmol 2'-FL with public goods 2'-FL metabolism, or 211 µmol lactose and 211 µmol 2'-FL without public goods 2'-FL metabolism, at the end of 21 days. n=10 per condition
(B,C) Visualisation of metabolic interactions in sample simulations with only *B. bifidum* and *B. vulgatus* for (B) the non-public goods metabolism and (C) the public goods metabolism. Line width is scaled with the flux per metabolite over the last 60 timesteps, multiplied by the carbon content of the molecule, with a minimum threshold of 100 µmol atomic carbon. Data from the last 3 hours, step 10020 to 10080.
(D) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 3 hours of the model, per condition of A. n=10 for each.
(E) Average relative abundance of *B. longum* and *B. vulgatus*, with 211 µmol lactose, 422 µmol lactose, or 211 µmol lactose and 211 µmol 2'-FL per 60 timesteps, at the end of 21 days. The simulations of the bar furthest to the right also include *R. inulinivorans*.
(F) Visualisation of metabolic interactions in a sample simulation with only *B. longum* and *B. vulgatus*. Line width is scaled with the flux per metabolite over the last 60 timesteps, multiplied by the carbon content of the molecule, with a minimum threshold of 100 µmol atomic carbon. Data from the last 3 hours, step 10020 to 10080.
(G) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 3 hours of the model, per condition of E. n=10 for each.
NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$

performed several additional simulations. With *B. longum*, *B. vulgatus* and *R. inulinivorans* without 2'-FL there was no reduction in mucin consumption compared to the same condition without *R. inulinivorans* (Fig. S7A&B, $p>0.05$). With *B. bifidum*, *B. vulgatus* and *R. inulinivorans* there was also no reduction in mucin consumption compared to the same condition without *R. inulinivorans* (Fig. S7C&D, $p>0.05$). We therefore concluded that *B. longum* could indeed reduce mucin consumption in the model with 2'-FL, in contrast to *B. bifidum*.

In conclusion, the model predicts that *B. bifidum* is exploited by *B. vulgatus* when it grows on 2'-FL. This happens because of the public good metabolism of *B. bifidum*. The model also predicts that when *B. longum* grows on 2'-FL it is not exploited by *B. vulgatus*, because *B. longum* does not produce public goods. This allows *B. longum* to out-compete *B. vulgatus*. Because *B. longum* does not consume mucin, mucin consumption is lower when it out-competes the mucin consumer *B. vulgatus*. *B. longum* metabolites can also feed a non-mucin consuming lactate consumer that can out-compete *B. vulgatus*, further lowering mucin consumption.

4.3 Discussion

To provide a mechanistic explanation for why breastfeeding modifies the microbiota in such a way that the microbiota consumes less mucin we created a multiscale mathematical model in this study. Concretely, the model predicts that the HMO 2'-FL stimulates non-mucin consuming species at the expense of mucin consumers, thus suggesting a plausible mechanism for the potentially beneficial effect of human milk on mucin consumption [219]. The mechanism predicted by the model is as follows: 2'-FL is consumed by *B. longum*, enabling it to out-compete mucin consuming species. *B. bifidum*, which consumes both mucin and 2'-FL, will lose the competition against *B. longum* in a complex community, because of its extracellular metabolism of 2'-FL. This extracellular metabolism allows competitors to consume intermediate digestion products. *B. longum* digests 2'-FL intracellularly and is, therefore, not sensitive to such exploitation of the public goods as produced in extracellular metabolism.

4.3.1 Comparisons with experimental data

Several *in vivo* and *in vitro* observations agree with our model predictions. Our model predicted that *B. bifidum* can consume mucin (Fig. 4.1C&D). In addition,

when the model included public goods metabolism of mucin, *A. hallii* could cross-feed on the public goods that *B. bifidum* produced from mucin, in particular galactose (Fig. 4.1D). These predictions agree with experimental observations: *B. bifidum* has been shown to grow on mucin and *A. hallii* has been shown to grow on digestion products of *B. bifidum*, including galactose, *in vitro* [27]. The model also predicted that *B. bifidum* can consume 2'-FL (Fig. 4.3C). When the model included public goods metabolism of 2'-FL, the model predicted that *B. bifidum* will produce lactose and fucose which could be taken up by other species (Fig. 4.3C&4.4C). This is in line with *in vitro* observations [220, 235]. Furthermore, the model predicted that 2'-FL will stimulate the growth of *B. longum* in the infant gut microbiota (Fig. 4.2C). This is a consequence of *B. longum* consuming 2'-FL intracellularly, thereby not producing public goods that could be used by competitors (Fig. 4.2C & 4.4E). In agreement with this prediction, a high abundance of *B. longum* has also been observed *in vivo* as an effect of 2'-FL supplementation [225] or breastfeeding [53]. The model also predicted that the microbiota consume less mucin in the presence of 2'-FL (Fig. 4.2D). This is a result of the non-mucin consumer *B. longum* getting a competitive advantage over mucin consumers by 2'-FL in the model (Fig. 4.2C). In agreement with this, less mucin is digested in *in vitro* fermentations using fecal samples of breastfed infants compared to fermentations using stool samples of formula-fed infants [219]. Within the simulations with 2'-FL, the model also predicted that a high abundance of *B. longum* reduces mucin consumption by the microbiota in the presence of 2'-FL (Fig. 4.3D). In agreement with this prediction, high abundance of *B. longum* is correlated with reduced mucin consumption *in vivo* in breastfed infants [30].

The simulations also predicted that in the presence of 2'-FL the public goods produced by *B. bifidum* get consumed by *B. vulgatus*, which prevents *B. bifidum* from becoming abundant. In agreement with this prediction, *B. bifidum* has been shown to stimulate a, as of yet undetermined, species within the Bacteroidaceae, the family that contains *B. vulgatus*, in the mouse gut [236]. No *in vitro* research is available on the competition or cooperation between *B. bifidum* and *B.vulgatus* in isolation. However, there are some *in vitro* studies on competition between *B. bifidum* and the non-HMO consumer *Bifidobacterium breve* when fed with HMOs [224, 83]. These studies reported contradictory results: *B. breve* became less abundant than *B. bifidum* in a short-term *in vitro* experiment [224], whereas it became more abundant than *B. bifidum* in a chemostat experiment with a continuous input flow of 2'-FL and output flow of metabolites [83]. The flow rate seems to explain

the latter results. In the chemostat experiment containing a co-culture of *B. bifidum* and *B. breve*, the two species together were much less abundant after the flow rate was reduced [83]. However, in a mono-culture of *B. bifidum* alone, reducing the flow rate only had a minor effect on the bacterial abundance [83]. This suggests that the flow rate particularly impacts bacterial interactions. The authors hypothesised that reduced substrate availability was the main cause of the observed reduction in bacterial abundance, but this does not explain why monocultures were less affected than co-cultures [83]. Interestingly, *in vitro* research on public goods production and consumption by two *Bacteroides* species actually found that high flow rates decreased public goods consumption and total abundance [237]. As both high and low flow rates caused a low total abundance and public goods consumption, intermediate flow rates may be important for bacterial abundance and public goods consumption in these systems. Computational modelling of a wide range of 2'-FL and public goods concentrations may help determine what factors are important in the outcome of bacterial competition in the infant gut.

4.3.2 Public good interactions in other microbial systems

The model predicts a crucial role for public goods metabolism and cross-feeding on public goods in determining the abundances of major species. Public goods mechanics has been previously studied *in vitro* in budding yeast (*S. cerevisiae*)[88]. Some members of this species break down sucrose into glucose and fructose extracellularly with the enzyme invertase, but loses 99% of the glucose produced to diffusion [88]. This leads to a dynamic community of invertase producers and non-producers. The non-producers can survive by feeding on the glucose produced by other cells, and have a competitive advantage by not having to expend any energy to produce invertase. A wide spectrum of yeast communities has been described, and co-existence between producers and non-producers is typical [88, 228]. In co-existence, the producers generate enough public goods to feed both themselves and a non-producing population. Co-existence is maintained because of a 'snowdrift game' where the less common strategy (producing or not producing) has an advantage over the more common strategy [88]. In the yeast system, the ratio of producers to non-producers depends on the cost associated with production [88]. High cost or low efficiency of production leads to a smaller fraction of producers [88]. In our model, *B. bifidum* is the only producer of public goods from 2'-FL, and all other species are non-producers.

Co-existence between *B. bifidum* and other species also occurred in most of our simulations, implying that our model is somewhat similar to the system studied in yeast [88]. However, the assumptions and predictions of our model also differ in several ways from the yeast system. In the first place, there is no direct cost associated with public good production for *B. bifidum* in our model, in contrast to the yeast system [88]. In yeast, the cells expend energy to create invertase, which is modelled by directly decreasing their growth [88]. We do not represent such a direct cost in our model. The yeast system predicts that producers are more abundant than non-producers when public goods production has low or no cost [88]. Despite the lack of a direct cost for the producer *B. bifidum* in our model, it was still less abundant than most other species. This may be because *B. bifidum* suffers an implicit cost from its metabolism, which gives it a competitive disadvantage. In contrast to *B. longum*, *B. bifidum* cannot actively transport 2'-FL into the cell [222]. Furthermore, *B. bifidum* is not able to consume fucose like *B. vulgatus* can [93]. An additional explicit metabolic cost for *B. bifidum*, as the producer has in the model of [88], may further reduce its abundance. However, this cannot change the model prediction that *B. bifidum* is much less abundant than *B. longum*. Secondly, in contrast to the yeast system where digestion of sucrose is always extracellular, *B. longum* digests 2'-FL intracellularly. Such an ecological role has been coined the 'loner' type in *in vitro* cultures of *Pseudomonas aeruginosa* strains. 'Loners' consume the substrate by themselves without producing public goods. In the *P. aeruginosa* system, loners exist next to producer and non-producer types. The loner type outcompetes non-producers, because the producers can obtain more substrate than the non-producers [238]. The producers outcompete the loners because they can obtain more substrate, as extracellular metabolism is more efficient [238]. 'Loners' have also been created through genetic modification in an *in vitro* budding yeast community [239]. As expected, yeast loners outcompete non-producers, but in contrast to *P. aeroginosa*, they also outcompete producers. In addition, yeast loners died out after they outcompeted producers. Our model predicts that, like the 'loner' type in yeast, *B. longum* typically outcompetes both non-producers and producers (Fig. 4.2C). However, in our model, *B. longum* does not die out, as its metabolism is efficient enough to maintain a large population in the absence of a producer (Fig. 4.3B&D). Thus, neither of the 'loner' types described previously are an exact equivalent to the role the model predicts for *B. longum*. A third difference between the yeast system and the model predictions is that in the yeast system producers are more abundant when public good production

is more efficient, due to a higher sucrose concentration. In contrast, the outcome of our model did not depend strongly on the amount of public goods produced per timestep by each population (Fig. S2). The lack of dependence on the rate of public goods production may be because of several reasons. The presence of other food sources, such as lactose, may dampen the effect of variation in public good availability in our model. Alternatively, *B. bifidum* may be so limited by its metabolism that more efficient public goods production cannot compensate sufficiently to significantly increase its abundance. Finally, diffusion may cause the local concentration of 2'-FL to become too low to allow for more public goods to be produced. Further study could determine what factor is decisive in our system.

4.3.3 Comparison to the infant gut microbiota composition

We can compare the model predictions for the relative abundances of species in the infant gut with available *in vivo* data. Our model predicts which species from our selection of 21 species become abundant in the infant gut in response to specific interventions. The simulations with lactose, as the only carbon source, i.e. without 2'-FL, predicted that *Escherichia*, *Bacteroides* and *Bifidobacterium* become the most abundant genera, in agreement with *in vivo* data on infants at the age of three weeks [3, 31]. The model simulations with 2'-FL also predicted a high abundance of *Bifidobacterium*, which matches *in vivo* data [53, 225]. However, also various discrepancies were found between the model predictions and *in vivo* data. Firstly, the model predicts a gut microbiota composition typically consisting of only a few species (Fig. 4.2C), whereas *in vivo* measurements have shown that the gut microbiota consists of dozens of species [31]. Most of these species have a low abundance *in vivo*, but they may still influence the dynamics of the total system [3, 74]. The inclusion of more species, a more extensive portrayal of bacterial metabolism combined with more complex nutritional input into the model will most likely further increase the accuracy and representativeness of the model. Secondly, the model does not predict a dominance of Bacillota, such as *Lactobacillus* or *Streptococcus* species, in any of the simulations (Fig. 4.2C), in contrast with *in vivo* data where such dominance is sometimes seen (e.g. in 18% of infants [16]). Although it is unclear why this is seen *in vivo*, there is a positive correlation with lower gestational age and defective mucin barrier [196, 33]. These are factors that we did not include in the model. Lastly, the model does not reproduce differences in *Bifidobacterium* abundance as a result of

mucin fucosylation (Fig. S4), as has been observed in an *in vivo* study [73]. It is unclear why different mucin fucosylation leads to different *Bifidobacterium* abundances [73]. Thus, it is also unclear what additions could be made to the model to improve these predictions.

4.3.4 Future extensions

Discrepancies are likely the result of the model's incomplete and simplified representation of many aspects of the infant gut and the infant gut microbiota. At the species level, the selection used in the model is incomplete, as we do not include the vast majority of bacterial species present in the early developing infant gut [3]. We also do not include any viruses, fungi or archaea, which may play a significant role in metabolic interactions in the infant gut [200]. Future versions of the model may include a larger consortium of species. On a metabolic level, flux balance analysis (FBA) requires many assumptions regarding bacterial metabolism, growth rate and biomass production [13, 240]. We have examined the effect of different biomass reactions on the model predictions (Fig. S6) and found that the predictions were similar when the biomass reaction included acetate in addition to ATP, but not when other biomass reactions were used. This shows that the model can reproduce many results when carbon is taken up and used for growth, but that it is sensitive to the specific manner in which carbon is used for growth. These differences are most probably caused by the differences in FBA solutions. These, in turn, depend on the GEMs. Further study of the internal reactions used with the different biomass reactions may shed more light on why the system is so sensitive to these variations. Furthermore, only biomass reactions consisting of a single carbohydrate or a single carbohydrate plus ATP were tested. In reality, however, bacterial metabolism requires many more components, including amino acids, nucleotides and cell wall components. Ideally, all these should be included in the biomass reaction. However, these cellular components all include nitrogen, which requires a good modelling of nitrogen availability and metabolism in the gut to be implemented in the system. The addition of such a system to future versions of the model will likely lead to more complex dynamics, as amino acids are also a cross-feeding substrate for some species [241] and protein content in infant formula influences microbiota composition [242]. Despite these discrepancies, the assumption that the ATP production rate corresponds with biomass production allowed the model to create accurate predictions for both the overall com-

position of the microbiota and the relative abundances of *B. longum* and *B. bifidum*.

The model is currently limited to a simplified representation of mucin. Only two different mucin structures are used in the model, but many more mucin structures exist. *In vivo* mucin structure is highly diverse and contains a greater number of (more complex) mucins compared to the mucin structures used in our model [87, 243]. A more realistic representation of mucin in the model would possibly result in more accurate model predictions about the developing gut microbiota when implemented. Furthermore, although human milk contains about 200 HMO structures, we focused our modeling on 2'-FL, as 2'-FL is the most abundant oligosaccharide in most human milk [244]. We also modelled GOS, as it is a common component of infant formulas [71]. If other HMOs were to be included in the model, this would lead to more complex metabolic interactions as these HMOs are digested using different enzymes [222]. A final limitation of the current model is that it does not include factors like gut pH or the influence of the host immune system [245, 246]. These factors could also influence interactions and competition between bacteria in the infant gut, and thus influence the predictions of the model.

4.3.5 Potential health implications

The model predictions may have implications for the health and well-being of infants. Although it is currently unclear how the consumption of mucin by the gut microbiota directly impacts infant health, the observation that there is lower consumption of mucin in breastfed infants may suggest that mucin is important for normal growth and development of infants. This could be illustrated by the fact that breastfed infants are more protected against IBD and Crohn's disease later in life than formula-fed peers [247], as both disorders having been associated with damage to the mucin layer in the gastrointestinal tract [248]. Strategies to prevent mucin loss, like nutritional interventions with 2'-FL or GOS as described in this paper, can be important in this light. Of course, other effects of breastfeeding, such as immune system modulation [187, 249], may also play a role.

4.3.6 Outlook

Future versions of the model may consider a greater number of bacterial species, as well as a more complex representation of metabolism and metabolite sharing that could for instance include amino acids, nucleotides, and other oligosaccharides. The

model can also be extended to simulate the adult gut. As dietary fiber becomes more abundant in the nutrition of infants after weaning [250], extracellular metabolism and public good dynamics may play an even larger role during childhood, adolescence and adulthood. Such extensions may provide further insights into the public good dynamics at work in the human gut.

4.4 Methods

4.4.1 Model overview

We have extended a multiscale spatial model of the infant gut. The model is based on our earlier models of the human microbiota [122] and of the infant microbiota in particular [103, 229]. The major additions to this extension of the model are the inclusion of mucins and the public goods metabolism of mucins, GOS, and 2'-FL. In short, the model consists of a regular square lattice of 225×8 lattice sites (Fig. 4.2A). Each lattice site can contain a single population of a single species (Section 4.4.2, 'Species composition'). Each species is represented by a genome-scale metabolic model, to which we have made some improvements to (Section 4.4.3, 'Changes to GEMs'). Each lattice site can contain any number of nutrients and metabolites in any concentration. Metabolism is calculated for each population using flux balance analysis (FBA, section 4.4.4, 'FBA approach'). We have developed both a non-spatial (section 4.4.5, 'Non-spatial model') and a spatial version of the model (section 4.4.6 ,'Spatial model'). In the non-spatial model all populations, nutrients, and metabolites diffuse throughout the whole system every timestep. In the spatial model the populations, nutrients, and metabolites diffuse locally (Fig. 4.2A 2&3). To mimic movement through the gut all nutrients and metabolites also advect by one lattice site each timestep in the spatial model (Fig. 4.2A 4). Nutrients are input into the system at regular intervals to allow for metabolism (Section 4.4.7, 'Nutrient input'). To represent the public-goods producing digestion of extracellular oligosaccharides, such as mucin, GOS, and 2'-FL, we have also created an additional model variant (Section 4.4.8 'Extracellular metabolism'). In this variant FBA is not used to calculate the breakdown of extracellular oligosaccharides. Finally, the model is initialised with randomly placed populations of all species, and the populations are allowed to grow and divide (Section 4.4.9 'Population dynamics'). This ultimately leads to a model that predicts a dynamic ecosystem with complex bacterial compositions and

interactions, as we describe in the results.

4.4.2 Species composition

We selected the list of species to represent in the model (Table 4.1) from [3], using sheet 2 of their Table S3. We selected the 20 entries with the highest prevalence in vaginally delivered newborns. After removing two duplicate entries we selected a genome-scale metabolic model (GEM) of a species from each genus from the Virtual Metabolic Human database, generated in the AGORA project [10]. We then added an additional *Bifidobacterium breve* and *Bifidobacterium bifidum* GEM to represent the diversity of *Bifidobacterium* species in the infant gut [164]. We also added a GEM of *Roseburia inulinivorans*, as in our previous model [229]. *Roseburia* spp. have been shown to be prevalent butyrate producing bacteria in infants in other studies [20]. The populations of *Roseburia inulinivorans*, *Eubacterium hallii*, and *Clostridium butyricum* are pooled and listed as "Butyrate producers" in the visualisations.

4.4.3 Changes to GEMs

We applied various changes and additions to the GEMs, as in our previous version of the model [229]. A full list of changed and added reactions is in table S1. We will

Table 4.2: Parameters of the model

Parameter	Value	Unit
Lattice side length	2	mm
Width of lattice	225	lattice sites
Height of lattice	8	lattice sites
Timestep	180	seconds
Initial populations	540	average number
New population placement probability	0.00005	per timestep per empty lattice site
Population death probability	0.0075	per timestep per population
Initial size per population	$5 \cdot 10^7$	no. of bacteria
Population size to create new population	$1 \cdot 10^{10}$	no. of bacteria
Maximum population size	$2 \cdot 10^{10}$	no. of bacteria
ATP to grow one cell	$1 \cdot 10^{-15}$	mol
Enzymatic constraint	2	µmol flux per timestep per $1 \cdot 10^{10}$ bacteria
Public goods production rate	2	µmol per nutrient per timestep per $1 \cdot 10^{10}$ bacteria
Nutrient input	211	µmol per nutrient every 60 timesteps
Mucin input	0.5	µmol per timestep
Initial oxygen	0.1	µmol per lattice site
Metabolic advection	2	mm per timestep
Diffusion of metabolites and bacteria	$6.3 \cdot 10^5$	square cm per second

highlight the more notable changes here. We disabled anaerobic L-lactate uptake for the *Bifidobacterium* and *E. coli* GEMs [158, 208] and added a lactose symporter to *Anaerobutyricum hallii* [27], all *Bifidobacterium* GEMs [209], *Roseburia inulinivorans* [210], *Haemophilus parainfluenzae* [211], and *Rothia mucilaginosa* [212]. We also added galactose metabolism to *R. inulinivorans* [213] and *R. mucilaginosa* [212]. We have added metabolism of GOS to the *B. longum*, *B. breve* and *B. bifidum* GEMs and 2'-FL metabolism to the *B. longum* and *B. bifidum* GEMs. *B. longum* imports 2'-FL using an ABC transporter [251] and digests it to lactose and fucose inside the cell. *B. bifidum* breaks down 2'-FL to lactose and fucose using an extracellular fucosidase [220]. *B. longum* and *B. breve* import DP3 fractions of GOS intracellularly, and digest them to to monosaccharides using beta-galactosidases [85]. *B. longum* and *B. breve* break down fractions longer than DP3 extracellularly using glycoside hydrolases [86, 231]. *B. bifidum* digests all fractions extracellularly to lactose and galactose using extracellular beta-galactosidases [230]. We also added a fucosidase reaction to *B. bifidum* to allow it to remove fucose groups from mucin in line with the available literature [252].

4.4.4 FBA approach

We use a modified version of flux balance analysis with an enzymatic constraint [13, 100], as in previous model versions [103]. First, each GEM is converted to a stoichiometric matrix S. Reversible reactions are converted to two irreversible reactions, so that fluxes will be greater than or equal to 0. Reactions identified in the GEM as 'exchange', 'sink', or 'demand' take up or deposit metabolites into the environment. We assume that intracellular regulation occurs at a much faster rate than any extracellular dynamics, including population growth and spread and diffusion of extracellular metabolites. Subject to this separation of timescales, we can assume that all reactions are in internal steady state:

$$S \cdot \vec{f} = 0, \tag{4.1}$$

where $\vec{f}$ is a vector of the metabolic fluxes through each reaction in the network, in mol per time unit per population unit. Thus we apply a flux-balance analysis approach [13] to predict exchange fluxes as a function of extracellular concentrations.

Each exchange reaction that takes up metabolites from the environment F_{in} is constrained by an upper bound F_{ub} which represents the availability of metabolites

from the environment. It is determined as follows:

$$\vec{F_{in}} \leq \vec{F_{ub}}, \tag{4.2}$$

where $\vec{F_{in}}$ is a vector of fluxes between the environment and the bacterial population. $\vec{F_{ub}}$ is a vector of upper bounds on these fluxes. $\vec{F_{ub}}$ is set dynamically at each timestep t by the spatial environment at each lattice site $\vec{x}$:

$$\vec{F_{ub}}(\vec{x},t) = \frac{\vec{c}(\vec{x},t)}{B(\vec{x},t)}, \tag{4.3}$$

where $\vec{c}$ is a vector of all metabolite concentrations in mol per lattice site, $\vec{x}$ is the location and $B(\vec{x},t)$ is the size of the local bacterial population [103]. $\vec{F_{ub}}$ is set to 0 for any metabolite for which the public goods metabolism system is enabled for the local bacterial population (see section 'Extracellular metabolism').

The total flux through the network in each FBA solution is constrained by the enzymatic constraint a, in mol per time unit per population unit [103, 100]. The enzymatic constraint represents the maximum, total amount of flux that can be performed per cell in each population:

$$\sum \vec{f} \leq a. \tag{4.4}$$

As both $\vec{f}$ and a are given as a flux per population unit, this limit scales linearly with population size. Given these constraints, FBA uses linear programming to identify a solution that optimizes the objective function, ATP production. The solution consists of a set of input and output exchange fluxes $\vec{F_{in}}(\vec{x},t)$ and $\vec{F_{out}}(\vec{x},t)$, and a growth rate $g(\vec{x},t)$. The exchange fluxes are taken as the derivatives of a set of partial-differential equations to model the exchange of metabolites with the environment. The size of the population increases proportionally to the growth rate in the FBA solution. Populations above $2 \cdot 10^{10}$ bacteria do not perform metabolism to mimic quiescence at high densities.

4.4.5 Non-Spatial model

For the simulations of Fig. 4.1 we used a non-spatial version of the model to represent well-mixed *in vitro* conditions. All nutrients and metabolites are spread across all lattice sites at the end of each timestep. All populations are moved to random

locations at the end of each timestep. This eliminates spatial variation in metabolism. The change in concentration per lattice site is thus determined as follows:

$$\frac{d\vec{c}(\vec{x},t)}{dt} = \vec{F_{out}}(\vec{x},t)B(\vec{x},t) - \vec{F_{in}}(\vec{x},t)B(\vec{x},t), \tag{4.5}$$

where $\vec{F_{out}}(\vec{x},t)$ is a vector of fluxes from the bacterial populations to the environment, in mol per time unit per population unit. All other parameters, including the lattice size, are as in the spatial model.

4.4.6 Spatial model

For all simulations after those of Fig. 4.1 we used a spatial version of the model. This version also used a regular square lattice of 225 x 8 lattice sites. In the spatial model mixing by colonic contraction is mimicked by diffusion. To each lattice site we apply the exchange fluxes as predicted by the FBA solution, yielding:

$$\frac{d\vec{c}(\vec{x},t)}{dt} = \vec{F_{out}}(\vec{x},t)B(\vec{x},t) - \vec{F_{in}}(\vec{x},t)B(\vec{x},t) + \frac{D}{L^2}\sum_{\vec{i}\in \mathrm{NB}(\vec{x})}\left(\vec{c}(\vec{i},t) - \vec{c}(\vec{x},t)\right), \tag{4.6}$$

where D is the diffusion constant, L is the lattice side length, and $NB(\vec{x})$ are the four nearest neighbours.

Diffusion is applied to the metabolite concentrations on each lattice site at each timestep to represent mixing by colonic contractions. Metabolic diffusion is applied twice during each timestep. Each time it is applied, 14.25% of each metabolite diffuses from each lattice site to each of the four nearest neighbours. This causes a net diffusion each timestep of $6.3 \cdot 10^5$ cm^2/s. To mimic advection all metabolites except oxygen are moved distally by one lattice site every timestep. The transit time through the colon is approximately 11 hours in the model, corresponding with *in vivo* observations in newborn infants [141, 142]. Metabolites at the most distal column of the lattice, the end of the colon, are removed from the system at each timestep. This represents the simulated feces.

4.4.7 Nutrient input

The secretion of mucin is mimicked as follows: each timestep a small concentration of mucin is added to the bottom-most row of the model. We used a core-2 mucin

with one additional fucose, known in the VMH database as 'MGlcn23_rl', except for the simulation of Fig. S4. Here a core-2 mucin without additional fucose was used, known in the VMH database as 'core2_rl'. In most simulations nutrients representing inflow from the small intestine are inserted into the first six columns of lattice sites every 60 timesteps, representing three hours, a realistic feeding interval for newborn infants [175]. Food intake contains 211 µmol of lactose ('lcts' in the VMH database) by default, a concentration in line with human milk [66], assuming 98% host uptake of carbohydrates before reaching the colon [131]. In some simulations 211 µmol of additional lactose, GOS, or 2'-FL is added. 2'-FL and GOS were not present in the VMH database. GOS is inserted as separate fractions of DP3, DP4, or DP5 based on analysis of the composition of Vivinal-GOS [206]. 64% is DP3, 28% is DP4 and 8% is DP5. Water ('h2o' in the VMH database) is provided in unlimited quantities.

4.4.8 Extracellular metabolism

We will now discuss the details of extracellular public-goods producing metabolism in our model. The GEMs of many species in our model contain reactions for the extracellular breakdown of oligosaccharides, such as mucins. We created a separate system for public goods-producing extracellular metabolism. In the simulations where this system was enabled, the oligosaccharides were excluded from FBA. Instead, $\vec{F_{in}}$ values for extracellular oligosaccharides were set directly by the environment as follows: The model applies each extracellular reaction at a rate of 2 µmol per $1 \cdot 10^{10}$ population per timestep. This is entered into the sets of input exchange fluxes $\vec{F_{in}}$ and output exchange fluxes $\vec{F_{out}}$. If two reactions can apply to a substrate, and insufficient substrate is available for each reaction to apply fully, each reaction is applied to half the remaining substrate. As no substrates have more than two reactions associated with them in our model, this ensures that breakdown is limited to the total amount of substrate available:

$$\vec{F_{in}}(\vec{x},t) \leq \frac{\vec{c}(\vec{x},t)}{B(\vec{x},t)}, \tag{4.7}$$

The change in concentration of metabolites is still determined by eq. 4.5 in the non-spatial model and eq. 4.6 in the spatial model.

Because the fluxes from the FBA and the separate extracellular metabolism are applied simultaneously they cannot interact within their own timestep. Diffusion

and advection are applied before the products can be used by FBA. The public goods metabolism is applied to mucins in all simulations of the model, except those in Fig. 4.1C. The public goods metabolism is also applied to extracellular 2'-FL or GOS, and is noted as "public goods" when used. All reactions related to 2'-FL and GOS breakdown are listed in S1 table. All mucin breakdown reactions are listed in table S2. In short, this approach lets us mimic the production of a larger variety and quantity of public goods from oligosaccharides than the FBA approach does.

4.4.9 Population dynamics

The model is initialized by giving each lattice site a probability of 0.3 to generate a population of $5 \cdot 10^7$ bacteria of a single random species. To this end, a GEM corresponding with this species (section 'Species composition') is associated with this lattice site. The exchange rates of metabolites for each population are calculated using FBA, based on the GEM, the enzymatic constraint a, its current population size $B(\vec{x},t)$ and the local concentrations of metabolites $\vec{c}(\vec{x},t)$. The outcome is applied to the environment (eq. 4.6) and the growth rate $g(\vec{x},t)$ to the local population size, as follows:

$$\frac{dB(\vec{x},t)}{dt} = B(\vec{x},t)g(\vec{x},t). \tag{4.8}$$

After initialisation, new bacterial populations can be created through reproduction (1) or introduction of new species (2). To mimic reproduction (1), each population of at least $1 \cdot 10^{10}$ bacteria (Table 4.2) creates a new population of the same species in an adjacent empty lattice site. Half the population size is transferred to the new population, so that biomass is conserved. To mimic introduction of species (2), a new species is selected at random with equal probability and is introduced into an empty lattice site with a probability of 0.00005 for each empty lattice site. We initialize these populations at the same population size B as the initial populations in the model (Table 4.2). Finally, population are removed from the system with a probability of 0.0075 per population per timestep.

To mix the bacterial populations, the lattice sites swap population contents each timestep. We use a random walk algorithm based on Kawasaki dynamics [176], also used previously [103, 122]. Each site is addressed in a random order, and the contents are swapped with a site randomly selected from the Moore neighbourhood. The contents consist of the bacterial population size $B(\vec{x},t)$ and the GEM. The swap

only occurs if both the origin and destination site have not already swapped in this timestep. With this mixing method the diffusion constant of the bacterial populations is $6.3 \cdot 10^5 cm^2/s$, equal to that of the metabolites. Bacterial populations at the most distal column, i.e. at the exit of the colon, are removed from the system. To increase the bacterial diffusion rate in the sensitivity analysis this process was executed five times, marking all sites as unswapped after each execution. To decrease the bacterial diffusion rate the number of swaps was limited to a fifth of the usual number of swaps.

4.4.10 Validity checks on FBA solutions

We performed a number of checks on the FBA solutions to ensure the model produces plausible predictions. We first checked whether the right species could grow on lactose. In line with *in vitro* literature all GEMs could grow on lactose, except *Veillonella disparans* [95], *Cutibacterium acnes* [195], *Eggerthella* sp. YY7918 [214], and *Gemella morbillorum* [215]. No GEM could grow without a substrate. We also checked each FBA solution for thermodynamic plausibility during the simulations using a database of Gibbs free energy values [173]. Values for 2'-FL, GOS, and mucin structures were generated using the values for their monosaccharides. All values assumed a pH of 7 and an ionic strength of 0.1 M, which are close to cytoplasmic values for common gut bacteria under acidic gut conditions [253, 254, 255]. Energy loss l in joules per timestep per population unit was recorded as follows, where i are metabolites, F is the exchange flux rate in mol per timestep per population unit and E contains the Gibbs free energy in joules per mol for each metabolite,

$$l = \sum_i F(i) \cdot E(i). \tag{4.9}$$

In the simulations of Fig 4.2 with 211 µmol of lactose per 60 timesteps (n=30) 99.6% of all FBA outcomes had a lower or equal amount of Gibbs free energy in the output compared to the input. The remaining 0.4% of FBA solutions was responsible for 0.07% of total bacterial growth. In the simulations with an additional 211 µmol of 2'-FL per 60 timesteps and public goods metabolism of 2'-FL (n=30) 99.3% of all FBA outcomes had a lower or equal amount of Gibbs free energy in the output compared to the input. The remaining 0.7% of FBA solutions was responsible for 0.04% of total bacterial growth.

4.4.11 Parameters

Parameters of the system are listed in table 4.2. We estimate that the infant colon has a volume of 90ml [135, 136]. This leads to a rough estimate on the order of 10^{12} bacteria in the newborn infant colon given an abundance per ml of around 10^{10} [5]. Values for free parameters were estimated and evaluated in the sensitivity analysis(Fig. S2).

4.4.12 Implementation details

The model is implemented in C++ 11 with libSBML 5.18.0 for C++ to load GEMs and the GNU Linear Programming Kit 4.65 (GLPK) to solve the FBA problems. Random numbers were generated with Knuth's subtractive random number generator algorithm [256]. Diffusion of metabolites was implemented using the Forward Euler method. The model is based on our own earlier models of the gut microbiota [122, 103, 229]. GEMs are sourced from the May 2019 update of AGORA, the latest at time of writing, from the Virtual Metabolic Human Project website (vmh.life). We used Python 3.6 to extract thermodynamic data from the eQuilibrator API (December 2018 update) [173]. All p-values were calculated with R 4.2.2. Unless noted otherwise p-values were calculated using the Mann-Whitney test. Model screenshots were made using the libpng16 and pngwriter libraries. Other visualisations were performed with R 4.2.2. Raincloud visualisations used a modified version of the Raincloud plots library for R [216].

4.5 Supplemental material

Available from
`drive.google.com/drive/folders/11pW77mica7ugorC0Gg8MwCaaniS9PWrz`

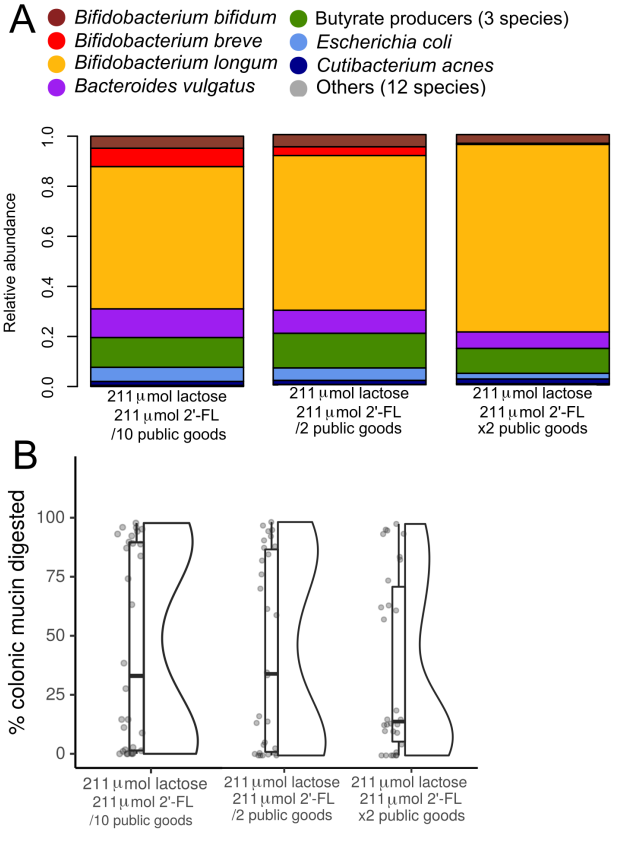
S1 Table

S1table.csv
A table of changes made to the AGORA models as a .csv file.

S2 Table

S2table.csv

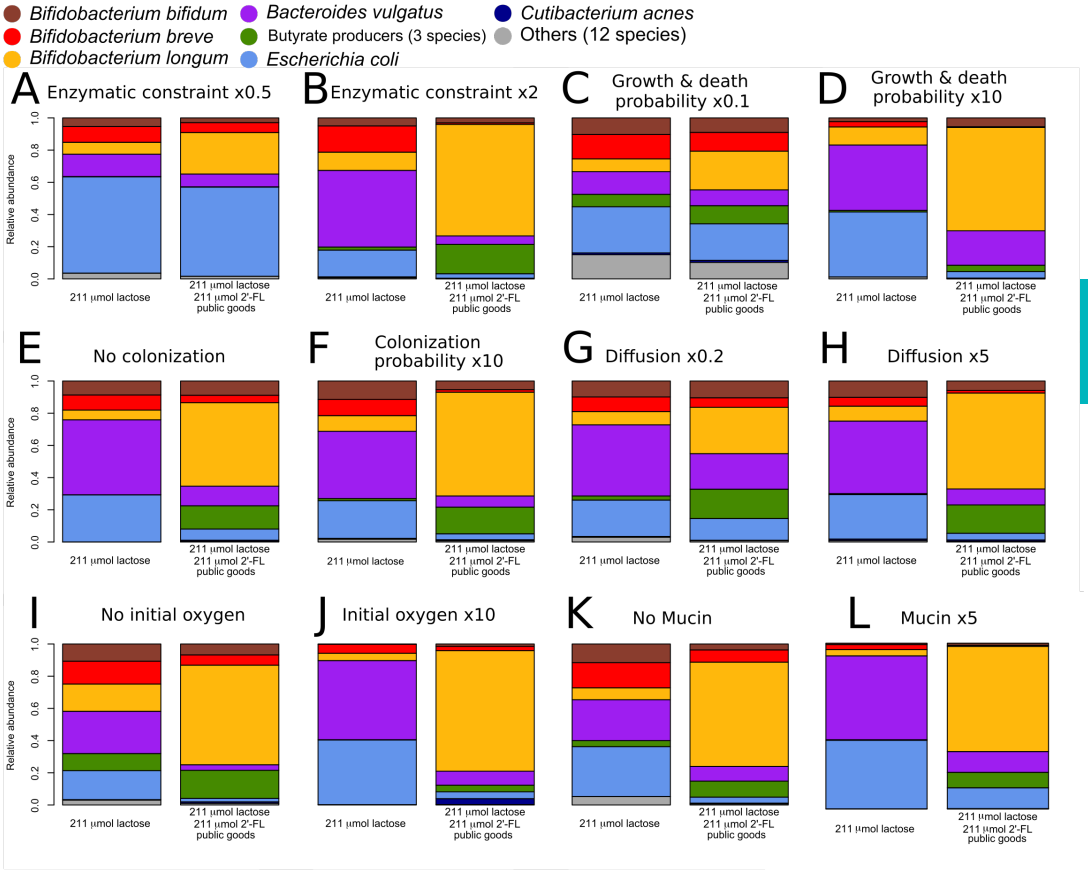
A table of mucin reactions used for public goods metabolism as a .csv file.



S1 Figure

(A) Average relative abundance of bacterial species in the condition with 2'-FL with either the public good production rate reduced to a tenth, halved, or increased to double. n=30 for each condition.

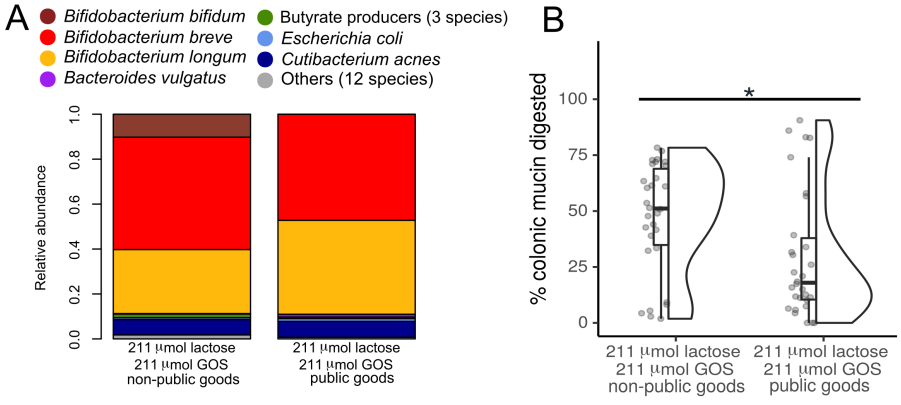
(B) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 60 timesteps of the model, per condition corresponding to A. n=30 for each condition



S2 Figure

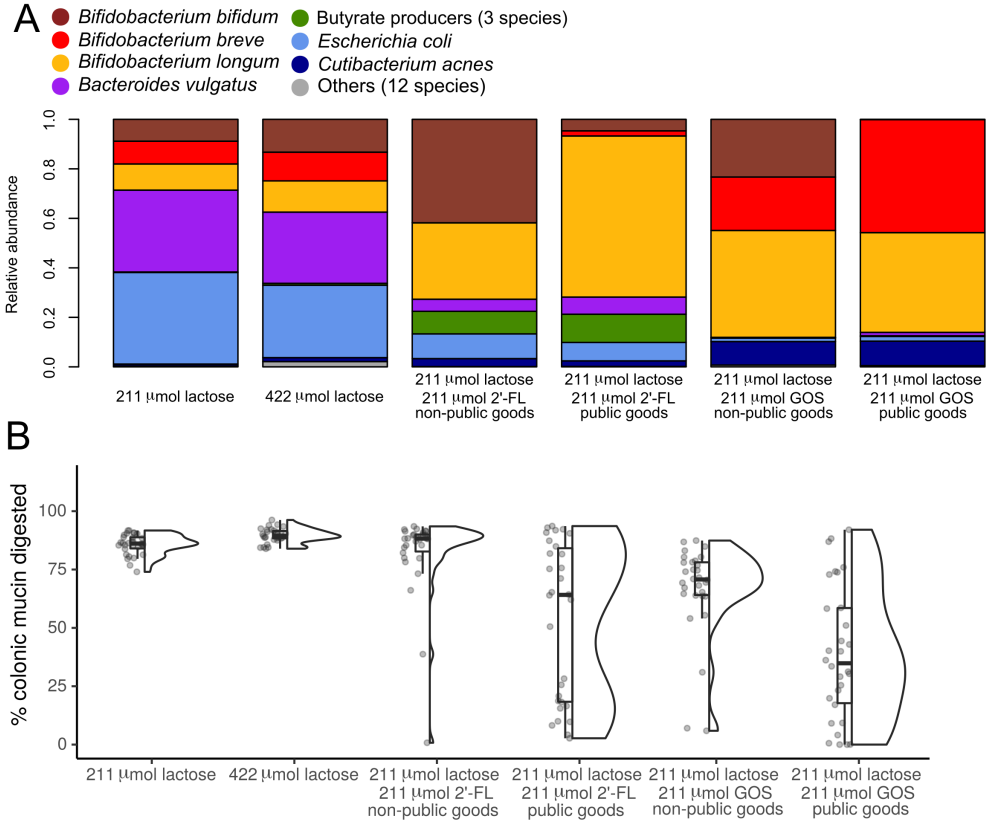
(A to J) Average relative abundance of bacterial species in the conditions with only lactose, or with 2'-FL and public goods 2'-FL metabolism, at the end of 21 days, with the following alteration from the baseline of Fig. 4.2A: (A) Enzymatic constraint loosened by a factor of 2, to 4 µmol flux per timestep per $1{\cdot}10^{10}$ population (B) Enzymatic constrained tightened by a factor of 2, to 1 µmol flux per timestep per $1 \cdot 10^{10}$ population (C) Growth decreased by a factor of 10, by increasing the ATP required to grow one bacteria to $1 \cdot 10^{-14}$, with the death probability decreased to 0.00075 per population per timestep. (D) Growth increased by a factor of 10 by decreasing the ATP required to grow one bacteria to $1 \cdot 10^{-16}$, with the death

probability increased to 0.075 per population per timestep (E) Colonisation removed by setting the probability for new populations to be placed after initialization to 0 (F) Colonisation increased by a factor of 10 by setting the probability per empty lattice to acquire a new population to 0.0005 per timestep (G) Diffusion of both metabolites and bacteria decreased by a factor of 5 to $1.26 \cdot 10^{-6}$ cm^2/s (H) Diffusion of both metabolites and bacteria increased by a factor of 5 to $3.15 \cdot 10^{-5}$ cm^2/s (I) No initial presence of oxygen (J) Initial oxygen increased to 1 µmol per lattice site (K) No secretion of mucin (L) Secretion of mucin increased to 2.5 µmol per timestep. For each figure: n=30 for each condition, each simulation is weighed equally.



S3 Figure

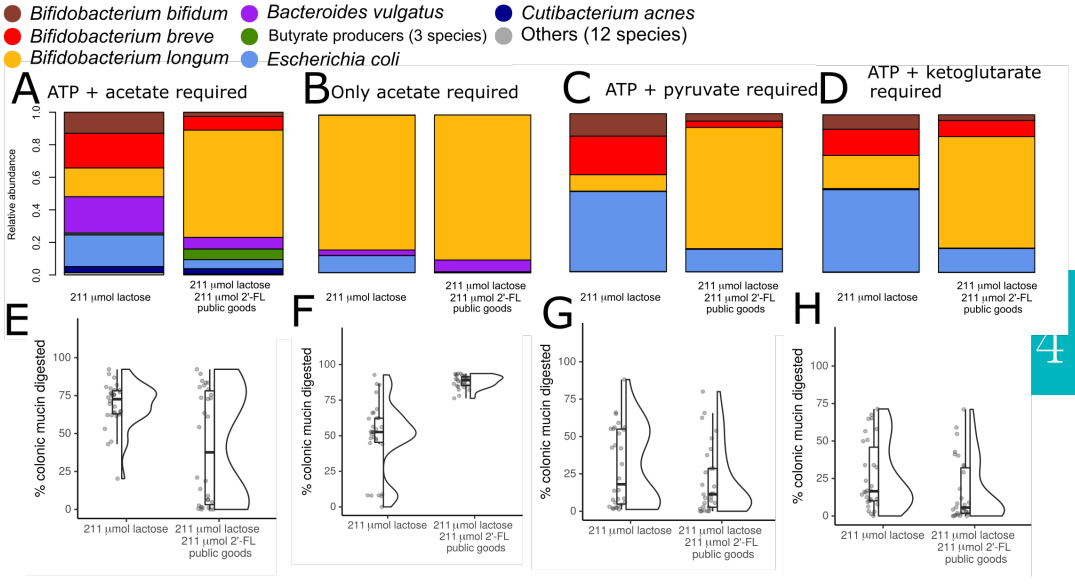
(A) Average relative abundance of bacterial species in the condition with GOS without public goods metabolism of GOS or GOS with public goods metabolism of GOS at the end of 21 days. n=30 for each condition.
(B) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 60 timesteps of the model, per condition. n=30 for each condition



S4 Figure

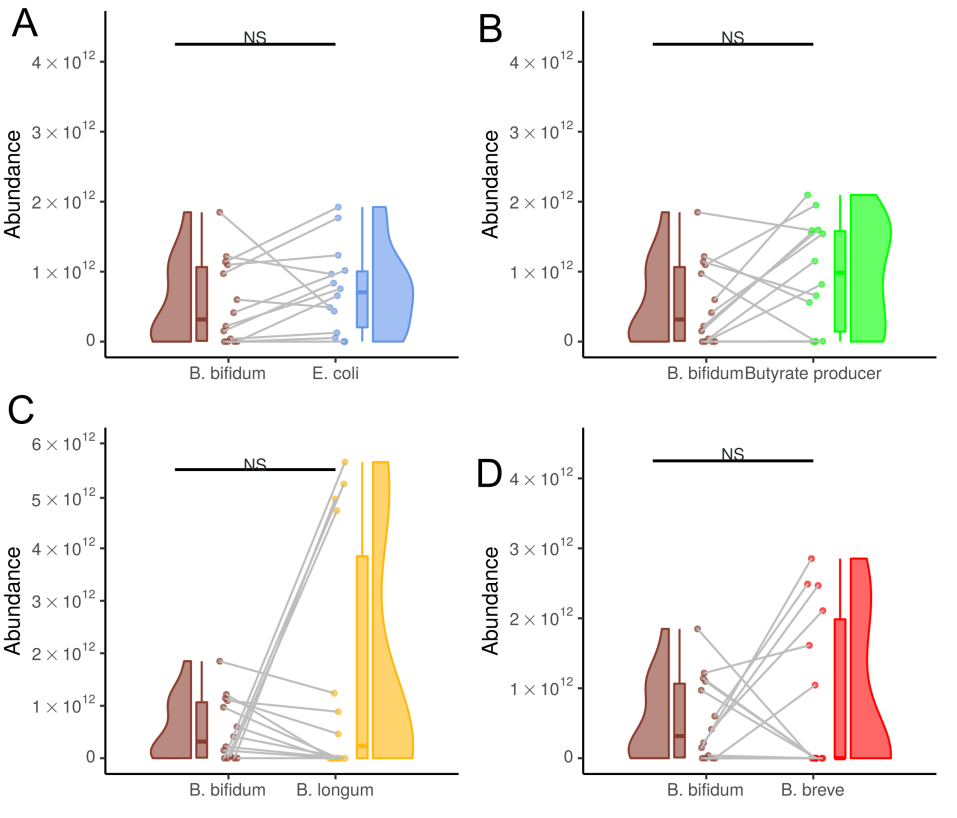
(A) Average relative abundance of bacterial species with non-fucosylated mucin instead of fucosylated mucin in each condition from Fig. 4.2and Fig. S4 at the end of 21 days. n=30 for each condition.

(B) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 60 timesteps of the model, per condition. n=30 for each condition



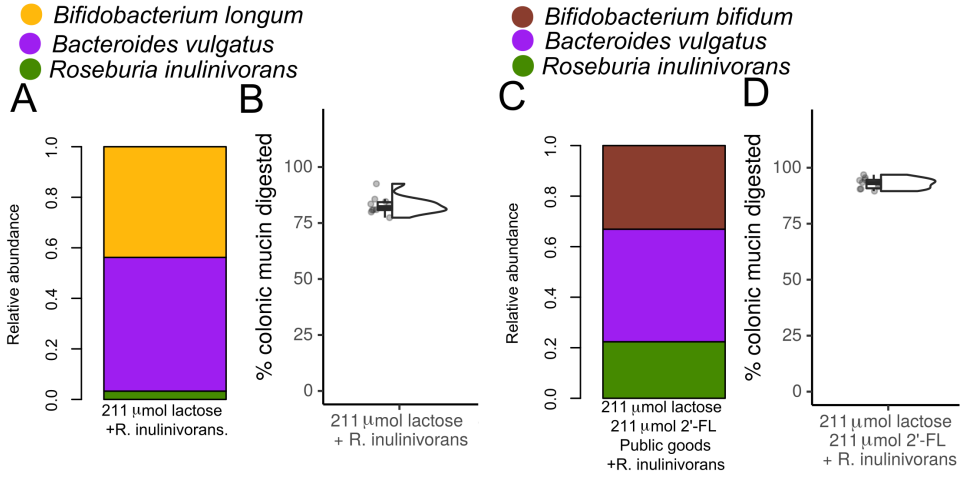
S5 Figure

(A-D) Average relative abundance of bacterial species at the end of 21 days with the following alternative biomass reactions: (A)ATP + acetate (B) acetate (C) ATP + pyruvate (D) ATP + ketoglutarate n=30 for each condition.
(E-H) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 60 timesteps of the model, associated with A-D. n=30 for each condition



S6 Figure

(A,B,C,D) Absolute abundance of *B. bifidum* compared to (A). *E. coli* (B) Butyrate producers (3 species) (C) *B. longum* (D) *B. breve* in the simulations with high mucin consumption of the public goods condition. Data from the same simulation is connected with a line. n=14.
NS: Not significant



S7 Figure

(A) Average relative abundance of *B. longum*, *B. vulgatus*, and *R. inulinivorans* without oligosaccharides at the end of 21 days. n=10

(B) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 60 timesteps of the model for A. n=10

(C) Average relative abundance of *B. bifidum*, *B. vulgatus*, and *R. inulinivorans* with 2'-FL and public good metabolism at the end of 21 days. n=10

(D) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 60 timesteps of the model for C. n=10

4.6 Contributions

J.M.W.G., and R.M.H.M acquired funding. D.M.V., J.M.W.G., C.W., and R.M.H.M. conceived and planned the simulations. D.M.V. wrote software used for the simulations. D.M.V. performed the simulations and analyzed the data. E.L., J.M.W.G., and R.M.H.M contributed to the interpretation of the results. J.M.W.G., and R.M.H.M. supervised the project. D.M.V. drafted the manuscript. D.M.V., E.L., J.M.W.G. and R.M.H.M. revised and edited the manuscript.

4.7 Acknowledgments

This study was financially supported by FrieslandCampina. E.L., and J.M.W.G. are currently employed by FrieslandCampina. This work was performed using the ALICE compute resources provided by Leiden University.

Chapter 5

Multiscale modelling of post-antibiotic recovery in the newborn infant gut microbiota

Authors
David M. Versluis[a], Ellen Looijesteijn[b], Jan M. W. Geurts[b],
Roeland M. H. Merks[a,c]
Affiliations
[a] Leiden University, Institute of Biology, Leiden, The Netherlands
[b] FrieslandCampina, Amersfoort, the Netherlands
[c] Leiden University, Mathematical Institute, Leiden, The Netherlands

In preparation

Abstract

The infant gut microbiota is a dynamic ecosystem that is crucial for infant health. The microbiota can be disturbed by antibiotics, infection, or diarrhoea. In this chapter we use a computational model to predict and explain the effects of such disturbances on the infant gut microbiota. We focus in particular on antibiotics, though we expect that our results generalize to other disturbances as well. Antibiotics decrease the abundance of pathogenic bacteria. However, antibiotics typically also cause an increase in potentially harmful Enterobacteriaceae, a decrease in beneficial *Bifidobacterium* bacteria, and an overall decrease in diversity. Many factors, such as antibiotic resistance and inflammation-induced host metabolites may be involved, but it is unclear which factors are most important. We examine in particular whether bacterial dynamics can explain, at least in part, the observed effects of disturbances, and if prebiotics can improve the resilience of the gut microbiota. We investigated these factors using a computational model that describes the metabolism and ecology of the infant gut microbiota. The model predicts that strong disturbances in the gut microbiota due to simulated antibiotics exposure lead to a decrease in the temporal stability of the microbiota, which is reflected by changes in the bacterial composition and lowering of the microbial diversity. The model only reproduces the increase in Enterobacteriaceae and decrease in *Bifidobacterium* observed *in vivo* when disturbances of intermediate strength are applied to the microbiota. The model predicts that disturbances impact acid production, particularly butyrate. Supplementation with 2'-fucosyllactose or galacto-oligosacharides, common prebiotics, make the model microbiota recover more consistently after antibiotic exposure. The model still predicts a disturbed acid production with these prebiotics, which may have health impacts. Altogether, the model can function as a framework for further investigations into the potential effects of different disturbances and their interaction with nutrition.

5.1 Introduction

The infant gut microbiota, which is important for normal growth and development [180], can be disturbed by antibiotics [257], bacterial and viral infections [258], or diarrhoea [259]. These kinds of disturbances have a complex effect on the composition and metabolic activity of the microbiota, which potentially has an impact on health. In this study we use a computational multiscale model to generate predictions of how antibiotics treatment disturbs the infant gut microbiota and affects microbiota composition and metabolic activity. We expect that many of our predictions will generalize to other disturbances as well, like bacterial and viral infections and diarrhoea. Disturbances of the microbiota by antibiotics are common in infants [257]. Antibiotics are administered to infants both preventatively and to treat bacterial infections [58]. Ampicillin and gentamicin are the most commonly used for infants, but many others are in use [257]. Antibiotics in infants are specifically associated with negative health outcomes such as asthma [260], wheezing [261] and gastro-intestinal disorders [262]. These health effects have been linked to differences in the composition and metabolism of the gut microbiota caused by the antibiotics [260, 261, 262] and have been described both *in vivo* and *in vitro*. Both *in vivo* and *in vitro* antibiotic treatment impacts the composition of the microbiota by increasing the relative abundance of Enterobacteriaceae, such as *E. coli* [263, 57, 59, 60, 58, 264, 52], decreases the abundance of *Bifidobacterium* spp. [263, 57, 58, 264, 52], and lowers the overall diversity [57, 58]. Breastfeeding seems to mitigate some of these negative effects, probably through its positive effect on the abundance of *Bifidobacterium*[265]. Some infants consume infant formula with *Bifidobacterium*-stimulating prebiotic oligosaccharides [71]. These infants may also be protected to some extent.

Mechanistic insights in how antibiotics impact the gut microbiota are sparse. Some bacterial species are more resistant to antibiotics than others [266], however, antibiotic resistance is generally low in the infant gut even after treatment [264]. It has also been hypothesized that the release of oxygen or nitrate by the gut wall may lead to an increase in the facultatively anaerobic Enterobacteriaceae over the strictly anaerobic *Bifidobacterium* [267, 268, 269]. However, the increase in Enterobacteriaceae has been replicated in an *in vitro* model that did not include a gut wall [263], suggesting that oxygen and nitrate input from the gut wall are not totally explanatory. With our model, we examined whether the observed effects can, at least partially, be explained by metabolic interactions between bacteria itself. Specifically,

we aimed to answer the following questions:
(1) Can the consequences of antibiotics exposure on the infant gut microbiota described *in vivo* be reproduced in our multiscale model on the basis of metabolic interactions between bacteria alone?
(2) How do model results depend on the strength of disturbance?
(3) Can the model predict potential health effects caused by an antibiotic disturbance?
(4) What role can the prebiotic oligosaccharides 2'-fucosyllactose (2'-FL) and galacto-oligosacharide (GOS) play in recovery from an antibiotic disturbance?

To answer these questions we used and further developed an existing multiscale computational model [122, 103, 229]. An earlier version of this model had already been used to study the effect of a diarrhea-like disruption of the microbiota, but focused on the disturbed spatial structure, not on a decline in population [122]. This version of the model predicted that increased flow through the gut decreases metabolic diversity by making cross-feeding more difficult. This may explain the decreased microbial diversity often observed in patients with diarrhea [122]. As an extension to the model, we will focus here on bacterial population dynamics and model how antibiotics may influence spatial separation and cross-feeding. For this we use a more concretely defined infant gut microbiota, with genome-scale metabolic models (GEMs) specific to infant bacteria, and detailed modelling of prebiotic metabolism, as reported in our previous study [229] and chapter 5. We disturb the microbiota by killing a large part of the population on days 8 and 9, and monitor the recovery both for simulated infants that have had no prebiotics in their nutrition, and those that have had prebiotics in their nutrition since birth.

Briefly, we show that:
(1) The model can reproduce the effects of an antibiotic disturbance observed *in vivo*: an increase of the Enterobacteriaceae population, the decrease of *Bifidobacterium*, and lower diversity, but the first two effects are only reproduced at an intermediate strength of disturbance.
(2) The model predicts that antibiotics may negatively affect the production of organic acids in general in the short-term, and decrease butyrate production specifically in the longer-term.
(3) The model predicts that ongoing supplementation with prebiotics, before, during, and after antibiotic treatment, leads to a consistently *Bifidobacterium*-dominated mi-

crobiota, with smaller *Bacteroides* and *E. coli* populations, and that infants who are given prebiotic supplementation return to this composition more consistently after antibiotic treatment.

5.2 Results

5.2.1 Stronger disturbances kill more bacteria but impact relative abundances in different ways

To investigate the effect of different strengths of antibiotics-related microbiota disruption, including the complex spatial and temporal bacterial interactions, we use a multiscale computational model based on our earlier models [122, 103]. The spatial component of the model consists of a regular square lattice of 225×8 (Fig. 5.1A), where each lattice site can contain a single population of a single species, as well as any number of metabolites in any concentration. Each species has its own GEM, selected from an existing database [101]. We use a set of 21 species derived from *in vivo* data [3] (Table 5.1, methods section 'Species composition'). We simulate the model in timesteps, each of which represents three minutes of simulated time. At each timestep, the model uses the nutrients and metabolites available in each lattice site, and the GEM of each population, to calculate metabolism per population using flux balance analysis (FBA). For each population, FBA returns a set of inputs and outputs, which we apply to the local environment as well as a growth rate, which we apply to the local population (Fig. 5.1A-1). Populations that grow large enough can spread to neighbouring lattice sites. There is also a small probability for each population to be removed from the system after each timestep, to represent natural turnover of populations. The model is initialised with a probability for each lattice site to acquire a population of a random species. Every timestep, there is small chance for each empty lattice site to acquire a new population of a random species, to represent the input of bacteria from the environment. Populations diffuse throughout the lattice (Fig. 5.1A-2). Nutrients and metabolites diffuse (Fig. 5.1A-3) and advect towards the rightmost edge of the system (Fig. 5.1A-4), where they are removed. Every 60 timesteps 211 µmol of lactose is placed distributed over the most proximal columns, and 0.5 µmol of mucin is placed at the bottom-most row of the lattice every timestep. We run the model for 20160 timesteps, equivalent to a simulated 42 days.

Over time, a complex microbiota and metabolic ecology forms (Fig. 5.1B).

Table 5.1: Species and subspecies included in the model. Color indicates group and color used in figures.

Name	Phylum
Bifidobacterium longum ssp. *infantis*	Actinomycetota
Bifidobacterium breve	Actinomycetota
Bifidobacterium bifidum	Actinomycetota
Collinsella aerofaciens	Actinomycetota
Cutibacterium acnes	Actinomycetota
Eggerthella sp. *YY7918*	Actinomycetota
Rothia mucilaginosa	Actinomycetota
Anaerobutyricum hallii	Bacillota
Clostridium butyricum	Bacillota
Roseburia inulinivorans	Bacillota
Enterococcus faecalis	Bacillota
Gemella morbillorum	Bacillota
Lactobacillus gasseri	Bacillota
Ruminococcus gnavus	Bacillota
Staphylococcus epidermidis	Bacillota
Streptococcus oralis	Bacillota
Veillonella dispar	Bacillota
Bacteroides vulgatus	Bacteroidota
Parabacteroides distasonis	Bacteroidota
Escherichia coli SE11	Pseudomonadota
Haemophilus parainfluenzae	Pseudomonadota

In our model we represent the effect of antibiotics treatment by increasing the population death probability on days 8 and 9 of the model, to mimic a typical duration for an antibiotic disturbance [57]. This causes the population size of all species to decrease greatly (Fig. 5.1C). Because nutrient input and bacterial metabolism continues, this leads to a unique new environment.

Our model is stochastic, which allows it to reach a wide variety of outcomes in terms of the abundances of species and metabolites, even when the parameters are constant. This spectrum of outcomes is similar to that of the *in vivo* infant gut microbiota [103, 31]. However, stochastic behaviour necessitates the comparison of bacterial populations of separate control and test simulations. We performed sets of 30 simulations without any disturbance, and with four levels of severity of disturbance. Each simulation had a unique random seed, for a total of 150 unique random

seeds. The severity of disturbances ranged from an increased death probability for all populations by a factor of 1.75 to a factor of 2.5. We first quantified how the different death probabilities impacted the absolute abundance immediately after the disturbance period ended at the end of day 9 (Fig. 5.1D). Disturbances caused much of the population to be killed, and more severe disturbances killed more of the population, ranging from 81% to 99%. We will use these percentages to indicate the disturbance strength from here on. When 99% of the population was killed, the system was effectively reinitialised in many simulations, as the model only maintained the concentration of metabolites and nutrients.

We next analysed the relative abundances of bacterial groups in the control and antibiotics-exposed simulations. To make comparisons easier, and allow focus on large-scale differences, we grouped the three *Bifidobacterium* species and the three butyrate producers together as separate groups, since they have similar metabolisms and ecological roles [90, 91, 20]. We also merged the 13 other species that remained at very low abundance in all simulations (indicated in Table 5.1). Where relevant, we also performed analysis with all species separately.

As shown in Fig. 5.1E, the model predicts a microbiota on day 42 composed mostly of *Bacteroides vulgatus*, *Bifidobacterium* spp., and *Escherichia coli*, in line with the most common genera in infants at this age [3, 31]. The antibiotics exposure that killed 81% of the total population had little effect, whereas an exposure leading to a reduction of 91% lowered the abundance of *Bifidobacterium* compared to the undisturbed simulations (p=0.001) and increased the abundance of *E.coli* (p=0.019) significantly, in agreement with reported *in vivo* observations [57, 58]. Interestingly, the microbiota composition changed differently for disturbances killing more than 91%. In these situations *Bifidobacterium* remained approximately at the level of the control situation. When the severity of the antibiotics disturbance was further increased to a 99% reduction of bacterial abundance, the relative abundance of *E. coli* was reduced (p=0.011), while butyrate producers became more abundant (p=0.002). These results show how the various degrees of antibiotic disturbances impact the model in a complex and non-linear way. To further analyse these effects we investigated the changes in more detail over time.

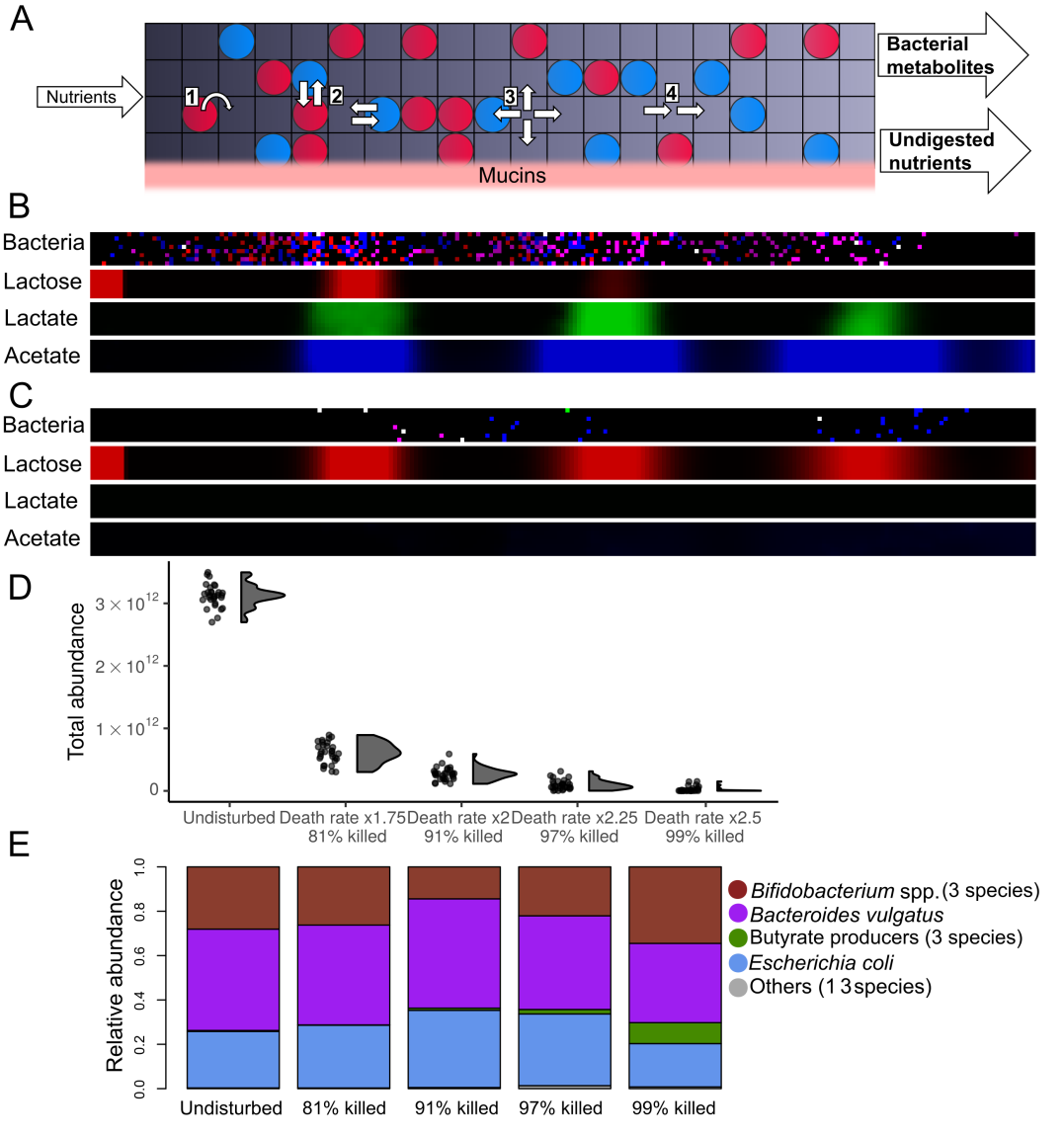


Figure 5.1: Model predicts complex effect of disturbance on relative bacterial abundances

(A) Schematic of the model. Circles represent bacterial populations. Numbers indicate processes in the model: (1) Bacterial growth and division (2) Bacterial diffusion (3) Metabolic diffusion (4) Metabolic advection **(B,C)** Screenshot of the model at the start of day 10, displaying the bacterial layer and three of the metabolic layers, in (B) an undisturbed simulation (C) a simulation with a modelled disturbance **(D)** Total bacterial abundance immediately after the disturbance (at the start of day 10) per condition. n=30 per condition. **(E)** Relative abundance of bacterial species on day 42 without disturbance or with a disturbance on the simulated day 8 and 9, n=30 per condition

5.2.2 Disturbances disrupt species composition and decrease diversity; correlation to pre-exposure populations is lost

Interestingly, we observed that in the conditions that killed 81% or 97% the relative abundance of all bacterial groups did not change. Did the antibiotics disturbances in these cases still have an effect on the microbiota, even when there was no difference in relative abundance? To answer this question we examined what impact each antibiotics disturbance had, by calculating the Bray-Curtis dissimilarity between abundances, as is commonly performed on microbial abundance data [20, 4]

We visualised the dissimilarity between the grouped relative abundances at the endpoint of the undisturbed simulations and at each timepoint in each set of simulations (Fig. 5.2 A&B and Fig. S1) and observed that all sets of simulations moved towards the same distribution (Fig. 5.2 A&B and Fig. S1), but stronger disturbances appeared to remain more distant from the undisturbed simulations. This was also the case when we analysed the dissimilarity without grouping bacterial abundances (Fig. S2), or when we analysed the Euclidean distance instead of the Bray-Curtis Dissimilarity (Fig. S3&S4, methods section 'Analysis'). To quantify the difference between disturbed and undisturbed simulations, we calculated the Bray-Curtis dissimilarity on day 42 between each simulation and the mean of the undisturbed simulations. ANOVA analysis revealed a significant effect of the disturbances on the dissimilarity ($p<0.0001$). Each disturbance that killed more than 81% was further away from the baseline than the disturbance that killed 81% ($p<0.001$ for all,Fig. 5.2C). Even the 97% condition, which did not differ from the undisturbed condition in mean bacterial abundances per group, did differ in this metric. When we did not group the species, the differences were still significant, but less so ($p=0.04$,$p=0.008$ and $p<0.001$, Fig. S5A). We also examined the Bray-Curtis dissimilarity between each pair of simulations within each condition, and found that conditions with a stronger disturbance had a higher dissimilarity ($p=0.005$ for the 81% disturbance, $p<0.001$ for all others, Fig. S5B). This shows that the disturbances not only cause a higher dissimilarity from a non-disturbed baseline, but also cause a wider variation of outcomes. We concluded that all disturbances influence the microbiota, and that disturbance strength matters for the effect.

We next examined whether the disturbances also affected species diversity, as is often observed *in vivo*. We calculated Shannon diversity, a common metric for bacterial diversity [57, 58], for each simulation to give a metric for how the distribu-

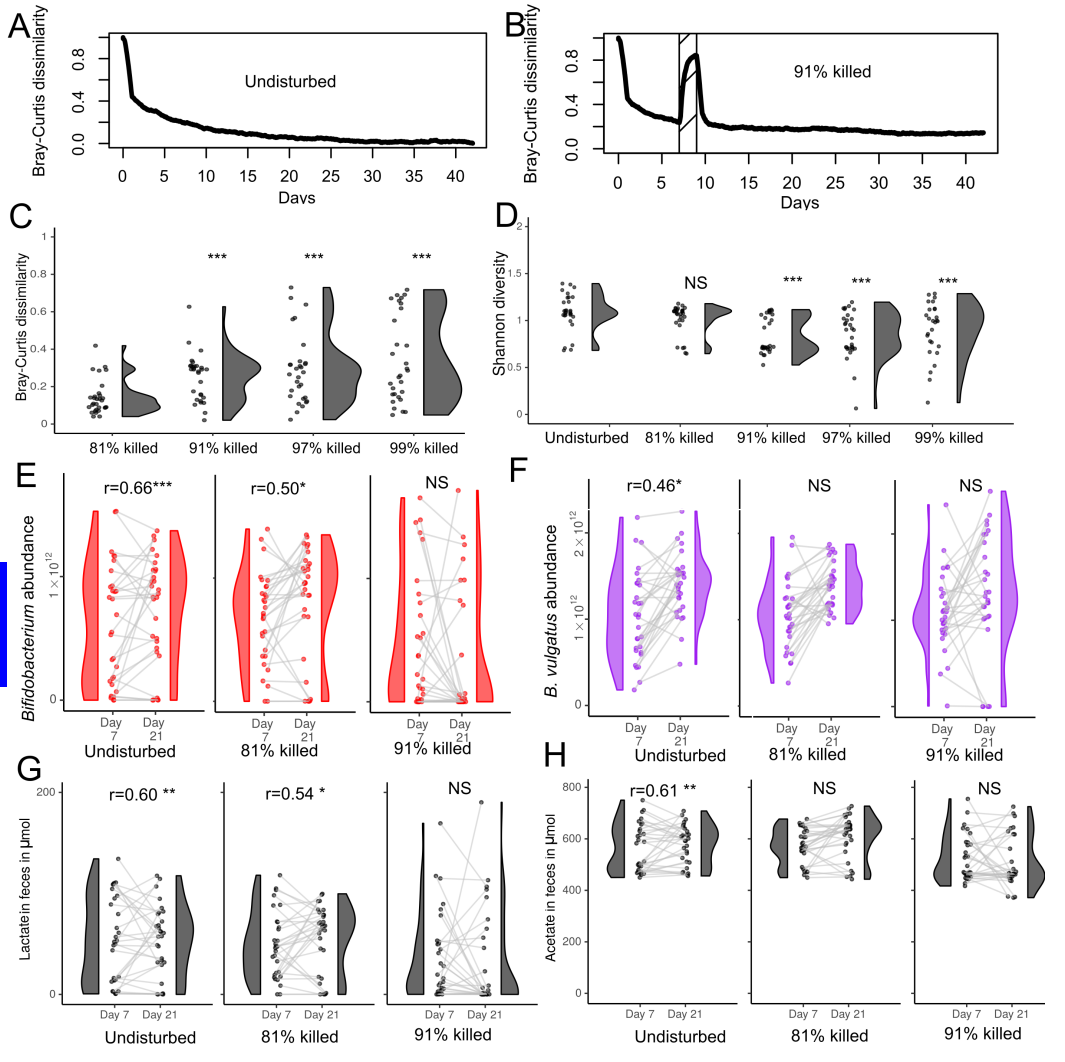


Figure 5.2: Stronger disturbances cause a larger distance from baseline, lower diversity, and disturb temporal stability more

(A,B) Bray-Curtis dissimilarity between the mean abundance at every timestep of the major bacterial groups (Table 5.1) the undisturbed condition and the condition where 91% were killed in Fig. 1D, compared to the mean abundance of the undisturbed simulations at day 42 (timestep 20160). Striped area indicates the duration of the disturbance. **(C)** Bray-Curtis dissimilarity by simulation between the mean relative abundance grouped by group, as in table 5.1, at day 42 (timestep 20160), compared to the mean relative abundance grouped in the same way of the undisturbed simulation at day 42. n = 30 per condition **(D)** Shannon diversity for all simulations in each condition at day 42 (timestep 20160). **(E,F)** Absolute abundance of (E) *Bifidobacterium* spp. (F) *B. vulgatus* per simulation at the end of day 7 and day 21 for the first three conditions of Fig. 5.1E. Points from the same simulations are connected with a line. Spearman correlation between day 7 and day 21 indicated where significant. n=30 per condition **(G,H)** Quantity of (G) lactate (H) acetate per simulation at the end of day 7 and day 21 for the first three conditions of Fig. 5.1E. Points from the same simulations are connected with a line. Spearman correlation with Bonferroni correction between day 7 and day 21 indicated where significant. n=30 per condition. NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$

tion of diversity differs. We calculated the Shannon diversity using the abundance of each individual species in each simulation, as in [57, 58]. An ANOVA analysis found a significant effect of disturbances on the diversity ($p<0.0001$). With pairwise comparisons we found that only the simulations with disturbances where more than than 81% was killed were significantly less diverse than undisturbed simulations ($p<0.001$), but the simulations also varied greatly in their diversity (Fig. 5.2D). In all tested conditions many simulations did reach a diversity similar to the typical undisturbed situation (Fig. 5.2D). Richness was equally high in all these simulations on day 42, as all species were present in all simulations. When we removed species with an abundance below 10^{10} from the data, species richness was between 2 and 5 for each simulation on day 42 (Fig. S6A). An ANOVA analysis showed a significant effect of disturbances ($p=0.005$), but in pairwise comparisons only the set of simulations that killed 99% was significantly less rich ($p=0.0001$). When we removed species with an abundance below 10^{10} from the data, diversity remained similar (Fig. S6B). We also calculated Sheldon's index, which is a measure of evenness. We found that it matched closely with the diversity when the less abundant species were included (Fig. S6C). However, when we excluded the less abundant species, only the condition that killed 99% had less evenness than the baseline ($p<0.001$, Fig. S6D). We conclude that a reduced evenness may explain the lower diversity in the disturbed conditions, but that reduced evenness alone does not explain the decreased diversity when we remove less abundant species. We further examined the distribution of outcomes with a principal component analysis on the abundance by genus (Fig. S6E), which showed a great variability between simulations and conditions. It also showed that *Bifidobacterium*, *Bacteroides* and *Escherichia* were the major drivers of differences in abundance. Thus our stochastic model predicts that after a disturbance in some infants the microbiota either maintains or recovers its diversity completely, while in others it loses some of its diversity and evenness. The model predicts that this leads to a wide variety of outcomes.

We proceeded to examine how disturbances may lead to new outcomes, and whether we can identify a factor that functions as the 'memory' of the system: a factor that persists throughout the disturbance and can be used to predict how a microbiota responds to a disturbance. We first examined whether we could link the bacterial community composition before the disturbance to that after the disturbance. To do so we calculated the Spearman correlation within the same simulations between day 7 and day 21 for the total abundance of each major group. We did so

separately for each bacterial group, so that we have a p and r value for each bacterial group for each condition. We observed that in the undisturbed simulations a correlation between abundance on day 7 and day 21 existed for *Bifidobacterium* and for *B. vulgatus* (Fig. 5.2E&F, r=0.66 & r=0.46, p<0.001& p=0.04 including Bonferroni correction). These correlations were not present with any disturbance for *B. vulgatus*, nor were they present with at least a disturbance that killed 91% for *Bifidobacterium*. *E. coli* never had a significant correlation (Fig. S7). We conclude that bacterial abundances before a disturbance are not indicative for the abundances after the disturbance. We next examined whether the quantity of nutrients or metabolites might serve as 'memory' for the system, by also examining their correlation within the same simulations between day 7 and day 21. We measured the amount of acetate and lactate leaving the system per 60 timesteps. Acetate and lactate were correlated in the undisturbed condition (r=0.61 & r=0.60, p=0.002&p=0.003, Fig. 5.2 G&H). Interestingly, lactate remained correlated even with the disturbance that killed 81% (r= 0.54, p=0.01), though the correlation was weaker than in the undisturbed condition. This indicates that the system does have a 'memory' for the excretion of lactate that persists through a disturbance. We also tested the correlation for lactose, mucin, and butyrate, but these did not even correlate significantly in the undisturbed condition (p>0.05). Oxygen was no longer present in any simulation at the start of the disturbance. We conclude that there is no temporal correlation between bacterial abundances or metabolite concentrations in the system with a disturbance that kills at least 91%, indicating that it does not have a 'memory'.

5.2.3 Disturbances disrupt acid production and cross-feeding

We previously noted that the model predicted a decrease in *Bifidobacterium* abundance and increase in *E. coli* abundance with the disturbance that killed 91% of the population. *Bifidobacterium* is positively associated with infant health due to, among other factors, its production of acids [19, 18]. We next investigated what effects of a disturbance the model predicts for acids produced by the bacteria. We used the condition that killed 91% of the population, as this provided the closest match to *in vivo* data [57].

We examined: (1) organic acids in general, which provide resistance to pathogens[18], (2) the short-chain fatty acid butyrate, which is associated with a reduction in allergies [20] and (3) lactate, which is an important acidifier and cross-feeding substrate

[74]. For each aspect we compared the simulation where 91 % was killed with the undisturbed, control simulations. Fig. 5.3 shows the results of these analysis. We found that there was significantly less total organic acid, butyrate, and lactate in feces 480 timesteps (one simulated day) after the antibiotic disturbance ($p<0.001$, $p=0.03$,$p<0.001$ Fig. 3A-F). Though the mean amount of lactate present was also lower before the disturbance due to stochasticity (Fig. 5.3E), there was no significant difference immediately before the disturbance (timestep 3360, $p=0.445$). We further examined why there was less lactate after the disturbance, as lactate is both produced and consumed in the infant gut and in our model [74, 103]. We found that there was less lactate in the feces due to a reduced production of lactate ($p<0.001$), not an increased consumption. In fact, consumption also decreased ($p<0.001$, Fig. 3G&H). We concluded that there are many potential impacts on metabolites that are relevant for infant health. These effects may explain some of the health effects observed in infants *in vivo* [260, 261, 262].

5.2.4 Prebiotic-fed microbiotas recover more consistently from a disturbance

The simulations predicted that simulated antibiotic disturbances affect many aspects of the infant gut microbiota, including various metabolites related to infant health. We next asked whether the resilience of the system could be improved in order to reverse these effects. *In vivo*, breastfeeding protects against some of the negative effects of antibiotic treatment, which is thought to be caused by the presence of prebiotic oligosaccharides in human milk [265]. These prebiotic oligosaccharides stimulate, amongst others, *Bifidobacterium* [63, 64]. We asked if and how prebiotics, in particular the prebiotic oligosaccharides 2'-fucosyllactose (2'-FL) and galacto-oligosaccharides (GOS), could affect the recovery after disturbance in the model. 2'-FL is the most abundant oligosaccharide in most human milk samples [66], and GOS is commonly used in infant formula [71]. We simulated three conditions. In each condition we added additional sugars to the base amount of 211 µmol of lactose: (1) Addition of 211 µmol 2'-FL (2)Addition of 42.2 µmol 2'-FL and 168.8 µmol GOS, a ratio as used in infant formula, (3) Addition of 211 µmol lactose, as a control condition. The third condition allowed us to discern between effects from the specific sugars we added, and an effect from the total amount of sugar present. We used the antibiotics disturbance that killed 91% of the population in earlier simulations, as it

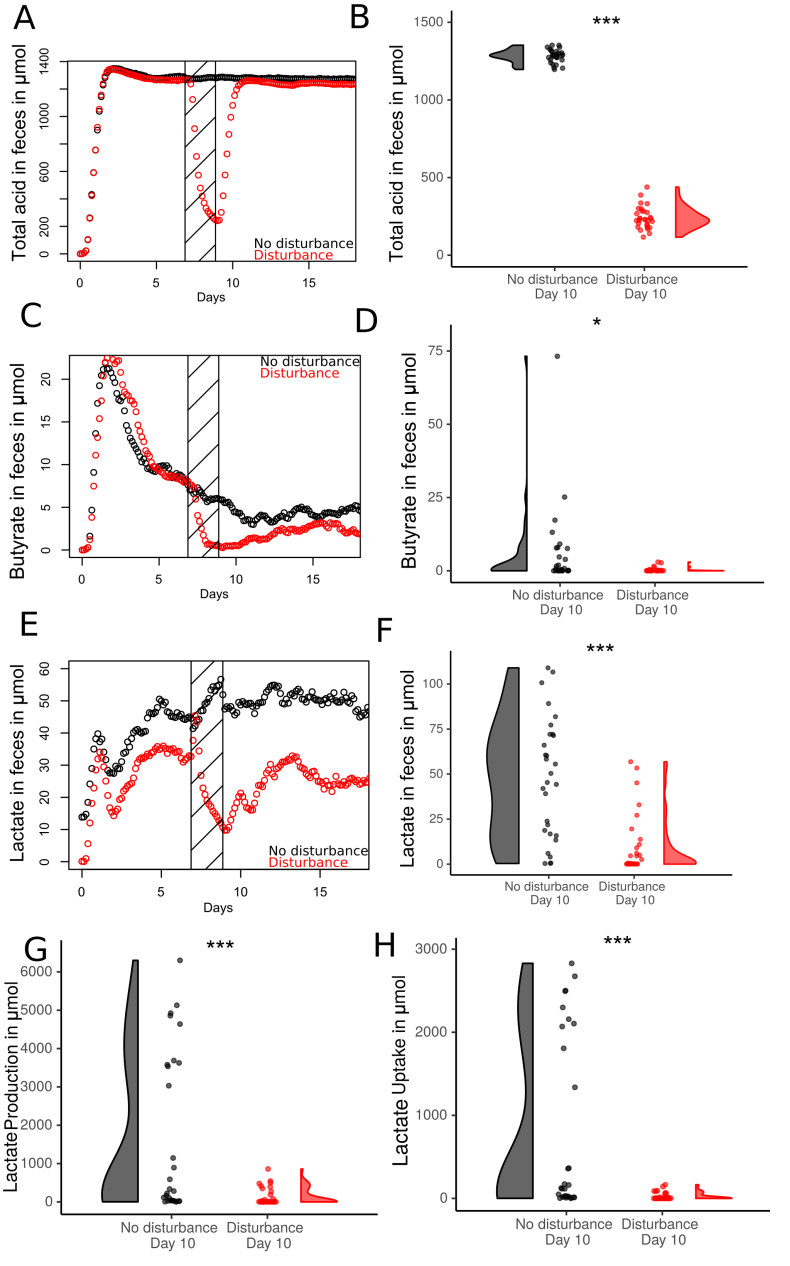


Figure 5.3: Disturbances influence metabolites and metabolism
(A,C,E) (A) Total acid, (C) butyrate, or (E) lactate leaving the system per 60 timesteps in the model with no disturbance (black) and the model with 91% killed disturbance (red). Striped area indicates the duration of the disturbance, n=30. **(B,D,F)** (B) Total acid, (D) butyrate, or (F) lactate leaving the system from timestep 4320 to 4360, immediately after the end of the disturbance, compared to the same steps in the undisturbed simulations, n=30. **(G)** Total lactate production per simulations from timestep 4320 to 4360 , immediately after the end of the disturbance, compared to the same steps in the undisturbed simulations, n=30. **(H)** Total lactate consumption per simulations from timestep 4320 to 4360, immediately after the end of the disturbance, compared to the same steps in the undisturbed simulations, n=30. NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$

produced the best match with *in vivo* data [57]. Fig. S8 shows that the combined addition of 2'-FL and GOS reduced the Bray-Curtis dissimilarity and Euclidean distance from the baseline on day 42 compared to the condition that killed 91% that did not have additional sugar (Fig. S8D&E, $p<0.001$). However, the control condition, where we only added lactose, also reduced the distance ($p=0.013$). Thus, we could not conclude that the prebiotics had an effect. The reduced Bray-Curtis dissimilarity may have been caused by a higher total bacterial abundance alone. The post-disturbance populations were indeed much larger than the post-disturbance populations in the simulations without added sugars (Fig. S8F, Fig. 5.1D).

To control for an increased total bacterial abundance, we next performed additional simulations where the total amount of sugar was kept the same as in the simulations of Fig. 5.1. We used two conditions: (1) 105.5 µmol of lactose replaced with 2'-FL (indicated with 1/2 2'-FL in fig. 5.4) (2) 21 µmol of lactose replaced with 2'-FL and 84.5 µmol of lactose replaced with GOS (indicated with 1/2 2'-FL&GOS in Fig. 5.4). We performed these simulations both without a disturbance and with a disturbance that killed 91%. The results are displayed in Fig. 5.4. Either condition led to a total bacterial abundance comparable to the abundance without prebiotics (Fig. 5.4A). Compared to the simulations with the same disturbance, but without prebiotics, there was much more *Bifidobacterium* (Fig. 5.4B& Fig. 5.1, $p<0.001$). There were no significant differences in relative abundance between the disturbed and undisturbed simulations (Fig. 5.4B, $p>0.05$). We calculated the dissimilarity between each prebiotic condition with and without a disturbance. There was a lower dissimilarity, compared to the simulations without prebiotics (Fig. 5.4C, $p=0.001$ for 2'-FL, $p<0.001$ for 2'-FL+GOS). With either prebiotic conditions there was also no reduction in diversity due to the disturbance (Fig. 5.4D, $p>0.05$). This indicates that in the model the prebiotics-influenced microbiota was more resilient to a disturbance.

We also analysed the effects on total acids, butyrate, and lactate, as we did for the simulations without prebiotics (Fig. 5.4E-J). Total acids were disturbed in both conditions on day 10 ($p<0.001$), but butyrate was only disturbed in the condition with only 2'-FL ($p<0.001$). Nonetheless, butyrate was much more abundant even with a disturbance in the condition with 2'-FL, compared to any other condition ($p<0.001$). Lactate was increased in the simulations with both 2'-FL and GOS ($p<0.001$), but not with 2'-FL alone. In conclusion, the model predicted that supplementation with prebiotics led to a more consistent return to baseline after disturbance in the model, but metabolites, particularly total acids, and total bacterial abundance were still affected.

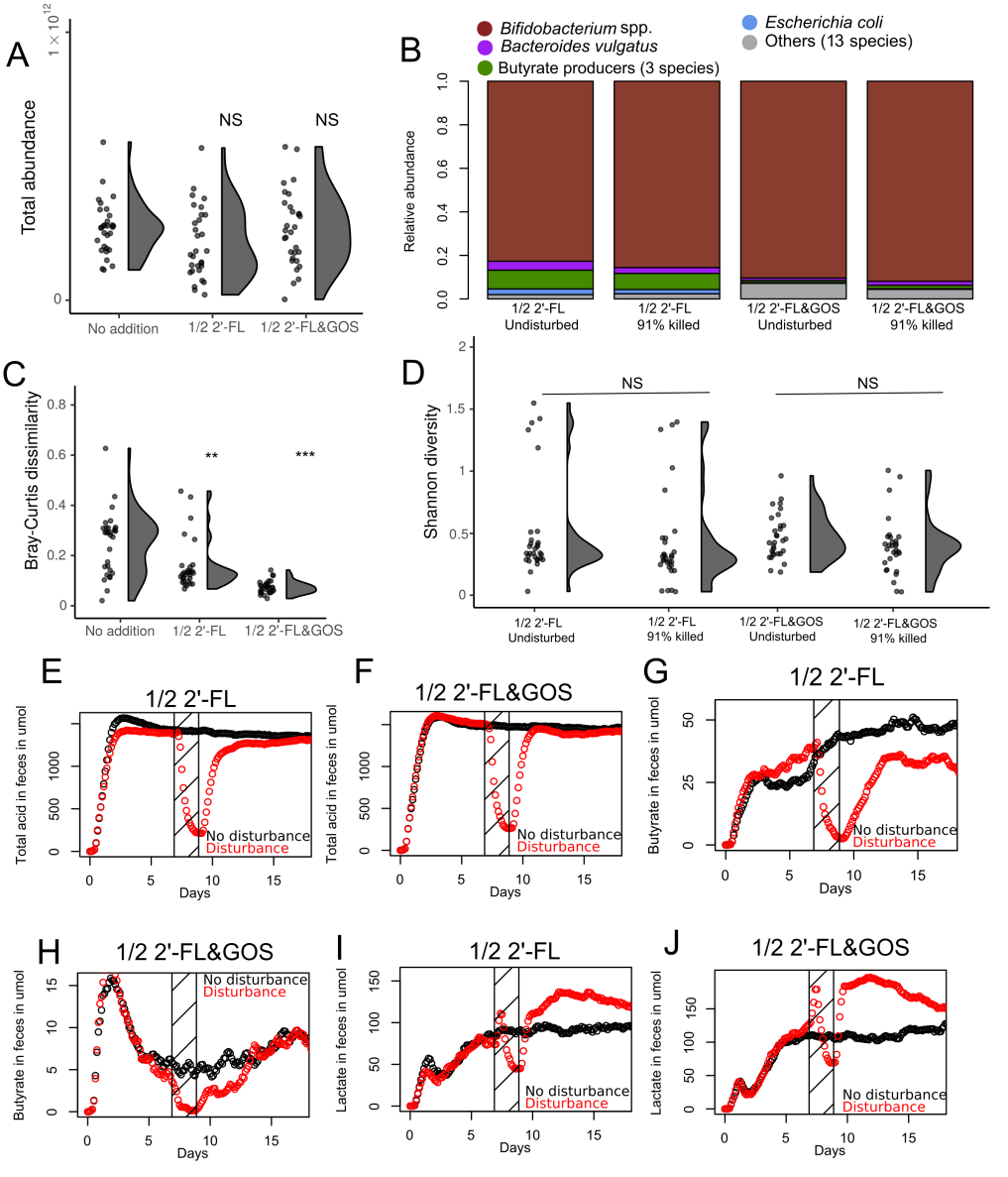


Figure 5.4: Prebiotics shape a more resilient microbiota
For all figures: '1/2 2'-FL' is 105.5 µmol of lactose replaced with 2'-FL, '1/2 2'-FL&GOS' is 21 µmol of lactose replaced with 2'-FL and 84.5 µmol of lactose replaced with GOS. n=30 per condition. NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$ **(A)** Total bacterial abundance at the start of day 10 per condition in the disturbed simulations. **(B)** Relative abundance of bacterial species on day 42 without disturbance or with a disturbance on the simulated day 8 and 9. **(C)** Bray-Curtis dissimilarity by simulation between the mean relative abundance on day 42 (timestep 20160) of the simulations of A, grouped as in table 5.1, compared to undisturbed simulations with the same nutrition at day 42. **(D)** Shannon diversity for all simulations in each condition of B. **(E-J)** Total acid, butyrate, or lactate leaving the system per 60 timesteps in the model with no disturbance (black) and the model with 91% killed (red). Striped area indicates the duration of the disturbance.

5.3 Discussion

In short, the simulations predicted that (1) an antibiotic disturbance that kills 91% of bacteria increases the abundance of Enterobacteriaceae and decreases the abundance *Bifidobacterium*, (2) any disturbance that kills at least 91% of bacteria decreases diversity, (3) a disturbance that kills 91% of bacteria may reduce production of acids, and in particular butyrate, which may have health effects, (4) prebiotics may create and sustain a more resilient microbiota that returns to the same state more consistently, but such a microbiota is still sensitive to disturbance of acid production. We will address the relation between each of these major predictions and the available *in vitro* and *in vivo* data in turn, and suggest some further simulations and analyses that may provide more insight.

Firstly, a reduction in the abundance of *Bifidobacterium* and increase in the abundance of *E. coli* after antibiotics are widely reported [57, 58, 264, 52]. The model also predicts this, but only for the disturbance that kills 91% (Fig. 5.1E). We hypothesise that the disturbance that killed 91% increased the relative abundance of *E. coli* and decreased *Bifidobacterium* because it interrupts and delays succession. Early in the model simulations, there is a higher abundance of *E. coli* (Fig. S7). *E. coli* is typically replaced by other species, such as *Bifidobacterium*, at later timepoints of the model (Fig. 5.1E). This same succession takes place in infants [32, 31]. An interruption of this succession in our model may then lead to a relatively large share of Enterobacteriaceae remaining in the gut compared to controls of the same simulated age. A delay in succession, and so a longer *E. coli*-dominated phase, has also been hypothesised to explain some of the increased *E. coli* observed in infants *in vivo* after an antibiotic disturbance [55]. Whether the model also predicts that this mechanism is important can be tested by adding an antibiotic disturbance at different stages of succession. Disturbances could be added at either a very early *E. coli*-dominated stage or at a stable state with low *E. coli* abundance, which the model typically reaches after several simulated weeks. This testing would reveal whether the increase in *E. coli* in the model depended on the succession stage during which the disturbance was applied. We further hypothesise that in our model weaker disturbances do not disturb the microbiota sufficiently to alter the course of the succession. Heavier disturbances may cause the population to go extinct, which resets the system and does not preserve the relatively high *E. coli* abundance. This would essentially re-initialize the system, but without oxygen. A re-initialisation would explain the increased abundance of the

strictly anaerobic butyrate producers in the model with the disturbance that killed 99% (Fig. 5.1E). Interactions with the level of oxygen present in the gut could also be modelled. We expect oxygen to still be abundant in the infant gut if the disturbance occurs in the first days after birth [103], and the model has previously predicted that variation in oxygen explains much of the variation in succession timing observed in the infant gut microbiota [103].

Secondly, a lower microbial Shannon diversity due to antibiotics is widely reported in infants [57, 58, 270]. The model reproduced the lowered diversity due to antibiotics with every disturbance that killed at least 91% and did not include prebiotics (Fig. 5.2D). The reduced diversity due to a disturbance has also been linked to a reduced stability of the microbiota [270], which the model also predicted (Fig. 5.2E-H). We found that the reduced diversity could be attributed to a lower evenness, not to the complete extinction of species. However, when we eliminated less common species from the data the lowered diversity was maintained, while the lower evenness was not, indicating that a lower evenness does not fully explain the lower diversity.

On a metabolic level the model predicted a reduced acid output during and after disturbance (Fig. 5.3). Such a reduction in acidity due to antibiotics has not been reported in infants, but it has in adults and mice [271, 272]. The metabolic effect of 2'-FL and GOS supplementation on an infant gut microbiota disturbed by antibiotics has been previously studied in an *in vitro* system [263]. This *in vitro* model predicted a positive effect of 2'-FL on butyrate production, and a lowered SCFA output with antibiotics, confirming predictions of our model. However, the *in vitro* model also predicted an increase in lactate concentration due to antibiotics in the absence of 2'-FL and GOS [263], while our model predicted a decrease in lactate concentration in all conditions except those with 2'-FL. It is unclear why this discrepancy is present. Further analysis of lactate production and consumption in our model may reveal what factors are responsible for the incorrect prediction. The authors indicate that *Enterococcus*-driven lactate production may explain the increase in lactate that they observed in their *in vitro* model[263]. *Enterococcus* was included in our bacterial composition, but did not reach a high abundance. Further analysis of the metabolism of *Enterococcus* in our model may reveal where it differs from the *in vitro* model.

Further analysis and simulations could also be performed to reveal why prebiotics cause a more consistent recovery in the model. The prebiotics we used caused *Bifidobacterium* species to become dominant in the model (Fig. 5.4B), while without prebiotics *E. coli* and *Bacteroides vulgatus* were also abundant (Fig. 5.1E). This

effect is also commonly seen *in vivo* [53]. We hypothesise that this stimulation of *Bifidobacterium* over other species causes the microbiota to consistently return to the same *Bifidobacterium*-dominated baseline in the model. As *Bifidobacterium* is associated with improved infant health [18], this would be a positive outcome for the infant. Whether the stimulation of *Bifidobacterium* by prebiotics is the crucial factor in determining the consistency of our outcomes could be tested by repeating the simulations, but only adding the prebiotics after the antibiotic disturbance. We hypothesise that this will lead to a similarly *Bifidobacterium*-dominated microbiota in most simulations, regardless of composition before the disturbance.

More generally, various discrepancies exist between the model predictions and available *in vivo* data. Most importantly, only the disturbance that killed 91% decreased *Bifidobacterium* abundance and increased *E. coli* abundance in the model, and the other disturbances did not. We hypothesised that this happened because the succession away from *E. coli* was interrupted. It is unclear how much of the infant gut microbiota is killed during an antibiotic disturbance, but decreased *Bifidobacterium* and increased Enterobacteriaceae (which includes *E. coli*) abundance are commonly seen in the *in vivo* infant gut microbiota as an effect of antibiotics, and no link to disturbance strength has been reported [57, 59, 60, 58]. In addition, Enterobacteriaceae also become more abundant due to antibiotic treatment in adult humans and adult mice [268], where there is no succession to disrupt. There must be some further factor, unrelated to succession, that causes the increase in Enterobacteriaceae in antibiotic-treated adults. This factor may also cause the increase in Enterobacteriaceae to be much more consistent in infants than what we find. This factor could be a higher resistance of Enterobacteriaceae to antibiotics [266], or Enterobacteriaceae benefiting more than other bacteria from the inflammation-related release of oxygen and nitrate [268, 269]. Neither of these factors are currently included in the model, but they could be included in future versions. Modelling may allow us to explain which of these factors can explain the increase in Enterobacteriaceae observed *in vivo*.

Further simulations should also be performed to account for the possible effects of changes to parameters on the model predictions. In previous work (e.g. [103]) we found a large impact of several of the model parameters on the predictions, such as the amount of oxygen and the diffusion speed of metabolites. These factors may interact with an antibiotic disturbance. Disturbance could also be modelled in different ways in future extensions of the model. For example, injected antibiotics may

affect the gut microbiota from the gut wall, while antibiotics taken orally might have a stronger effect in the lumen. Each of these effects could be modelled to have spatially different effects in our model. In addition, our modelling of the disturbance as an increased death rate is a large simplification of the effect antibiotics have on the microbiota. Most antibiotics used in infants interfere with cell wall synthesis or ribosomal function, and so reduce growth, or make growth more energetically expensive [257]. In our model, this may be more appropriately considered as an altered biomass reaction, instead of an increased death rate. This may lead to more accurate predictions of the effect of antibiotics on microbial metabolism. Finally, additional model simulations with variation in the timing, quantity, and type of prebiotic oligosaccharides could provide more insight into the circumstances required for prebiotics to shape or restore the composition of the microbiota. In short, the presented work provides a starting points for making further predictions on the effects of antibiotics, and for the effects of prebiotic supplementation.

5.4 Methods

5.4.1 Model overview

In this study we have extended a multiscale model of the infant gut. The model is based on our earlier microbiota models [122, 103, 229]. The major addition in this study is the inclusion of disturbances of the microbiota. All parameters of the system are listed in table 5.2.

5.4.2 Species composition

The list of species in the model (table 5.1) is based on *in vivo* data from [3], using sheet 2 of their Table S3. We selected the 20 entries with the highest prevalence in vaginally delivered newborns. After removing two duplicate entries we selected a GEM of a species from each genus from [10]. We added GEMs of *Bifidobacterium breve* and *Bifidobacterium bifidum* to represent the diversity of *Bifidobacterium* species in the infant gut [164]. We also added a GEM of the butyrate producer *Roseburia inulinivorans*, as in our previous model [229]. *Roseburia inulinivorans*, *Eubacterium hallii*, and *Clostridium butyricum* are combined and listed as "Butyrate producers" in the visualisations.

5.4.3 Changes to GEMs

To improve the metabolic predictions of the model we applied a number of changes to the GEMs compared to the version in the AGORA database, as in our previous studies [103, 229]. No further changes were made after those in chapter 4. A full list of changed and added reactions is in table S1.

5.4.4 FBA approach

We continued to use the same FBA approach as in previous versions of the model [103, 229] and earlier chapters. This is a modified version of flux balance analysis with an enzymatic constraint [13, 100]. Each GEM is first converted to a stoichiometric matrix S. Reversible reactions are then converted to two irreversible reactions, so that flux is always greater than or equal to 0. Reactions identified in the GEM as 'exchange', 'sink', or 'demand' are allowed to take up or deposit metabolites into the environment. Each timestep, all reactions are assumed to be in internal steady state:

$$S \cdot \vec{f} = 0, \tag{5.1}$$

where $\vec{f}$ is a vector of the metabolic fluxes through each reaction in the network, in mol per time unit per population unit.

Table 5.2: Parameters of the model

Parameter	Value	Unit
Lattice side length	2	mm
Width of lattice	225	lattice sites
Height of lattice	8	lattice sites
Timestep	180	seconds
Initial populations	540	average number
New population placement probability	0.00005	per timestep per empty lattice site
Population death probability	0.0075	per timestep per population
Initial size per population	$5 \cdot 10^7$	no. of bacteria
Population size to create new population	$1 \cdot 10^{10}$	no. of bacteria
Maximum population size	$2 \cdot 10^{10}$	no. of bacteria
ATP to grow one cell	$1 \cdot 10^{-15}$	mol
Enzymatic constraint	2	µmol flux per timestep per $1 \cdot 10^{10}$ bacteria
Public goods production rate	2	µmol per nutrient per timestep per $1 \cdot 10^{10}$ bacteria
Nutrient input	211	µmol per nutrient every 60 timesteps
Mucin input	0.5	µmol per timestep
Initial oxygen	0.1	µmol per lattice site
Metabolic advection	2	mm per timestep
Diffusion of metabolites and bacteria	$6.3 \cdot 10^5$	square cm per second

Each reaction that takes up metabolites from the environment F_{in} is constrained by an upper bound F_{ub}, which represents the availability of metabolites from the environment. It is determined as follows:

$$\vec{F_{in}} \leq \vec{F}_{ub}, \tag{5.2}$$

where $\vec{F_{in}}$ is a vector of fluxes between the environment and the bacterial population. $\vec{F}_{ub}$ is a vector of upper bounds on these fluxes. $\vec{F}_{ub}$ is set dynamically at each timestep t by the spatial environment at each lattice site $\vec{x}$:

$$\vec{F}_{ub}(\vec{x}, t) = \frac{\vec{c}(\vec{x}, t)}{B(\vec{x}, t)}, \tag{5.3}$$

where $\vec{c}$ is a vector of all metabolite concentrations in mol per lattice site, $\vec{x}$ is the location and $B(\vec{x}, t)$ is the size of the local bacterial population. $\vec{F}_{ub}$ is set to 0 for any metabolite that is digested outside the cell (see section 'extracellular metabolism'). The total flux in each FBA solution is constrained by the enzymatic constraint a, in mol per time unit per population unit:

$$\sum \vec{f} \leq a. \tag{5.4}$$

Given these constraints, FBA identifies the solution that optimizes the objective function, ATP production. The solution consists of a set of input and output exchange fluxes $\vec{F_{in}}(\vec{x}, t)$ and $\vec{F_{out}}(\vec{x}, t)$, and a growth rate $g(\vec{x}, t)$. The exchange fluxes are taken as the derivatives of a set of partial-differential equations to model the exchange of metabolites with the environment. The size of the population increases proportionally to the growth rate in the FBA solution. Populations above $2 \cdot 10^{10}$ bacteria do not perform metabolism, to mimic quiescence at high densities.

Diffusion is applied to the metabolite concentrations on each lattice site at each timestep to represent mixing by colonic contractions. Metabolic diffusion is applied twice during each timestep. Each time it is applied, 14.25% of each metabolite diffuses from each lattice site to each of the four nearest neighbours. This causes a net diffusion each timestep of $6.3 \cdot 10^5$ cm^2/s. Metabolites are also added and removed by bacterial populations as a result of the FBA solutions. The change in concentration per lattice site is thus determined as follows:

$$\frac{d\vec{c}(\vec{x},t)}{dt} = \vec{F_{out}}(\vec{x},t)B(\vec{x},t) - \vec{F_{in}}(\vec{x},t)B(\vec{x},t) + \frac{D}{L^2}\sum_{\vec{i}\in \mathrm{NB}(\vec{x})}\left(\vec{c}(\vec{i},t) - \vec{c}(\vec{x},t)\right), \quad (5.5)$$

where $\vec{F_{out}}(\vec{x},t)$ is a vector of fluxes from the bacterial populations to the environment, in mol per time unit per population unit, and D is the diffusion constant, L is the lattice side length, and $NB(\vec{x})$ are the four nearest neighbours.

To represent advection all metabolites except oxygen are moved distally by one lattice site every timestep. Metabolites at the most distal column of the lattice, the end of the colon, are removed from the system at each timestep. This represents the simulated feces.

5.4.5 Nutrient input

Mucin is added each timestep in equal concentration to the lowest row of the model, to represent secreted mucin. The mucin struture we use is identified in the VMH database as MGlcn23_rl. Nutrients representing inflow from the small intestine are inserted into the first six columns of lattice sites every 60 timesteps, representing three hours. Three hours is a realistic feeding interval for newborn infants [175]. Food intake contains 211 µmol of lactose by default, a concentration in line with human milk [66], assuming 98% host uptake of carbohydrates before reaching the colon [131]. In some simulations additional lactose, GOS, or 2'-FL is added. GOS is inserted as separate fractions of DP3, DP4, or DP5 based on analysis of the composition of Vivinal-GOS [206]. 64% is DP3, 28% is DP4 and 8% is DP5. Water is provided in unlimited quantities.

5.4.6 Extracellular metabolism

Extracellular metabolism is handled through a system separate from FBA in the model. Mucins are always digested extracellularly, as are GOS chains of four or five sugars. *Bifidobacterium bifidum* also digests 2'-FL and GOS chains of length three extracellularly. The system uses the set of extracellular reactions from the GEM. Extracellularly digested oligosaccharides are excluded from FBA. Instead, the model directly applies each extracellular reaction at a rate of 2 µmol per $1 \cdot 10^{10}$ local population per timestep. This alters the set of input exchange fluxes $\vec{F_{in}}$ and

output exchange fluxes $\vec{F_{out}}$, as used in eq. 5.5. If multiple reactions can apply to a substrate, and insufficient substrate is available for each reaction to apply fully, each reaction is applied to half the remaining substrate. No substrate has more than two reactions associated with it in the model, so breakdown can never exceed the available substrate:

$$\vec{F_e}(\vec{x},t) \leq \frac{\vec{c}(\vec{x},t)}{B(\vec{x},t)}. \tag{5.6}$$

All oligosaccharide-related reactions are listed in S1 table. All mucin breakdown reactions are listed in S2 table.

5.4.7 Population dynamics

We initialize the model by giving each lattice site a population of $5 \cdot 10^7$ bacteria of a single random species at a probability of 0.3. Each timestep, each population solves the FBA problem based on its GEM, the enzymatic constraint a, its current population size $B(\vec{x},t)$ and the local concentrations of metabolites $\vec{c}(\vec{x},t)$. The outcome is applied to the environment (eq. 5.5) and the growth rate $g(\vec{x},t)$ is applied to the local population size:

$$\frac{dB(\vec{x},t)}{dt} = B(\vec{x},t)g(\vec{x},t). \tag{5.7}$$

New bacterial populations can be created in two ways. (1) each population of at least $1 \cdot 10^{10}$ bacteria creates a new population of the same species in an adjacent empty lattice site. Half the population size is transferred to the new population. (2) Each empty lattice site has a probability of 0.00005 to get a new population of $5 \cdot 10^7$ bacteria each timestep. All species have an equal probability to be selected.

Bacterial populations are mixed by swapping population contents between lattice sites during each timestep. We use an algorithm inspired by Kawasaki dynamics [176] to mix the bacterial populations. Each site is addressed in a random order, and the bacterial population size $B(\vec{x},t)$ and the GEM are swapped with a site randomly selected from the Moore neighbourhood. The swap only occurs if both the origin and destination site have not swapped in this timestep. With this mixing method the diffusion constant of the bacterial populations is $6.3 \cdot 10^5 cm^2/s$, equal to that of the metabolites. Bacterial populations at the most distal column, i.e. at the exit of the colon, are removed from the system. Populations are also removed from the system at a probability of 0.0075 per population per timestep. Disturbances are implemented

in the model by increasing the probability for a population to be removed by a factor K on days 8 and 9 of the model. We use values for K of 1.75, 2, 2.25, and 2.5. These correspond with a population approximately 81%, 91%, 97%, and 99% lower than in the undisturbed simulations immediately after the disturbance period (Fig. 5.1C).

5.4.8 Analysis

The model records the the size, species, location, and important exchange fluxes $\vec{F_{in}}(\vec{x},t)$ and $\vec{F_{out}}(\vec{x},t)$ for each population. In addition, the metabolites and nutrients leaving the system distally were recorded every timestep, and summed over every 60 timesteps for visualisations. Euclidean distance between relative abundances was calculated as follows:

$$d(\vec{p},\vec{q}) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}, \tag{5.8}$$

where d is the distance between the vectors $\vec{p}$ and $\vec{q}$. $\vec{p}$ and $\vec{q}$ are vectors of the bacterial abundance, either grouped by species or as in table 5.1. This distance metric indicates how far away each set of simulations is from the undisturbed endpoint.

5.4.9 Implementation details

The model is implemented in C++ 11 with libSBML 5.18.0 for C++ to load GEMs and the GNU Linear Programming Kit 4.65 (GLPK) to solve the FBA problems. Random numbers were generated with Knuth's subtractive random number generator algorithm [256]. Diffusion of metabolites was implemented using the Forward Euler method. The model is based on our own earlier models of the gut microbiota [122, 103, 229]. GEMs are sourced from the May 2019 update of AGORA, from the Virtual Metabolic Human Project website (vmh.life). We used Python 3.6 to extract thermodynamic data from the eQuilibrator API (December 2018 update) [173]. All p-values were calculated with R 4.2.2. Unless noted otherwise p-values were calculated using the Mann-Whitney test ('wilcox.test') from the 'stats' package. ANOVA tests used the 'aov' function from the 'stats' package. Principal component analysis was done with the 'prcomp' function from the 'stats' package. Shannon diversity was calculated using the function 'diversity' from the 'vegan' package. Sheldon's index was calculated using the 'sheldon' function from the 'seqtime' package [273]. Model

screenshots were made using the libpng16 and pngwriter libraries. Other visualisations were performed with R 4.2.2. Raincloud visualisations used a modified version of the Raincloud plots library for R [216].

5.5 Supplemental material

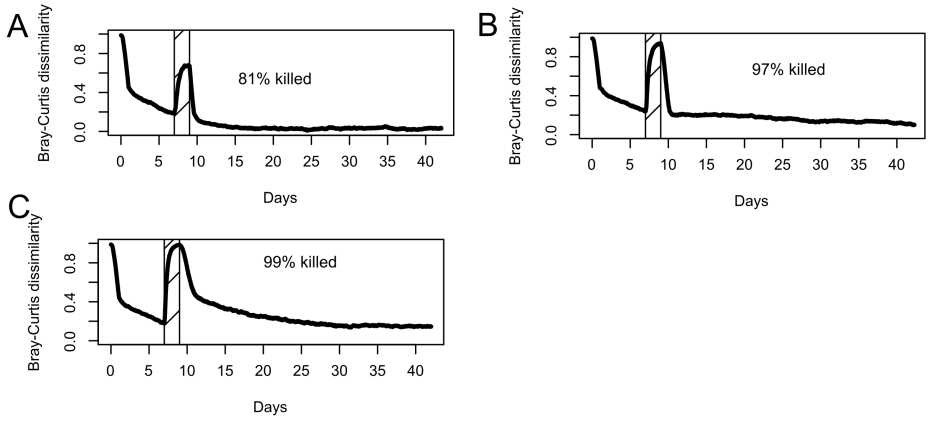
Available from
`drive.google.com/drive/folders/11pW77mica7ugorC0Gg8MwCaaniS9PWrz`

S1 Table

S1table.csv
A table of changes made to the AGORA models as a .csv file.

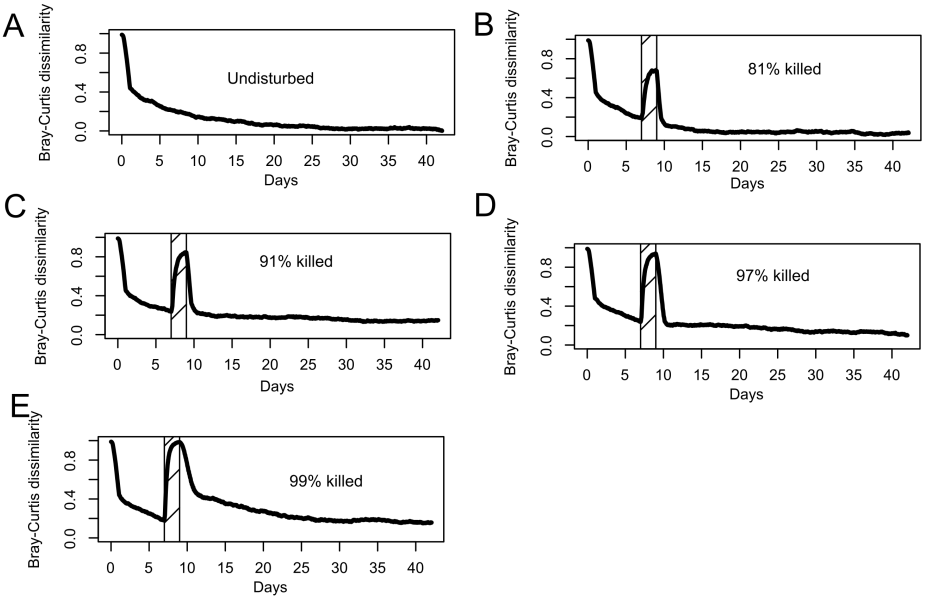
S2 Table

S2table.csv
A table of mucin reactions used for public goods metabolism as a .csv file.



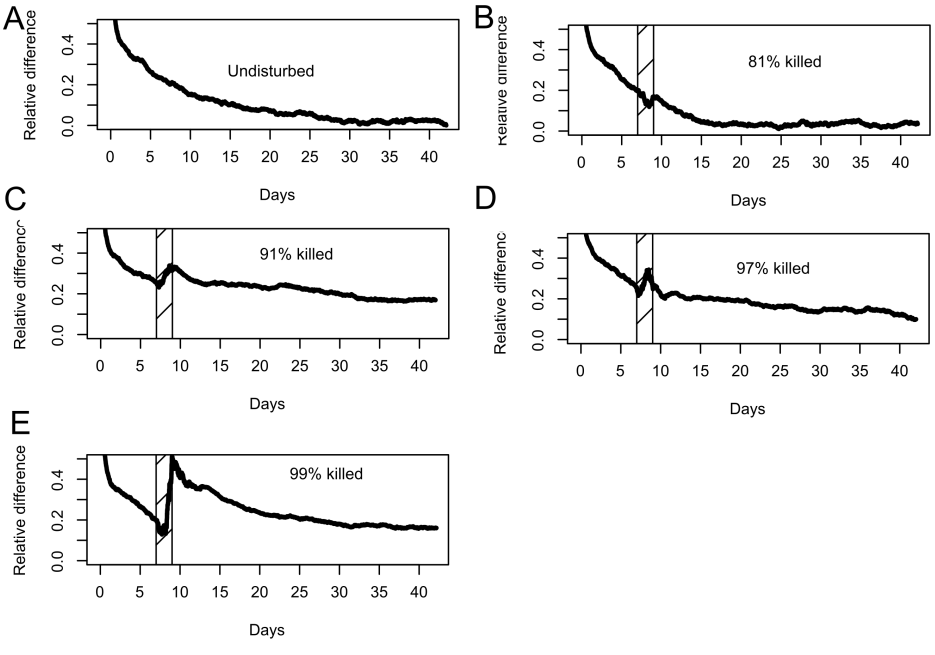
S1 Figure

(A-C) Bray-Curtis dissimilarity between the mean relative abundance per species grouped as in table 5.1 of the conditions that killed 81%, 97%, and 99%, compared to the mean relative abundance of the undisturbed simulations at day 42 (timestep 20160). Striped area indicates the duration of the disturbance. n=30 per condition.



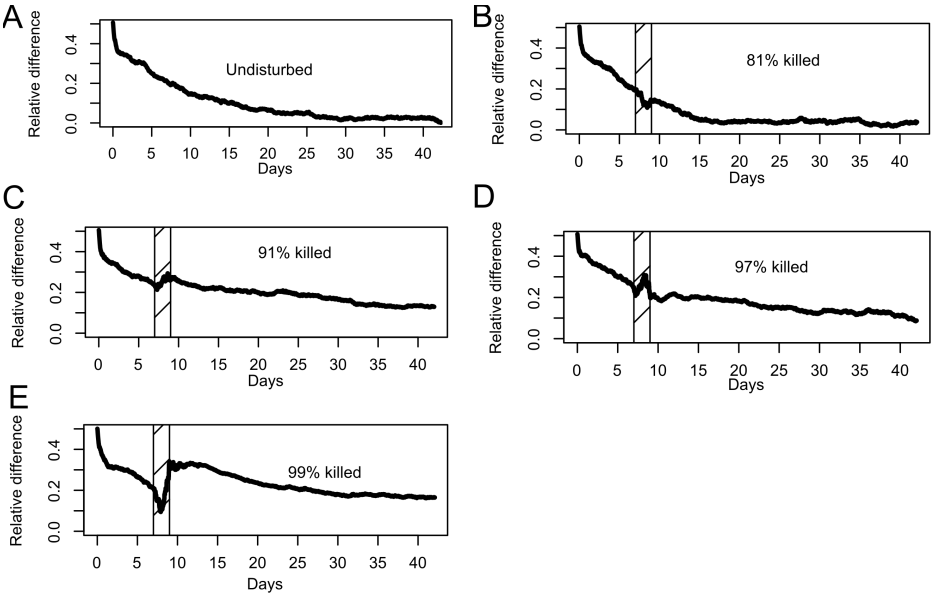
S2 Figure

(A-E) Bray-Curtis dissimilarity between the mean relative abundance per species (ungrouped) of each condition of Fig. 5.2 compared to the mean relative abundance per species (ungrouped) of the undisturbed simulations at day 42 (timestep 20160). Striped area indicates the duration of the disturbance. n=30 per condition.



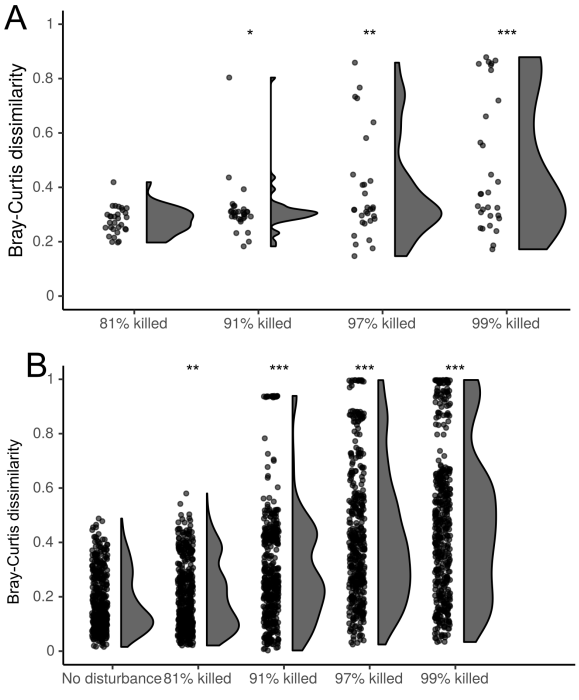
S3 Figure

(A-E) Euclidean distance between the mean relative abundance per species grouped as in table 5.1 of the undisturbed conditions and the conditions that killed 81%, 91%, 97%, and 99%, compared to the mean relative abundance of the undisturbed simulations at day 42 (timestep 20160). Striped area indicates the duration of the disturbance. n=30 per condition.



S4 Figure

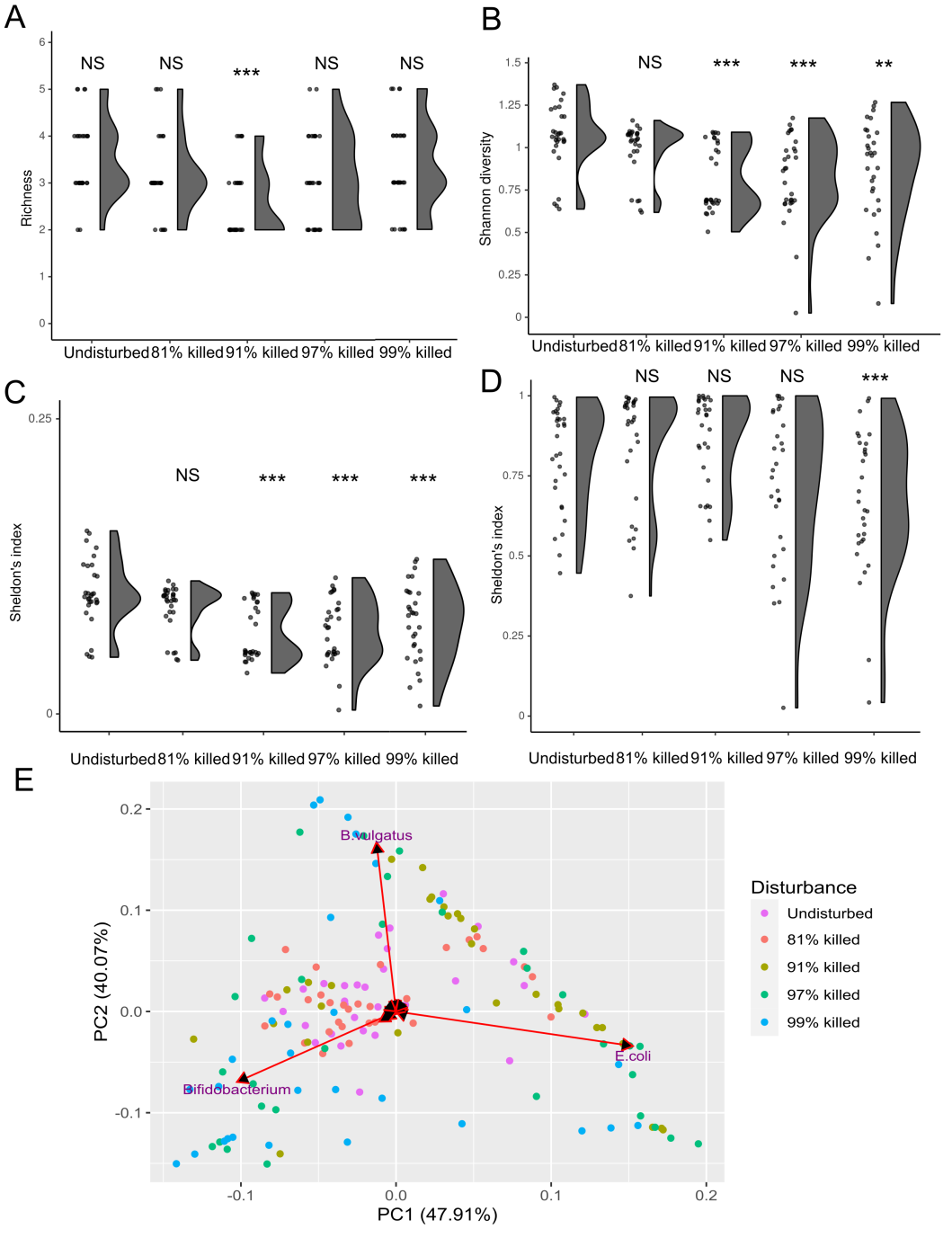
(A-E) Euclidean distance between the mean relative abundance per species (ungrouped) of each condition compared to the mean relative abundance per species (ungrouped) of the undisturbed simulations at day 42 (timestep 20160). Striped area indicates the duration of the disturbance. n=30 per condition.



S5 Figure

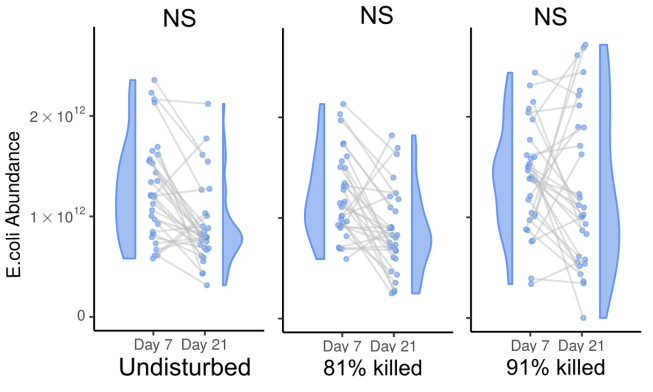
(A) Bray-Curtis dissimilarity by simulation between the mean relative abundance per species (ungrouped) of each condition in Fig. 1D at day 42 (timestep 20160), compared to the mean relative abundance of the undisturbed simulations at day 42. n = 30 per condition (B) Bray-Curtis dissimilarity of the grouped abundance at day 42 (timestep 20160) of each simulation compared to each other simulation within each condition. n=30 per condition, leading to 450 unique comparisons for each condition. NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$

5.5. Supplemental material



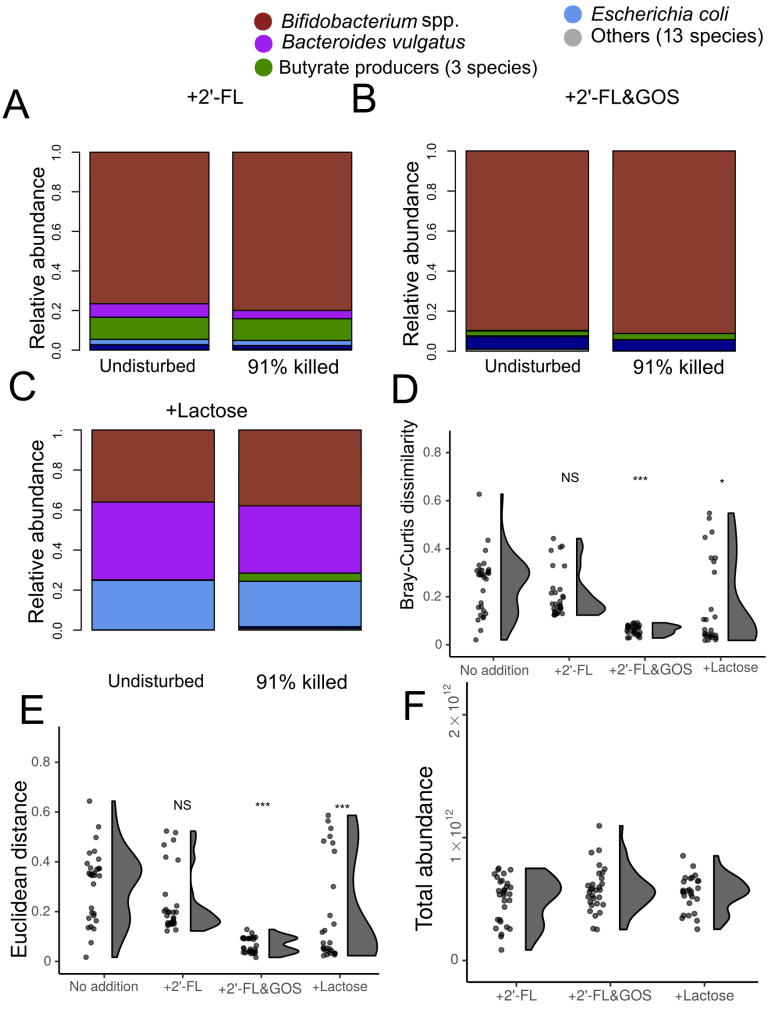
S6 Figure

(A) Richness for all simulations of Fig.5.2 in each condition at day 42 (timestep 20160), with a threshold of 10^{10} bacteria for each species to be counted. (B) Shannon diversity for all simulations of Fig.5.2 in each condition at day 42 (timestep 20160), with a threshold of 10^{10} bacteria for each species to be counted. (C) Sheldon's index for all simulations of Fig.5.2 at day 42 (timestep 20160) (D) Sheldon's index for all simulations of Fig.5.2 at day 42 (timestep 20160), where we set the abundance of each species at 0 if it was below 10^{10} bacteria. (E) Principal component analysis for all simulations of Fig.5.2 at day 42, using the abundance of each bacterial genus. Arrows displayed for each genus, labels included for the largest three. NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$



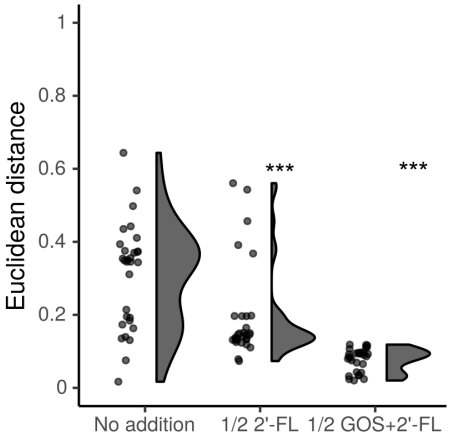
S7 Figure

Absolute abundance of *E. coli* per simulation at the end of day 7 and day 21 for the first three conditions of Fig. 1D. Points from the same simulation are connected with a line. Spearman correlation between day 7 and day 21 indicated where significant. n=30 per condition. NS= Not significant.


Bifidobacterium spp.
Bacteroides vulgatus
Butyrate producers (3 species)
Escherichia coli
Others (13 species)
A
+2'-FL
B
+2'-FL&GOS
Relative abundance
Undisturbed
91% killed
C
+Lactose
D
Bray-Curtis dissimilarity
NS
***
*
No addition
+2'-FL
+2'-FL&GOS
+Lactose
E
Euclidean distance
F
Total abundance

S8 Figure

(A-C) Relative abundance of bacterial species on day 42 without disturbance or with the death rate increased on the simulated day 8 and 9, with the addition of either 211 µmol lactose and 211 µmol FL per 60 timesteps, 211 µmol lactose, 42 µmol 2FL and 169 µmol GOS, or 422 µmol lactose. n=30 per condition. (D) Bray-Curtis dissimilarity by simulation between the mean relative abundance grouped as in table 5.1, at day 42 (timestep 20160) and the mean relative abundance grouped in the same way of the corresponding undisturbed simulations at day 42. All use the disturbance that killed 91%. n=30 per condition (E) Euclidean distance by simulation between the mean relative abundance grouped as in table 5.1, at day 42 (timestep 20160), compared to the mean relative abundance grouped in the same way of the corresponding undisturbed simulations at day 42. All use the disturbance that killed 91%. n=30 per condition (F) Total bacterial abundance at the start of day 10 per condition in the simulations with disturbance of A-C. n=30 per condition. NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$



S9 Figure

Euclidean distance by simulation between the mean relative abundance per species grouped as in table 5.1 of each condition in Fig. 5.4A at day 42 (timestep 20160), compared to the mean relative abundance of the undisturbed simulations at day 42. n = 30 per condition. NS: Not significant, *: $p<0.05$, **:$p<0.01$, ***:$p<0.001$

5.6 Contributions

J.M.W.G., and R.M.H.M acquired funding. D.M.V., J.M.W.G., E.L., and R.M.H.M. conceived and planned the simulations. D.M.V. wrote software used for the simulations. D.M.V. performed the simulations and analyzed the data. E.L., J.M.W.G., and R.M.H.M contributed to the interpretation of the results. J.M.W.G., and R.M.H.M. supervised the project. D.M.V. drafted the manuscript. D.M.V., E.L., J.M.W.G. and R.M.H.M. revised and edited the manuscript.

5.7 Acknowledgments

This study was financially supported by FrieslandCampina. E.L., and J.M.W.G. are currently employed by FrieslandCampina. This work was performed using the ALICE compute resources provided by Leiden University.

Chapter 6

Summarizing discussion, other works, and outlook

In this chapter we will first summarize the methods and results described in chapters 2 through 5. We then proceed to analyse the limitations of our approach, the opportunities for improvement, the field of metabolic modelling in general, and finally some future directions for this modelling approach and the study of the infant gut microbiota in general.

6.1 Thesis overview

In **chapter 1** we introduced the infant gut microbiota, its typical compositions, and some of the factors influencing it. We then introduced our modelling approach, constraint-based modelling, and discussed how it can be applied to modelling bacterial metabolism and bacterial interactions.

In **chapter 2** we examined the role of oxygen in the newborn infant gut, and introduced the multiscale mathematical model based on the previous model by Van Hoek and Merks [122]. The model was extended in several ways, and focused much more tightly on a specific *in vivo* situation: the infant gut microbiota in the first three weeks after birth. We showed how an enzymatic constraint was crucial to correctly model how the metabolism of *Bifidobacterium* spp. depends on the concentration of lactose. On the scale of the infant gut community we showed that the initial presence

of oxygen in the infant gut could explain the succession from *Escherichia coli* to *Bifidobacterium* spp. commonly observed in the infant gut microbiota. Furthermore, the model we created provided a suitable basis for further studies in the next chapters.

In **chapter 3** we extended the nutritional input of the model with prebiotic oligosaccharides, which are common in human milk and some infant formulas. We also further refined the model by using the same diffusion speed for nutrients, metabolites, and bacterial populations. Previously, populations diffused at a slightly different speed from nutrients and metabolites, which was not supported by the available data. Further refinements to the genome-scale metabolic models and the selection of species allowed us to better model cross-feeding relationships on the *Bifidobacterium* metabolites lactate and 1,2-PD. The model predicted that several different microbiota compositions can exist that are dominated by *Bifidobacterium*, and that the cross-feeding species present depends on the substrate available to *Bifidobacterium*. The prebiotic oligosaccharide 2'-FL was predicted to lead to a microbiota rich in *Bifidobacterium* with cross-feeding butyrate producers, such as *Anaerobutyricum hallii*. The prebiotic oligosaccharide GOS was also predicted to lead to a *Bifidobacterium*-rich microbiota, but with the non-butyrate producing cross-feeder *Cutibacterium acnes* instead of butyrate producers. This may have health implications, as butyrate is associated with improved infant health.

In **chapter 4** we further extended the model by including the production of mucin by the simulated gut epithelium and the extracellular digestion of both mucins and the prebiotic oligosaccharides 2'-FL and GOS. The model predicted that the decreased mucin consumption capacity observed in the microbiota of breastfed infants may be explained through the selective stimulation of the non-public goods producing *Bifidobacterium* species *Bifidobacterium longum* by 2'-FL. In contrast to the public goods producing *Bifidobacterium bifidum*, *B. longum* cannot consume intestinal mucin. Although its public goods metabolism allows *B. bifidum* to efficiently consume the prebiotic oligosaccharides in human milk as well as mucins, the model predicted that it also becomes vulnerable to having its public goods 'stolen'. Other bacteria can cross-feed on the public goods, mainly sugars, that *B. bifidum* releases into the medium. The model predicted that because *B. longum* can consume the prebiotic oligosaccharides without creating public goods, it is stimulated more in the infant gut microbiota by breastfeeding, and so reduces mucin consumption.

In **chapter 5** we further examined how the infant gut microbiota may form and reform communities by introducing disturbances into the model. These disturbances

kill much of the microbiota. The microbiota then slowly recovers to a state similar to, but often distinct from, that found in undisturbed infants. The model reproduced the increase in *E. coli* abundance and decrease in *Bifidobacterium* abundance that are often observed *in vivo*, but only for intermediate disturbance strengths. We hypothesised that this occurs because the disturbance disrupts the succession from *E. coli* to *Bifidobacterium* that we described in chapter 2. The model further predicted that supplementation with prebiotics in the model leads to disturbed simulations returning more consistently to a composition similar to that found in undisturbed simulations. We hypothesized that this occurs because prebiotics consistently lead to a large relative abundance of *Bifidobacterium* in the model. This implies that prebiotics may help shape an infant gut microbiota that recovers more consistently from disturbances.

In short, the model predicted that the development of the infant microbiota is regulated by metabolic factors as follows: the initial presence of oxygen stimulates a high abundance of Enterobacteriaceae, such as *E. coli*, with an aerobic metabolism. As oxygen is depleted, other species that grow faster anaerobically on lactose, particularly *Bifidobacterium*, become more abundant. The model further predicted that the prebiotics GOS and 2'-FL affect the development of the infant microbiota by steering the development of the microbial ecosystem to a high abundance of *Bifidobacterium*, particularly *Bifidobacterium longum*. A microbiota dominated by *B. longum*, in turn, consumes less mucin in the model. Through cross-feeding, 2'-FL also stimulates butyrate producers. Taken together, the model predicted a crucial role for oxygen and nutrition in determining the relative abundance of bacterial species. The abundance of species determines the nutrients consumed and metabolites produced in the gut, including metabolites that may have health effects.

In this chapter we will first discuss some limitations and simplifications of the model that should be emphasised, and that may be improved in future versions of this model. We will then discuss other related models and their relation to our modelling approach. Finally, we will discuss several interesting topics and questions that could be investigated with new models.

6.2 Model limitations

6.2.1 The difficulties of modelling metabolism

The model predictions are limited in several ways. We will focus on three aspects: metabolism, spatial variation, and microbial diversity. The model predictions are firstly limited in the quality of their metabolic predictions by their limited scope. The model only included a limited set of genome-scale metabolic models (GEMs), used only their carbon metabolism, and only in the context of the gut of the non-premature vaginally born infant. This means that the model cannot currently make predictions for any situations outside of this context, such as on the effect of different protein concentrations in nutrition, or the effect of disturbances on infants born by cesarean section. Even within the scope of the model it does not capture the full extent of carbon metabolism. As described in previous chapters, we use flux balance analysis (FBA) to predict metabolism. For FBA, an objective function needs to be chosen to serve as a proxy for biomass production. We have previously discussed the effects of our assumptions related to the objective function in chapter 4, but the core issue bears repeating: FBA assumes that bacterial metabolism is optimised to fulfill a specific reaction that represents biomass [13, 240]. However, the evolution of bacterial regulation has not been influenced solely by how efficiently biomass can be produced in isolation. Bacteria also, for example, produce bacteriocins [274], spores [20], adhesion proteins [163], or proteins that interact with the host immune system [17]. Such factors may influence the internal distribution of fluxes, and so may also influence the uptake of nutrients and production of metabolites. It is difficult to assess the impact of this omission on our model predictions. The uptake and production of metabolites is vital to the ecological interactions, such as competition and cross-feeding, that occur in the model, so even small changes to bacterial metabolism may have far-reaching effects. By producing bacteriocins some bacteria may even kill other bacteria [275], which may lead to a higher diversity in the infant gut [276, 274]. Furthermore, oxygen tolerance varies between species, and the presence of oxygen in the newborn infant gut requires bacteria to produce protective enzymes to mitigate oxygen toxicity [277, 278]. The GEMs used in this thesis included oxygen metabolism for relevant species, but did not include the necessity for species that do not use oxygen to protect themselves from oxygen in the environment. *In vitro*, the production of enzymes differs between species and even within subspecies for *Bifidobacterium*

[278], which means that this variation may be important to include in future version of the model. If additional factors were known exactly they could be incorporated in FBA through, for example, minimum fluxes through certain bacteriocin-producing reactions, minimum fluxes through oxygen-tolerating reactions, or techniques such as resource balance analysis [279]. This is difficult in practice, as the circumstances and extent of production in an *in vivo* context are unknown. Nonetheless, their incorporation could improve the quality of metabolic predictions.

This brings us to the next issue, which is that it is generally not known whether FBA predictions are correct, because no *in vitro* data is available to verify them. For some bacteria, the factors influencing metabolism are well documented, and the response to different nutrients is well characterised. This allowed us to verify, in chapter 2, that the predictions that the model makes for the ratio of metabolites produced by *Bifidobacterium longum* from lactose are correct, and that the predictions the model makes for how these ratios change based on the abundance of substrate are also correct [90, 103]. However, for most other species this information is not available, and the FBA predictions remain unvalidated. Wider sets of data have been used to validate GEMs [10], but these often focus largely on growth assays, which merely distinguish growth or no growth. This overlooks the effect of correctly predicting bacterial metabolites, which may greatly influence other bacteria in the ecosystem [89]. Even metabolic assays for particular conditions cannot validate whether the metabolites produced by the bacteria under different conditions are correctly predicted. Correct metabolic predictions for one substrate do not necessarily mean that the metabolic prediction for a different substrate is correct, as we saw for the *B. longum* products of 2'-FL digestion in chapter 3. Here, the model correctly reproduced the metabolic switch in *Bifidobacterium* on low vs. high concentrations of lactose, but not the very similar switch on low vs. high concentrations of 2'-FL. Predictions may be improved through incorporating additional constraints on the FBA solutions, such as rate of change constraints [116] or thermodynamic realizability constraints [110]. Work remains to be done on validating other existing predictions, to better estimate how good the overall model predictions fit reality.

6.2.2 Modelling the spatial distribution of bacteria in the human colon

The next limitation to discuss is our handling of space in the model. Most *in vitro* and *in silico* microbial models use well-mixed conditions to simulate the bacterial environment [280]. This allows for consistent results, but may miss much of the variability that occurs in the colon [281]. The colon represents a unique spatial system, where not only physical effects such as diffusion and flow impact the distribution of bacteria and metabolites, but also active mixing through contractions and peristalsis [282]. In the models used in this thesis, we included a two-dimensional space and bacterial diffusion, but not advection. This led to a spatial separation of primary and secondary consumers in our model, as we described in chapter 2, and as was also observed in earlier work by Van Hoek et al. using a similar spatial approach [122]. This modelling approach is a large simplification, and the model cannot represent or make more complex predictions that take factors such as colonic motility into account. This may influence many of the predictions of the model, as the spatial structure of populations determines whether they can cross-feed, and the model predicts that cross-feeding is crucial for some species.

It is unclear what the distribution and diffusion of bacteria actually is in the infant or adult gut. In the adult gut diffusion has been estimated at values around 3 orders of magnitude higher than what we used in the model [159, 103]. Some spatial *in silico* models of the adult gut have applied advection and similar diffusion values as [159] to bacterial populations, and predict that most bacterial metabolism occurs very close to the entrance of the gut [283], with little spatial separation. In this model bacteria exist throughout the gut, but the high rate of diffusion and advection requires the growth and metabolism of bacteria to occur nearly exclusively in the proximal colon, very close to the entrance from the small intestine.

Another *in silico* model also included estimates of viscosity and active movement of the gut [281]. This model predicted that the effect of any plausible diffusion is negligible compared to the known active movement of the gut, viscosity effects, and an assumed motility of bacteria [281]. They argue that, due to a high viscosity close to the gut wall, a small amount of bacterial motility best explains the persistence of bacteria in the gut. From this, their model predicts that most metabolism occurs in the transverse colon - roughly the middle section [281]. However, metabolism was also present in the other sections. Both of these *in silico* models lack experimental

validation. Additional *in vivo* data would help determine what model best approximates reality. In addition, these studies all model the adult colon - localization of bacteria and bacterial metabolism may be different altogether in the infant gut. Important factors for the spatial distribution of bacteria, such as stool viscosity and colonic motility, are different in infants [284, 162]. A thorough modelling of viscosity and active movement in the infant gut, coupled with *in vivo* data, may provide useful insights into the spatial distribution of metabolism and bacterial abundance specific to the infant gut. These insights could then be used to better estimate what spatial distributions of bacteria, and so what cross-feeding interactions, are plausible in the infant gut.

6.2.3 Modelling more microbial diversity

There are several ways in which the models used in this thesis cannot capture the full breadth of variation that exists between infant gut microbiotas, and how future versions may come closer. This would allow us to provide predictions for the effect of infant gut microbiota in different conditions, and may allow us to answer new questions such as: (1) What is the interaction between nutrition and premature or cesarean birth on the composition of the infant gut microbiota? (2) What is the influence of nutrition when we consider the true diversity of species in the gut, including less abundant bacterial species, Fungi, and Archeae? (3) What is the influence of nutrition on the abundance of specific strains, and how do differences between strains impact the digestion of nutrients? We will discuss each in turn.

Firstly, there is more variability between infant birth conditions than what we model. For example, infants are often born prematurely, or by cesarean section. Premature infants typically have a microbiota with a much higher Enterobacteriaceae or Bacillota abundance and lower *Bifidobacterium* abundance [200, 285]. Cesarean-born infants typically have more *E. coli* and less *Bifidobacterium* and *Bacteroides* [3, 16]. These effects may be related to both an altered colonization by bacteria, as well as altered physiological conditions in the gut. Further modelling could be done to explain why the microbiota develops differently in these infants, and how it can be steered to a healthier composition. The model could provide predictions for the interactions of nutrition with different birth conditions. Nutrition is a factor that may compensate for the typically less health-promoting cesarean section microbiota [50].

Secondly, the model only includes a small subset of all species in the infant gut. In reality, the infant gut microbiota consists of hundreds of species [3], with tens of different species co-existing at the same time [5]. We have selected GEMs of species from the most abundant genera in fecal samples [3, 229], but abundances may be different in the gut microbiota itself [37, 35]. Within the variety of species used, we also limit ourselves to bacterial species. Though their abundance is limited, Fungi and Archaea are also present, and may interact with the bacterial microbiota [200]. Further versions of the model may use more species from these groups, and so provide further predictions and insights into the effect of other factors, such as nutrition, on the abundance of these species.

Finally, genetic variation exists within species. Variation occurs even in genes very crucial to survival in the infant gut, such as genes necessary for the consumption of prebiotic oligosaccharides [286]. This variation can have large repercussions for the overall metabolism of the infant gut microbiota. For example, some infants fed prebiotics do not develop a microbiota that can digest prebiotics, because the microbiota lacks the right genes for prebiotic digestion [32]. Ultimately, these issues come down to the model not incorporating the diversity of bacterial metabolisms that exist in the infant gut, and so missing some of the metabolic diversity. To provide new predictions that apply more broadly to the diversity of infants, future version of the model should also include these metabolisms.

Two new resources for microbial modelling may allow future models to better represent these metabolisms. Firstly, the GEM collection AGORA 2 [287] is greatly expanded compared to the AGORA 1 collection we used in this thesis, and now includes more species and subspecies, as well as more detailed modelling of existing species [287]. For example, where relevant the new models include a separate periplasm compartment. Secondly, the improved thermodynamic database Equilibrator 3 [288] includes support for the dynamic generation of thermodynamic values for any nutrient or metabolite, such as complex mucin structures. This would allow reactions in GEMs to be checked for thermodynamic plausibility, and could potentially guide further refinements of GEMs.

6.3 Related modelling approaches

The work we performed on modelling the infant gut microbiota took place in the context of many other *in silico* approaches to modelling bacterial interactions. A

common question in the study of the gut microbiota is whether interactions are competitive or commensalist [289, 97, 290, 291]. Many approaches exist that focus on different aspects of interactions in the gut microbiota, and aim to answer different kinds of questions around the microbiota. A common approach is to focus on population dynamics, and how these may be influenced by internal and external factors. This is commonly modelled with a set of ordinary differential equations which describe a generalized Lotka-Volterra model [280, 292, 200]. For example, Roa et al. [200] used this approach, and used the absolute abundances of bacteria in the prematurely born infant gut as input. This allowed them to predict the positive and negative effect of bacteria on the abundance of other species, and so predict what relations between bacteria shaped the succession they observed in premature infants. However, their model could not explain how these interactions occurred [200], which is a weakness of Lotka-Volterra models in general [280].

Some other approaches make concrete assumptions or predictions on the means by which populations interact. Interactions may occur in many ways, such as through competition for nutrients, cross-feeding on metabolites [89], antimicrobial toxins produced by bacteria [276] or through phages [293]. Different models altogether focus on the eco-evolutionary aspects of bacterial interactions [294, 295]. By including mutation and selection these models aim to explain how the interactions between species have evolved in a dynamic environment. We will focus here on models that examine interactions between bacteria through a concretely modelled metabolism. Such interactions can be competitive, but through cross-feeding these interactions can also be commensalist or mutualist [89]. Metabolic modelling allows for mechanistic predictions that describe what molecules are important for certain interactions between bacterial population [280]. Many approaches exist to model metabolism [280], but we will focus here on approaches that use genome-scale metabolic model to examine how metabolic interactions can occur, and how these can interact with bacterial abundances to form communities of interacting microbes. We will discuss several important historical models that use genome-scale metabolic networks, as well as a number of state-of-the-art modelling frameworks. We will highlight how these approaches have provided insight into the interactions within the microbiota, and how they have informed our modelling choices. A persistent problem in the creation of useful models with genome-scale metabolic models is selecting realistic uptake constraints and objective functions [240]. These choices, as we will discuss, also influence the degree to which models predict competitive and cooperative interactions. The models we discuss are summarized in table 6.1.

Table 6.1: Important historical and current approaches for the genome-scale metabolic modelling of microbial communities

Name	Year	Scope	Dynamic	Spatial	Uptake constraints	Objective function	Metabolic model source	Notable aspects or conclusions
Cobra toolbox[296]	2007	Single microbes	No	No	Various	Various	Various	First widely used FBA toolbox
Stolyar et al. approach [115]	2007	Carbon metabolism of two soil bacteria species	No	No	Set manually	Total biomass or biomass of one species	Own models	First multispecies model
Zhuang et al. approach [118]	2011	Carbon metabolism of two soil bacteria species	Yes	No	Set manually, then dynamically	Species biomass	Various	First dynamic multispecies model
OptCom [297]	2012	Soil bacteria and bioreactors	No	No	Set manually	Community biomass	Various, including [115]	First multispecies framework
COMETS (2014) [121]	2014	Two- and three-species *in vitro* consortia	Yes	Yes	Set manually, then dynamically, minor components constant	Population biomass	Various	First dynamic spatial framework
DFBALAB [298, 299]	2016	Bioreactors and wound biofilms	Yes	Yes	Set manually, then dynamically	Species biomass and uptake	Various	Bacterial metabolism differs spatially
Van Hoek and Merks approach [122]	2017	Carbon metabolism of human gut microbiota	Yes	Yes	Set dynamically	Species biomass	[300] with adjustments	Bacteria can evolve spatial differentiation
SteadyCom [131]	2017	9 species from the human gut microbiota	No	No	Randomized	Equal biomass production rate by each species	Various	Bacterial composition influenced by fiber uptake.
BacArena [123]	2017	Many systems, including gut microbiota	Yes	Yes	Set manually, then dynamically	Population biomass	Magnusdottir et al. [10] (AGORA)	Bacterial diversity is increased by a mucin gradient
MAMBO [301]	2018	Various human microbiotas	Yes	No	Randomized	Population biomass	Own models	Predicts metabolomics directly from genomics
Chan et al. approach [124]	2019	5 species from the human gut microbiota	Partial	Yes	Randomized	Population biomass and equal biomass production rate	Various	Bacterial diversity is increased by an oxygen gradient
MICOM [302]	2020	Entire human gut microbiota	No	No	From [303]	Population biomass	Own models	Uses abundances as inputs to calculate growth rates and dilutions
µBialSim [304]	2020	773 species from the human gut microbiota	Yes	No	Flat rate for all, then set dynamically	Population biomass	Magnusdottir et al. [10] (AGORA)	First model to use hundreds of metabolic models dynamically
ACBM [305]	2020	Bioreactors	Yes	Yes	Set manually or from transcriptomics, then dynamically	Population biomass	Various	First model to use three-dimensional space
COMETS (2021) [132]	2021	Many microbial systems, including the human gut microbiota [306]	Yes	Yes	Set manually, then dynamically	Population biomass	Various	COMETS can cover many situations
Our approach	2022-2023	Carbon metabolism of 21 species from the infant gut microbiota	Yes	Yes	Set manually, then dynamically	Population biomass	Magnusdottir et al. [10] (AGORA)	Summarized in this chapter

Early work with genome-scale metabolic models focused on analyzing the behaviour of single species using either steady-state FBA [296] or dynamic FBA [113]. This provided useful insights into microbial metabolism, but could not predict the interactions of species. Stolyar et al. [115] created the first multi-species model to predict and explain the interactions between two soil microbes. In this model the carbon metabolisms of two GEMs were linked, so that they could directly exchange metabolites and shared one objective function. This was a 'community' objective function, which maximised either total biomass or the biomass of one of the two component species. Implicitly, this assumed that the bacterial regulation evolved to be cooperative. Stolyar et al. concluded that the bacteria were mutalistic, that syntrophy influences soil bacteria, and that this multi-species approach can give insights that were not possible with a single-species approach [115]. Later work by Zhuang et al. [118] on soil bacteria for uranium bio-remediation performed dynamic FBA on interacting GEMs, where each GEM represented a single species. They termed this approach DMMM, for 'dynamic multi-species metabolic modelling'. Zhuang et al. used multiple GEMs, each with its own objective function, so that each species optimised its own biomass production in each step of the dynamic FBA. This allowed the species to grow separately, and adjust their relationship over time [118], which in turn led to a competitive relationship[118]. Cooperation could still have occurred under these conditions, but is no longer assumed. They concluded that competition can greatly influence bacterial interactions and relative abundances. The model we presented in this thesis used a similar approach by assuming that only population-level objective functions exist, and indeed we also observed both competitive and commensalist relationships, particularly in chapters 3 and 4.

DMMM models were further developed into several frameworks which could model a wide variety of conditions, such as OptCom [297] and COMETS [121]. OptCom used a combination of species-level and community-level objective functions for its FBA, similar to Stolyar et al. The community-level objective functions are necessary in OptCom to model complex systems, such as phototrophic microbial communities. These communities exist in nature, but are not predicted by OptCom without assuming a community-level objective function [297]. This assumption again implies that bacterial behaviour is not entirely selfish, but evolved to limit their own growth in favor of a stable community, which is not necessarily true. In contrast to OptCom, COMETS did not assume a community-level objective function, instead using objective functions unique to each species and calculated per bacterial popu-

lation. This assumes that bacterial behaviour is entirely selfish, and has not evolved to limit growth in favor of stability, which is not necessarily true. COMETS could still represent diverse communities without the community-level objective functions by incorporating spatial diversity [121]. COMETS showed how complex interactions that involve both cooperation and competition can occur when species start in different spatial locations. While in principle extendable to very complex systems, the COMETS publication only included simple systems with up to three species and one colony per species, and notably did not include complex population dynamics. The environment simulated was carbon-limited laboratory medium, which allowed uptake bounds to be set easily. Further work with the COMETS framework tested many pairwise relations between bacteria, and concluded that many competitive, mutualistic, and commensalist interactions could exist in this framework [307]. COMETS has formed the basis for many current approaches, including the one in this thesis, both by using a spatial lattice and by assuming that only population-level objective functions exist. However, it does not yet explore the possibility of complex population dynamics on this lattice.

Work with DFBALAB, another framework that included dynamic FBA and spatial diversity [298], did show how metabolism and species abundances can become varied over one-dimensional space, even when the initial composition of a system is homogeneous. Like COMETS, DFBALAB used individual objective functions. DFBALAB went beyond what was shown with COMETS by demonstrating how, through competition and commensalism, spatially distinct ecological niches may form over time. These niches then lead to a self-reinforcing spatial variability [298]. Further work with DFBALAB also showed that this spatial separation may occur in biofilms [299]. We observed similar spontaneous spatial separation in our model, as we highlight in chapter 2.

Van Hoek and Merks developed a spatial dynamic FBA model that focused on the human gut [122], which the model in this thesis is directly based on. The gut is a unique environment that introduces additional complexities in modelling. As we discussed in section 6.2, there are many more unknown factors in the gut, and data is scarce. To simplify this system, Van Hoek and Merks used an objective function that only required ATP generation for growth. This focused the model to be on carbon metabolism, and circumvented the problem of setting unknown metabolite uptake constraints for rare nutrients. Like COMETS, the model calculated FBA solutions per bacterial population. All populations were derived from the same GEM,

but diverged and evolved over the run of the model, representing an overlap with the eco-evolutionary models that we discussed previously. This led to the formation of different 'species', similar to models that are initialised with distinct species, and so still allows for spatial heterogeneity of metabolism. The authors concluded that spatial heterogeneity involving both commensalism and competition can evolve spontaneously in a spatial gut environment, but that spatial disturbance of the population causes this diversity to collapse.

The SteadyCom framework, by Chan et al., also covered the human gut microbiota, but used static genome-scale metabolic models and did not include spatial or temporal variability [131]. It focused on an adult system of nine representative species, each represented by a GEM, from various sources. It used species-specific objective function, but introduced the additional constraint that all species must have an equal growth rate. This forced the solution to represent a steady state microbiota, which corresponds to the usual state of the adult human microbiota [308]. Steadycom did not limit the metabolism of its models to carbon metabolism. Instead, Steadycom uses a random sampling method to determine metabolite uptake bounds for unknown metabolites. When averaged over 1000 sets, this led to population ratios close to those found *in vivo*, but with a low bacterial diversity, as most species go extinct [131]. Further work with Steadycom concluded that the previous assumptions of equal growth could not support a more diverse microbiota [309]. By assuming that bacteria did not always grow optimally, the number of commensalist relationships increased and many more species could become abundant [309]. This had a similar effect as the community objective functions in OptCom. The authors noted that a spatial approach might also lead to sub-optimal growth, and so allow for more diversity [309]. Steadycom was later combined with a spatial and dynamic FBA approach, which indeed allowed it to predict a more varied human gut microbiota without assuming suboptimal growth, and furthermore predict that a spatial gradient of oxygen from the gut wall increased bacterial diversity in the gut microbiota [124]. This is similar to our results in chapter 2, where we describe how a temporal gradient of oxygen leads to a higher bacterial diversity.

The spatial dynamic FBA framework BacArena was similar in construction to COMETS, but focused more on the heterogeneity in bacterial populations by modelling individual bacteria that have different metabolisms [123],. When adapted to model a section of the human gut, the model predicted that a gradient of mucin concentration from the gut wall creates new niches, and may so explain some of the

microbial diversity in the human gut [123], as we also examined in chapter 4. The BacArena framework was also used to predict what nutrients were important for competition and cooperation in this system. Initial uptake constraints were set by determining a set of essential nutrients, and setting uptake constraints for these at 1µM, while all other uptake constraints were set at 0.1µM. These constraints allowed all species to grow in the model, without explicitly trying to model all nutrient concentrations in the gut. BacArena represented an advanced modelling framework that includes spatial and temporal variability, and provides many of the same features as our model, although its features focused on representing bacterial heterogeneity.

The new 2021 version of the COMETS modelling framework [132] provides further interesting opportunities for standardizing metabolic modelling. This new version extends the framework to also include extracellular enzymes and eco-evolutionary interactions. There is also potential to include other complex bacterial interactions, such as those through toxins and phages [276, 293]. Taken together, it should be possible to translate the model presented in this thesis fully into the new COMETS framework. Testable hypothesis for many of the open questions could potentially be generated with this approach. As-is, it is very difficult to compare the outcome of our studies with these other models, or the outcomes of these models with each other. As we have discussed, several models exist that broadly agree in their modelling approach with the model in this thesis: the gut is modelled with populations of bacteria on a lattice, and metabolism is modelled with dynamic FBA using population-level objective functions [123, 132, 122]. However, there are many small differences in model assumptions, such as how to treat the availability of rare nutrients, whether bacterial advection should be included, or whether the microbiota in the lumen should be considered to be in steady state. Our model is also unique in that it focuses on the infant gut microbiota, and the unique microbiota and physiology associated with infants. Nonetheless, to allow for the better comparison of models, it would be useful for the future development of the field of microbiota modelling if more studies used the same framework.

Newer frameworks have also become more complex in various other ways. For example, the spatial dynamic ACBM framework includes three-dimensional space, in contrast to the earlier models we discussed that all included only two dimensions [305]. This allows ACBM to make more quantitatively accurate predictions of growth rates and nutrient consumption in bioreactors with two-species communities [305]. The µbialsim framework used modern computing power to model the entirety of the

773 species then available from the AGORA database in a single simulation environment [304]. This allowed the framework to model a very complex community, with many different interactions between species [304]. Some newer models also use GEMs differently from the methods we have discussed previously. For example, MAMBO does not use existing genome-scale metabolic models, but takes the meta-genomics of fecal samples as input [301]. It dynamically reconstructs GEMs, and then combines these with bacterial abundances to predict the metabolism of a microbiota, including competitive interactions [301]. This may be particularly useful in the study of microbiotas that have not been well-characterised, as no existing genome-scale metabolic models are necessary. Similarly, the MICOM pipeline takes bacterial abundances as input, but MICOM combines these with existing genome-scale metabolic models to calculate what growth rates and dilutions can plausibly form the observed abundances [302]. MICOM then uses these growth rates as constraints when predicting bacterial metabolism. By working this way MICOM evades the problem of setting unknown uptake bounds, at the cost of requiring bacterial abundances and being unable to predict the development of a microbiota over time [302].

Overall, the modelling of the gut microbiota using genome-scale metabolic methods has given several useful insights. Firstly, it shows the usefulness of taking a multi-scale approach, by considering temporal or spatial variability. Models that incorporate this variability can model more complex systems, without including far-reaching assumptions about community cooperation. Secondly, it has shown that many different ecological interactions are plausible in the gut microbiota, from competition to mutualism, but also that assumptions greatly matter in what kind of interactions are predicted.

6.4 Future directions

We will now highlight how the model presented in this thesis, and its metabolic modelling approach in general, may be useful to the field in the future. Firstly, a closer integration with the eco-evolutionary methods we discussed previously may allow for evolutionary questions to be answered. For example, it is unclear how the mutualistic relationship between humans and *Bifidobacterium* came into existence [70]. In humans, milk oligosaccharides specifically stimulate *Bifidobacterium*, which are beneficial to the host, and this is often used as an example of co-evolution [70, 220]. Milk oligosaccharides were an early development in the evolution of mammals, and

are present in nearly all mammals, including marsupials, monotremes, and all orders of placental mammals [68]. However, *Bifidobacterium* are not typically dominant in the gut of non-human newborn mammals [310, 311, 312, 313, 314]. For example, in newborn Asian elephants (*Elephas maximus*) bacteria of the order Enterobacteriales are initially dominant, as in human infants, but these are then replaced by Lactobacillales and Bacteroidales [313], not *Bifidobacterium* as in human infants. Asian elephant milk has been characterised, and contains a similar composition of oligosaccharides as human milk [68], at a higher concentration [315]. It is unclear what occurs in the guts of these animals, but the milk oligosaccharides may be digested by other bacteria, such as *Bacteroides* or *Lactobacillus* species [313, 310]. It is possible that these bacteria have developed their own mutualistic relationships with non-human animals. Such relationships may be even more complex in marsupials and monotremes, whose milk contains little lactose, but whose milk oligosaccharides are also directly consumed by the newborn [69, 316]. Modelling using eco-evolutionary methods [295, 122, 294] may provide insights into the evolution of unique newborn gut microbiotas, and the evolution of the link between different bacteria and milk oligosaccharides. For example, van Dijk et al. already created a model that predicted that serial transfer causes populations of interacting microbes to either focus on a high yield or a high growth rate [295]. Such a serial transfer between hosts must be an essential part of an evolutionary model of microbiota development. Baijic et al. used a GEM of *E. coli* to examine in detail how its metabolism may evolve and how this may, in turn, influence the environment [294]. Such an examination could also be performed on the GEMS of other common gut bacteria, such as *Bacteroides*, *Bifidobacterium* and *Lactobacillus* species, to potentially uncover mechanisms of co-evolution.

Secondly, a further refinement of the modelling approach may allow for a better understanding of the observed compositions of the human infant gut microbiota. "Everything is everywhere, but the environment selects" is a common tenet in microbiological research [317]. We now know that in the infant gut not everything is always present, but it is clear that the infant gut is exposed to many more bacterial species than the few that become abundant [200, 54]. Thus, selection by the environment is crucial in determining the composition of the gut microbiota. However, for many aspects we do not yet know why the environment selects in the way it does. Traditionally, studies in this field have been mostly observational. These studies have consisted of analyses of fecal samples of infants born and raised under

different conditions. With metagenomic techniques it is now possible to analyse the bacterial and metabolic composition of many fecal samples from the same infant, collected over time. Using these techniques, work by Tsukuda *et al.* [31] and Rao *et al.* [200] has provided extensive descriptions of the large-scale patterns of microbial and metabolic succession, as well as of the variability in development between infants. These analyses can show how the distribution of bacteria differs between infants, and may correlate different abundances with different conditions. However, it cannot answer the more fundamental 'why' questions. For example, as we discussed in chapter 2, *E. coli* is typically more abundant than *Bifidobacterium* because of the initial presence of oxygen in the gut. But why are the strictly anaerobic *Bacteroides* spp. often abundant in the infant gut, while the related *Prevotella* spp. are not? And why is *E. coli* typically abundant in the early infant gut, but sometimes the similarly facultatively anaerobic *Klebsiella* spp. is abundant? Without this understanding it is difficult to predict how these patterns may be influenced to create a healthier infant gut microbiota. This is where we believe *in silico* modelling can play a crucial role, by taking the bacterial and metabolic data as input, and providing further testable mechanistic predictions, as we have shown in this thesis.

In vitro studies may allow for these predictions to be tested. *In vitro* studies have already allowed for some more insight into, for example, the metabolism of *Bifidobacterium* [90]. Next to studies of infant fecal samples, these *in vitro* studies have been the primary source of input into our computational model. However, even in *in vitro* studies it often remains unclear why certain ecological effects occurs - e.g. why *Cutibacterium avidum* outcompetes *Anaerobutyricum hallii* when fed with lactate [194]. It is difficult to gather the right data and do the right experiments to find a mechanistic explanation for such complex effects without existing hypothesis. The wide variety of experimental conditions used in different *in vitro* studies (e.g. [245, 83, 160, 188]) further complicates creating a complete view of the interactions in the microbiota. Some studies include *in vivo* experiments in which germ-free mice or pigs are inoculated with a human microbiota. This allows relevant physiological factors to be included, such as the production of mucin by the gut wall, the absorption of metabolites, and immune interactions [318]. However, this approach also introduces its own biases [318]. Many species establish at different abundances in germ-free mice and pigs, compared to humans, and some species do not establish at all [319]. A better approach may be the recently developed gut-on-a-chip models. These *in vitro* models incorporate a living portion of the gut wall, better imitating the infant

gut conditions than an *in vitro* model that only includes bacteria. A gut-on-a-chip setup can also include physiological interactions, and conditions can be much more tightly controlled than in a mouse or pig model [320, 321]. *In silico* modelling may help guide these models to test more concrete hypothesis, such as the hypotheses we developed in chapter 3 and 4 on cross-feeding relations in the infant gut microbiota.

In conclusion, we expect that the metabolic modelling approach described in this thesis will be refined much further. We also expect that this approach will help explain the composition of the infant gut microbiota, and how it can be shaped.

Bibliography

[1] K. M. Kennedy, M. J. Gerlach, T. Adam, M. M. Heimesaat, L. Rossi, M. G. Surette, D. M. Sloboda, and T. Braun, "Fetal meconium does not have a detectable microbiota before birth," *Nature Microbiology*, vol. 6, no. 7, pp. 865–873, 2021.

[2] K. M. Kennedy, M. C. de Goffau, M. E. Perez-Muñoz, M. C. Arrieta, F. Bäckhed, P. Bork, T. Braun, F. D. Bushman, J. Dore, W. M. de Vos, A. M. Earl, J. A. Eisen, M. A. Elovitz, S. C. Ganal-Vonarburg, M. G. Gänzle, W. S. Garrett, L. J. Hall, M. W. Hornef, C. Huttenhower, L. Konnikova, S. Lebeer, A. J. Macpherson, R. C. Massey, A. C. McHardy, O. Koren, T. D. Lawley, R. E. Ley, L. O'Mahony, P. W. O'Toole, E. G. Pamer, J. Parkhill, J. Raes, T. Rattei, A. Salonen, E. Segal, N. Segata, F. Shanahan, D. M. Sloboda, G. C. Smith, H. Sokol, T. D. Spector, M. G. Surette, G. W. Tannock, A. W. Walker, M. Yassour, and J. Walter, "Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies," *Nature*, vol. 613, no. 7945, pp. 639–649, 2023.

[3] F. Bäckhed, J. Roswall, Y. Peng, Q. Feng, H. Jia, P. Kovatcheva-Datchary, Y. Li, Y. Xia, H. Xie, H. Zhong, M. T. Khan, J. Zhang, J. Li, L. Xiao, J. Al-Aama, D. Zhang, Y. S. Lee, D. Kotowska, C. Colding, V. Tremaroli, Y. Yin, S. Bergman, X. Xu, L. Madsen, K. Kristiansen, J. Dahlgren, and W. Jun, "Dynamics and stabilization of the human gut microbiome during the first year of life," *Cell Host and Microbe*, vol. 17, no. 5, pp. 690–703, 2015.

[4] G. Alessandri, F. Fontana, L. Mancabelli, G. A. Lugli, C. Tarracchini, C. Argentini, G. Longhi, A. Viappiani, C. Milani, F. Turroni, D. van Sinderen, and M. Ventura, "Exploring species-level infant gut bacterial biodiversity by meta-analysis and formulation of an optimized cultivation medium," *npj Biofilms and Microbiomes*, vol. 8, no. 1, 2022.

[5] C. Palmer, E. M. Bik, D. B. DiGiulio, D. A. Relman, and P. O. Brown, "Development of the human infant intestinal microbiota," *PLoS Biology*, vol. 5, no. 7, pp. 1556–1573, 2007.

[6] Y. Tsunoda, T. Asahara, K. Nomoto, Y. Yoshioka, and E. Fukuma, "Bacterial profile of infant feces associated with lactation infectious breasts," *Pediatric Health, Medicine and Therapeutics*, vol. Volume 9, pp. 173–180, 2018.

[7] T. Escherich, *Die darmbakterien des säuglings und ihre beziehungen zur physiologie der Verdauung*. Enke, 1886.

[8] P. A. Scholtens, R. Oozeer, R. Martin, K. B. Amor, and J. Knol, "The Early Settlers: Intestinal Microbiology in Early Life," *Annual Review of Food Science and Technology*, vol. 3, no. 1, pp. 425–447, 2012.

[9] S. Wang, S. Zeng, M. Egan, P. Cherry, C. Strain, E. Morais, P. Boyaval, C. A. Ryan, E. M. Dempsey, R. P. Ross, and C. Stanton, "Metagenomic analysis of mother-infant gut microbiome reveals global distinct and shared microbial signatures," *Gut Microbes*, vol. 13, no. 1, pp. 1–24, 2021.

[10] S. Magnúsdóttir, A. Heinken, L. Kutt, D. A. Ravcheev, E. Bauer, A. Noronha, K. Greenhalgh, C. Jäger, J. Baginska, P. Wilmes, R. M. Fleming, and I. Thiele, "Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota," *Nature Biotechnology*, vol. 35, no. 1, pp. 81–89, 2017.

[11] N. T. Devika and K. Raman, "Deciphering the metabolic capabilities of Bifidobacteria using genome-scale metabolic models," *Scientific Reports*, vol. 9, no. 1, pp. 1–9, 2019.

[12] K. Korpela and W. M. de Vos, "Infant gut microbiota restoration: state of the art," *Gut Microbes*, vol. 14, no. 1, 2022.

[13] J. D. Orth, I. Thiele, and B. O. Palsson, "What is flux balance analysis?," *Nature Biotechnology*, vol. 28, no. 3, pp. 245–248, 2010.

[14] S. Matamoros, C. Gras-Leguen, F. Le Vacon, G. Potel, and M. F. De La Cochetiere, "Development of intestinal microbiota in infants and its impact on health," *Trends in Microbiology*, vol. 21, no. 4, pp. 167–173, 2013.

[15] J. M. Johnson and E. D. Adams, "The Gastrointestinal Microbiome in Infant Colic," *MCN: The American Journal of Maternal/Child Nursing*, pp. 1–12, 2022.

[16] S. Dogra, O. Sakwinska, S. E. Soh, C. Ngom-Bru, W. M. Brück, B. Berger, H. Brüssow, N. Karnani, Y. S. Lee, F. Yap, Y. S. Chong, K. M. Godfrey, and J. D. Holbrook, "Rate of establishing the gut microbiota in infancy has consequences for future health," *Gut Microbes*, vol. 6, no. 5, pp. 321–325, 2015.

[17] I. O'Neill, Z. Schofield, and L. J. Hall, "Exploring the role of the microbiota member Bifidobacterium in modulating immune-linked diseases," *Emerging Topics in Life Sciences*, vol. 1, no. 4, pp. 333–349, 2017.

[18] R. M. Duar, D. Kyle, and G. Casaburi, "Colonization Resistance in the Infant Gut: The Role of B. infantis in Reducing pH and Preventing Pathogen Growth," *High-Throughput*, vol. 9, no. 2, p. 7, 2020.

[19] S. Fukuda, H. Toh, K. Hase, K. Oshima, Y. Nakanishi, K. Yoshimura, T. Tobe, J. M. Clarke, D. L. Topping, T. Suzuki, T. D. Taylor, K. Itoh, J. Kikuchi, H. Morita, M. Hattori, and H. Ohno, "Bifidobacteria can protect from enteropathogenic infection through production of acetate," *Nature*, vol. 469, no. 7331, pp. 543–549, 2011.

[20] O. Appert, A. R. Garcia, R. Frei, C. Roduit, F. Constancias, V. Neuzil-Bunesova, R. Ferstl, J. Zhang, C. Akdis, R. Lauener, C. Lacroix, and C. Schwab, "Initial butyrate producers during infant gut microbiota development are endospore formers," *Environmental Microbiology*, vol. 22, no. 9, pp. 3909–3921, 2020.

[21] A. Cait, E. Cardenas, P. A. Dimitriu, N. Amenyogbe, D. Dai, J. Cait, H. Sbihi, L. Stiemsma, P. Subbarao, P. J. Mandhane, A. B. Becker, T. J. Moraes, M. R. Sears, D. L. Lefebvre, M. B. Azad, T. Kollmann, S. E. Turvey, and W. W. Mohn, "Reduced genetic potential for butyrate fermentation in the gut microbiome of infants who develop allergic sensitization," *Journal of Allergy and Clinical Immunology*, vol. 144, no. 6, pp. 1638–1647.e3, 2019.

[22] L. Nylund, M. Nermes, E. Isolauri, S. Salminen, W. M. De Vos, and R. Satokari, "Severity of atopic disease inversely correlates with intestinal microbiota diversity and butyrate-producing bacteria," *Allergy: European Journal of Allergy and Clinical Immunology*, vol. 70, no. 2, pp. 241–244, 2015.

[23] H. Wopereis, K. Sim, A. Shaw, J. O. Warner, J. Knol, and J. S. Kroll, "Intestinal microbiota in infants at high risk for allergy: Effects of prebiotics and role in eczema development," *Journal of Allergy and Clinical Immunology*, vol. 141, no. 4, pp. 1334–1342.e5, 2018.

[24] C. De Weerth, S. Fuentes, P. Puylaert, and W. M. De Vos, "Intestinal microbiota of infants with colic: Development and specific signatures," *Pediatrics*, vol. 131, no. 2, 2013.

[25] D. R. Donohoe, N. Garge, X. Zhang, W. Sun, T. M. O'Connell, M. K. Bunger, and S. J. Bultman, "The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon," *Cell Metabolism*, vol. 13, no. 5, pp. 517–526, 2011.

[26] K. Meijer, P. De Vos, and M. G. Priebe, "Butyrate and other short-chain fatty acids as modulators of immunity: What relevance for health?," *Current Opinion in Clinical Nutrition and Metabolic Care*, vol. 13, no. 6, pp. 715–721, 2010.

[27] V. Bunesova, C. Lacroix, and C. Schwab, "Mucin Cross-Feeding of Infant Bifidobacteria and Eubacterium hallii," *Microbial Ecology*, vol. 75, no. 1, pp. 228–238, 2018.

[28] M. S. Desai, A. M. Seekatz, N. M. Koropatkin, N. Kamada, C. A. Hickey, M. Wolter, N. A. Pudlo, S. Kitamoto, N. Terrapon, A. Muller, V. B. Young, B. Henrissat, P. Wilmes, T. S. Stappenbeck, G. Núñez, and E. C. Martens, "A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility," *Cell*, vol. 167, no. 5, pp. 1339–1353.e21, 2016.

[29] S. Rokhsefat, A. Lin, and E. M. Comelli, "Mucin–Microbiota Interaction During Postnatal Maturation of the Intestinal Ecosystem: Clinical Implications," *Digestive Diseases and Sciences*, vol. 61, no. 6, pp. 1473–1486, 2016.

[30] S. Karav, G. Casaburi, and S. A. Frese, "Reduced colonic mucin degradation in breastfed infants colonized by Bifidobacterium longum subsp. infantis EVC001," *FEBS Open Bio*, vol. 8, no. 10, pp. 1649–1657, 2018.

[31] N. Tsukuda, K. Yahagi, T. Hara, Y. Watanabe, H. Matsumoto, H. Mori, K. Higashi, H. Tsuji, S. Matsumoto, K. Kurokawa, and T. Matsuki, "Key bacterial taxa and metabolic pathways affecting gut short-chain fatty acid profiles in early life," *The ISME Journal*, vol. 15, pp. 2574–2590, sep 2021.

[32] T. Matsuki, K. Yahagi, H. Mori, H. Matsumoto, T. Hara, S. Tajima, E. Ogawa, H. Kodama, K. Yamamoto, T. Yamada, S. Matsumoto, and K. Kurokawa, "A key genetic factor for fucosyllactose utilization affects infant gut microbiota development," *Nature Communications*, vol. 7, no. May, pp. 1–12, 2016.

[33] S. Dogra, O. Sakwinska, S.-E. E. Soh, C. Ngom-Bru, W. M. Brück, B. Berger, H. Brüssow, Y. S. Lee, F. Yap, Y.-S. S. Chong, K. M. Godfrey, J. D. Holbrook, and GUSTO Study Group, "Dynamics of infant gut microbiota are influenced by delivery mode and gestational duration and are associated with subsequent adiposity," *mBio*, vol. 6, no. 1, pp. 1–9, 2015.

[34] D. Matharu, A. J. Ponsero, E. Dikareva, K. Korpela, K. L. Kolho, W. M. de Vos, and A. Salonen, "Bacteroides abundance drives birth mode dependent infant gut microbiota developmental trajectories," *Frontiers in Microbiology*, vol. 13, 2022.

[35] Q. Tang, G. Jin, G. Wang, T. Liu, X. Liu, B. Wang, and H. Cao, "Current Sampling Methods for Gut Microbiota: A Call for More Precise Devices," *Frontiers in Cellular and Infection Microbiology*, vol. 10, no. April, 2020.

[36] J. Rromano-Keeler, D. J. Moore, C. Wang, R. M. Brucker, C. Fonnesbeck, J. C. Slaughter, H. Li, D. P. Curran, S. Meng, H. Correa, H. N. Lovvorn, Y. W. Tang, S. Bordenstein, A. L. George, and J. H. Weitkamp, "Early life

establishment of site-specific microbial communities in the gut," *Gut Microbes*, vol. 5, no. 2, pp. 192–201, 2014.

[37] M. Rajilić-Stojanović, H. Smidt, and W. M. De Vos, "Diversity of the human gastrointestinal tract microbiota revisited," *Environmental Microbiology*, vol. 9, no. 9, pp. 2125–2136, 2007.

[38] D. N. Frank and N. R. Pace, "Gastrointestinal microbiology enters the metagenomics era," *Current Opinion in Gastroenterology*, vol. 24, no. 1, pp. 4–10, 2008.

[39] S. R. Gill, M. Pop, R. T. DeBoy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. A. Relman, C. M. Fraser-Liggett, and K. E. Nelson, "Metagenomic Analysis of the Human Distal Gut Microbiome," *Science*, vol. 312, pp. 1355–1359, jun 2006.

[40] L. Beller, W. Deboutte, G. Falony, S. Vieira-Silva, R. Y. Tito, M. Valles-Colomer, L. Rymenans, D. Jansen, L. Van Espen, M. I. Papadaki, C. Shi, C. K. Yinda, M. Zeller, K. Faust, M. Van Ranst, J. Raes, and J. Matthijnssens, "Successional Stages in Infant Gut Microbiota Maturation," *mBio*, vol. 12, no. 6, 2021.

[41] M. C. Arrieta, L. T. Stiemsma, N. Amenyogbe, E. Brown, and B. Finlay, "The intestinal microbiome in early life: Health and disease," *Frontiers in Immunology*, vol. 5, no. AUG, pp. 1–18, 2014.

[42] H. M. Timmerman, N. B. Rutten, J. Boekhorst, D. M. Saulnier, G. A. Kortman, N. Contractor, M. Kullen, E. Floris, H. J. Harmsen, A. M. Vlieger, M. Kleerebezem, and G. T. Rijkers, "Intestinal colonisation patterns in breastfed and formula-fed infants during the first 12 weeks of life reveal sequential microbiota signatures," *Scientific Reports*, vol. 7, no. 1, pp. 1–10, 2017.

[43] V. Sagheddu, V. Patrone, F. Miragoli, E. Puglisi, and L. Morelli, "Infant early gut colonization by Lachnospiraceae: High frequency of Ruminococcus gnavus," *Frontiers in Pediatrics*, vol. 4, pp. 7–12, 2016.

[44] V. N. Rocha Martin, C. Schwab, L. Krych, E. Voney, A. Geirnaert, C. Braegger, and C. Lacroix, "Colonization of Cutibacterium avidum during infant gut microbiota establishment," *FEMS Microbiology Ecology*, vol. 95, no. 1, pp. 1–14, 2018.

[45] Y. Liu, Z. Zheng, L. Yu, S. Wu, L. Sun, S. Wu, Q. Xu, S. Cai, N. Qin, and W. Bao, "Examination of the temporal and spatial dynamics of the gut microbiome in newborn piglets reveals distinct microbial communities in six intestinal segments," *Scientific Reports*, vol. 9, no. 1, pp. 1–8, 2019.

[46] X. Cao, A. Dong, G. Kang, X. Wang, L. Duan, H. Hou, T. Zhao, S. Wu, X. Liu, H. Huang, and R. Wu, "Modeling spatial interaction networks of the gut microbiota," *Gut Microbes*, vol. 14, no. 1, 2022.

[47] M. V. Gryaznova, Y. D. Dvoretskaya, M. Y. Syromyatnikov, S. V. Shabunin, P. A. Parshin, E. V. Mikhaylov, N. A. Strelnikov, and V. N. Popov, "Changes in the Microbiome Profile in Different Parts of the Intestine in Piglets with Diarrhea," *Animals*, vol. 12, no. 3, 2022.

[48] K. R. James, T. Gomes, R. Elmentaite, N. Kumar, E. L. Gulliver, H. W. King, M. D. Stares, B. R. Bareham, J. R. Ferdinand, V. N. Petrova, K. Polański, S. C. Forster, L. B. Jarvis, O. Suchanek, S. Howlett, L. K. James, J. L. Jones, K. B. Meyer, M. R. Clatworthy, K. Saeb-Parsy, T. D. Lawley, and S. A. Teichmann, "Distinct microbial and immune niches of the human colon," *Nature Immunology*, vol. 21, no. 3, pp. 343–353, 2020.

[49] C. Tropini, K. A. Earle, K. C. Huang, and J. L. Sonnenburg, "The Gut Microbiome: Connecting Spatial Organization to Function," *Cell Host and Microbe*, vol. 21, no. 4, pp. 433–442, 2017.

[50] Y. Liu, S. Qin, Y. Song, Y. Feng, N. Lv, Y. Xue, F. Liu, S. Wang, B. Zhu, J. Ma, and H. Yang, "The Perturbation of Infant Gut Microbiota Caused by Cesarean Delivery Is Partially Restored by Exclusive Breastfeeding," *Frontiers in Microbiology*, vol. 10, no. March, pp. 1–11, 2019.

[51] S. J. Song, M. G. Dominguez-Bello, and R. Knight, "How delivery mode and feeding can shape the bacterial community in the infant gut," *Cmaj*, vol. 185, no. 5, pp. 373–374, 2013.

[52] E. Van Daele, K. Kamphorst, A. M. Vlieger, G. Hermes, C. Milani, M. Ventura, C. Belzer, H. Smidt, R. M. van Elburg, and J. Knol, "Effect of antibiotics in the first week of life on faecal microbiota development," *Archives of Disease in Childhood: Fetal and Neonatal Edition*, pp. 603–610, 2022.

[53] B. Berger, N. Porta, F. Foata, D. Grathwohl, M. Delley, D. Moine, A. Charpagne, L. Siegwald, P. Descombes, P. Alliet, G. Puccio, P. Steenhout, A. Mercenier, and N. Sprenger, "Linking human milk oligosaccharides, infant fecal community types, and later risk to require antibiotics," *mBio*, vol. 11, no. 2, pp. 1–18, 2020.

[54] D. H. Taft, Z. T. Lewis, N. Nguyen, S. Ho, C. Masarweh, V. Dunne-Castagna, D. J. Tancredi, M. N. Huda, C. B. Stephensen, K. Hinde, E. von Mutius, P. V. Kirjavainen, J. C. Dalphin, R. Lauener, J. Riedler, J. T. Smilowitz, J. B. German, A. L. Morrow, and D. A. Mills, "Bifidobacterium Species Colonization in Infancy: A Global Cross-Sectional Comparison by Population History of Breastfeeding," *Nutrients*, vol. 14, no. 7, 2022.

[55] N. A. Bokulich, J. Chung, T. Battaglia, N. Henderson, M. Jay, H. Li, A. D. Lieber, F. Wu, G. I. Perez-Perez, Y. Chen, W. Schweizer, X. Zheng, M. Contreras, M. G. Dominguez-Bello, and M. J. Blaser, "Antibiotics, birth mode, and diet shape microbiome maturation during early life," *Science Translational Medicine*, vol. 8, no. 343, pp. 1–14, 2016.

[56] P. Ferretti, E. Pasolli, A. Tett, F. Asnicar, V. Gorfer, S. Fedi, F. Armanini, D. T. Truong, S. Manara, M. Zolfo, F. Beghini, R. Bertorelli, V. De Sanctis, I. Bariletti, R. Canto, R. Clementi, M. Cologna, T. Crifò, G. Cusumano, S. Gottardi, C. Innamorati, C. Masè, D. Postai, D. Savoi, S. Duranti, G. A. Lugli, L. Mancabelli, F. Turroni, C. Ferrario, C. Milani, M. Mangifesta, R. Anzalone, A. Viappiani, M. Yassour, H. Vlamakis, R. Xavier, C. M. Collado, O. Koren, S. Tateo, M. Soffiati, A. Pedrotti, M. Ventura, C. Huttenhower, P. Bork, and N. Segata, "Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome," *Cell Host and Microbe*, vol. 24, no. 1, pp. 133–145.e5, 2018.

[57] F. Fouhy, C. M. Guinane, S. Hussey, R. Wall, C. A. Ryan, E. M. Dempsey, B. Murphy, R. P. Ross, G. F. Fitzgerald, C. Stanton, and P. D. Cotter, "High-throughput sequencing reveals the incomplete, short-term recovery of infant gut microbiota following parenteral antibiotic treatment with ampicillin and gentamicin," *Antimicrobial Agents and Chemotherapy*, vol. 56, no. 11, pp. 5811–5820, 2012.

[58] S. Tanaka, T. Kobayashi, P. Songjinda, A. Tateyama, M. Tsubouchi, C. Kiyohara, T. Shirakawa, K. Sonomoto, and J. Nakayama, "Influence of antibiotic exposure in the early postnatal period on the development of intestinal microbiota," *FEMS Immunology and Medical Microbiology*, vol. 56, no. 1, pp. 80–87, 2009.

[59] C. Greenwood, A. L. Morrow, A. J. Lagomarcino, M. Altaye, D. H. Taft, Z. Yu, D. S. Newburg, D. V. Ward, and K. R. Schibler, "Early empiric antibiotic use in preterm infants is associated with lower bacterial diversity and higher relative abundance of enterobacter," *Journal of Pediatrics*, vol. 165, no. 1, pp. 23–29, 2014.

[60] S. Arboleya, B. Sánchez, C. Milani, S. Duranti, G. Solís, N. Fernández, C. G. De Los Reyes-Gavilán, M. Ventura, A. Margolles, and M. Gueimonde, "Intestinal microbiota development in preterm neonates and effect of perinatal antibiotics," *Journal of Pediatrics*, vol. 166, no. 3, pp. 538–544, 2015.

[61] O. Brunser, M. Gotteland, S. Cruchet, G. Figueroa, D. Garrido, and P. Steenhout, "Effect of a milk formula with prebiotics on the intestinal microbiota of infants after an antibiotic treatment," *Pediatric Research*, vol. 59, no. 3, pp. 451–456, 2006.

[62] D. A. Medina, F. Pinto, A. Ovalle, P. Thomson, and D. Garrido, "Prebiotics mediate microbial interactions in a consortium of the infant gut microbiome," *International Journal of Molecular Sciences*, vol. 18, no. 10, pp. 1–16, 2017.

[63] V. Bunesova, C. Lacroix, and C. Schwab, "Fucosyllactose and L-fucose utilization of infant Bifidobacterium longum and Bifidobacterium kashiwanohense," *BMC Microbiology*, vol. 16, no. 1, pp. 1–12, 2016.

[64] M. Böger, S. S. Van Leeuwen, A. Lammerts Van Bueren, and L. Dijkhuizen, "Structural Identity of Galactooligosaccharide Molecules Selectively Utilized by Single Cultures of Probiotic Bacterial Strains," *Journal of Agricultural and Food Chemistry*, vol. 67, no. 50, pp. 13969–13977, 2019.

[65] M. A. Underwood, J. B. German, C. B. Lebrilla, and D. A. Mills, "Bifidobacterium longum subspecies infantis: Champion colonizer of the infant gut," *Pediatric Research*, vol. 77, no. 1, pp. 229–235, 2015.

[66] O. Ballard and A. L. Morrow, "Human Milk Composition. Nutrients and Bioactive Factors," *Pediatric Clinics of North America*, vol. 60, no. 1, pp. 49–74, 2013.

[67] P. Chaturvedi, C. D. Warren, M. Altaye, A. L. Morrow, G. Ruiz-Palacios, L. K. Pickering, and D. S. Newburg, "Fucosylated human milk oligosaccharides vary between individuals and over the course of lactation," *Glycobiology*, vol. 11, no. 5, pp. 365–372, 2001.

[68] S. D. Durham, Z. Wei, D. G. Lemay, M. C. Lange, and D. Barile, "Creation of a milk oligosaccharide database, MilkOligoDB, reveals common structural motifs and extensive diversity across mammals," *Scientific Reports*, vol. 13, no. 1, pp. 1–26, 2023.

[69] H. C. Bergman and C. Housley, "Chemical analyses of American opossum (didelphys virginiana) milk," *Comparative Biochemistry And Physiology*, vol. 25, pp. 213–218, apr 1968.

[70] S. Kijner, O. Kolodny, and M. Yassour, "Human milk oligosaccharides and the infant gut microbiome from an eco-evolutionary perspective," *Current Opinion in Microbiology*, vol. 68, p. 102156, 2022.

[71] G. Boehm, B. Stahl, J. Jelinek, J. Knol, V. Miniello, and G. E. Moro, "Prebiotic carbohydrates in human milk and formulas," *Acta Paediatrica, International Journal of Paediatrics, Supplement*, vol. 94, no. 449, pp. 18–21, 2005.

[72] R. J. Gibbons and B. Kapsimalis, "Estimates of the overall rate of growth of the intestinal microflora of hamsters, guinea pigs, and mice.," *Journal of Bacteriology*, vol. 93, no. 1, pp. 510–512, 1967.

[73] A. W. Thorman, G. Adkins, S. C. Conrey, A. R. Burrell, Y. Yu, B. White, R. Burke, D. Haslam, D. C. Payne, M. A. Staat, A. L. Morrow, and D. S. Newburg, "Gut Microbiome Composition and Metabolic Capacity Differ by FUT2 Secretor Status in Exclusively Breastfed Infants," *Nutrients*, vol. 15, no. 2, pp. 1–13, 2023.

[74] V. T. Pham, C. Lacroix, C. P. Braegger, and C. Chassard, "Lactate-utilizing community is associated with gut microbiota dysbiosis in colicky infants," *Scientific Reports*, vol. 7, no. 1, pp. 1–13, 2017.

[75] R. Francavilla, M. Calasso, L. Calace, S. Siragusa, M. Ndagijimana, P. Vernocchi, L. Brunetti, G. Mancino, G. Tedeschi, E. Guerzoni, F. Indrio, L. Laghi, V. L. Miniello, M. Gobbetti, and M. De Angelis, "Effect of lactose on gut microbiota and metabolome of infants with cow's milk allergy," *Pediatric Allergy and Immunology*, vol. 23, no. 5, pp. 420–427, 2012.

[76] J. Cummings and G. Macfarlane, "The control and consequences of bacterial fermentation in the human colon," *Journal of Applied Bacteriology*, vol. 70, pp. 443–459, jun 1991.

[77] A. Cederlund, Y. Kai-Larsen, G. Printz, H. Yoshio, G. Alvelius, H. Lagercrantz, R. Strömberg, H. Jörnvall, G. H. Gudmundsson, and B. Agerberth, "Lactose in Human Breast Milk an Inducer of Innate Immunity with Implications for a Role in Intestinal Homeostasis," *PLoS ONE*, vol. 8, no. 1, 2013.

[78] L. T. Weaver, M. F. Laker, and R. Nelson, "Neonatal intestinal lactase activity," *Archives of Disease in Childhood*, vol. 61, no. 9, pp. 896–899, 1986.

[79] W. C. Maclean, B. B. Fink, D. A. Schoeller, W. Wong, and P. D. Klein, "Lactose assimilation by full-term infants: Relation of [13C] and H2 breath tests with fecal [13C] excretion," *Pediatric Research*, vol. 17, no. 8, pp. 629–633, 1983.

[80] R. G. Heine, F. Alrefaee, P. Bachina, J. C. De Leon, L. Geng, S. Gong, J. A. Madrazo, J. Ngamphaiboon, C. Ong, and J. M. Rogacion, "Lactose intolerance and gastrointestinal cow's milk allergy in infants and children - Common misconceptions revisited," *World Allergy Organization Journal*, vol. 10, no. 1, pp. 1–8, 2017.

[81] R. B. Jones, P. K. Berger, J. F. Plows, T. L. Alderete, J. Millstein, J. Fogel, S. N. Iablokov, D. A. Rodionov, A. L. Osterman, L. Bode, and M. I. Goran, "Lactose-reduced infant formula with added corn syrup solids is associated with a distinct gut microbiota in Hispanic infants," *Gut Microbes*, vol. 12, no. 1, 2020.

[82] C. Schwab, H. J. Ruscheweyh, V. Bunesova, V. T. Pham, N. Beerenwinkel, and C. Lacroix, "Trophic interactions of infant bifidobacteria and eubacterium hallii during L-fucose and fucosyllactose degradation," *Frontiers in Microbiology*, vol. 8, pp. 1–14, 2017.

[83] M. Centanni, S. A. Ferguson, I. M. Sims, A. Biswas, and G. W. Tannocka, "Bifidobacterium bifidum ATCC 15696 and bifidobacterium breve 24b metabolic interaction based on 2-O-Fucosyl-lactose studied in steady-state cultures in a freter-style chemostat," *Applied and Environmental Microbiology*, vol. 85, no. 7, pp. 1–17, 2019.

[84] M. O'Connell Motherway, M. Kinsella, G. F. Fitzgerald, and D. Van Sinderen, "Transcriptional and functional characterization of genetic elements involved in galacto-oligosaccharide utilization by Bifidobacterium breve UCC2003," *Microbial Biotechnology*, vol. 6, no. 1, pp. 67–79, 2013.

[85] A. H. Viborg, T. Katayama, M. Abou Hachem, M. C. Andersen, M. Nishimoto, M. H. Clausen, T. Urashima, B. Svensson, and M. Kitaoka, "Distinct substrate specificities of three glycoside hydrolase family 42 β-galactosidases from Bifidobacterium longum subsp. infantis ATCC 15697," *Glycobiology*, vol. 24, no. 2, pp. 208–216, 2014.

[86] S. W. Hinz, M. I. Pastink, L. A. Van Den Broek, J. P. Vincken, and A. G. Voragen, "Bifidobacterium longum endogalactanase liberates galactotriose from type I galactans," *Applied and Environmental Microbiology*, vol. 71, no. 9, pp. 5501–5510, 2005.

[87] L. E. Tailford, E. H. Crost, D. Kavanaugh, and N. Juge, "Mucin glycan foraging in the human gut microbiome," *Frontiers in Genetics*, vol. 6, mar 2015.

[88] J. Gore, H. Youk, and A. Van Oudenaarden, "Snowdrift game dynamics and facultative cheating in yeast," *Nature*, vol. 459, no. 7244, pp. 253–256, 2009.

[89] N. W. Smith, P. R. Shorten, E. Altermann, N. C. Roy, and W. C. McNabb, "The Classification and Evolution of Bacterial Cross-Feeding," *Frontiers in Ecology and Evolution*, vol. 7, no. May, pp. 1–15, 2019.

[90] L. De Vuyst, F. Moens, M. Selak, A. Rivière, and F. Leroy, "Summer Meeting 2013: Growth and physiology of bifidobacteria," *Journal of Applied Microbiology*, vol. 116, no. 3, pp. 477–491, 2014.

[91] V. T. Pham, C. Lacroix, C. P. Braegger, and C. Chassard, "Early colonization of functional groups of microbes in the infant gut," *Environmental microbiology*, vol. 18, no. 7, pp. 2246–2258, 2016.

[92] S. H. Duncan, P. Louis, and H. J. Flint, "Lactate-Utilizing Bacteria, Isolated from Human Feces, That Produce Butyrate as a Major Fermentation Product," *Applied and Environmental Microbiology*, vol. 70, no. 10, pp. 5810–5817, 2004.

[93] J. M. Garber, H. Nothaft, B. Pluvinage, M. Stahl, X. Bian, S. Porfirio, A. Enriquez, J. Butcher, H. Huang, J. Glushka, E. Line, J. A. Gerlt, P. Azadi, A. Stintzi, A. B. Boraston, and C. M. Szymanski, "The gastrointestinal

pathogen Campylobacter jejuni metabolizes sugars with potential help from commensal Bacteroides vulgatus," *Communications Biology*, vol. 3, no. 1, pp. 1–11, 2020.

[94] K. M. Ng, J. A. Ferreyra, S. K. Higginbottom, J. B. Lynch, P. C. Kashyap, S. Gopinath, N. Naidu, B. Choudhury, B. C. Weimer, D. M. Monack, and J. L. Sonnenburg, "Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens," *Nature*, vol. 502, no. 7469, pp. 96–99, 2013.

[95] M. Rogosa, "the Genus Veillonella. I. General Cultural, Ecological, and Biochemical Considerations," *Journal of Bacteriology*, vol. 87, pp. 162–170, 1964.

[96] B. Zabel, C. C. Yde, P. Roos, J. Marcussen, H. M. Jensen, K. Salli, J. Hirvonen, A. C. Ouwehand, and W. Morovic, "Novel Genes and Metabolite Trends in Bifidobacterium longum subsp. infantis Bi-26 Metabolism of Human Milk Oligosaccharide 2'-fucosyllactose," *Scientific Reports*, vol. 9, no. 1, pp. 1–11, 2019.

[97] D. Gonze, K. Z. Coyte, L. Lahti, and K. Faust, "Microbial communities as dynamical systems," *Current Opinion in Microbiology*, vol. 44, pp. 41–49, 2018.

[98] K. Faust and J. Raes, "Microbial interactions: From networks to models," *Nature Reviews Microbiology*, vol. 10, no. 8, pp. 538–550, 2012.

[99] P. Sen and M. Orešič, "Metabolic Modeling of Human Gut Microbiota on a Genome Scale: An Overview," *Metabolites*, vol. 9, no. 2, p. 22, 2019.

[100] R. A. Majewski and M. M. Domach, "Simple constrained-optimization view of acetate overflow in E. coli," *Biotechnology and Bioengineering*, vol. 35, no. 7, pp. 732–738, 1990.

[101] S. Magnúsdóttir and I. Thiele, "Modeling metabolism of the human gut microbiome," *Current Opinion in Biotechnology*, vol. 51, pp. 90–96, 2018.

[102] N. Martyushenko and E. Almaas, "ModelExplorer - software for visual inspection and inconsistency correction of genome-scale metabolic reconstructions," *BMC Bioinformatics*, vol. 20, no. 1, pp. 1–8, 2019.

[103] D. M. Versluis, R. Schoemaker, E. Looijesteijn, D. Muysken, P. V. Jeurink, M. Paques, J. M. W. Geurts, and R. M. H. Merks, "A Multiscale Spatiotemporal Model Including a Switch from Aerobic to Anaerobic Metabolism Reproduces Succession in the Early Infant Gut Microbiota," *mSystems*, vol. 7, oct 2022.

[104] T. E. Harris and F. S. Ross, "Fundamentals of a method for evaluating rail net capacities," tech. rep., RAND Corporation, 1955.

[105] G. T. Cocks, J. Aguilar, and E. C. Lin, "Evolution of L-1,2 propanediol catabolism in Escherichia coli by recruitment of enzymes for L-fucose and L-lactate metabolism," *Journal of Bacteriology*, vol. 118, no. 1, pp. 83–88, 1974.

[106] L. Bode, "Human milk oligosaccharides: Every baby needs a sugar mama," *Glycobiology*, vol. 22, no. 9, pp. 1147–1162, 2012.

[107] S. Gudmundsson and I. Thiele, "Computationally efficient flux variability analysis," *BMC Bioinformatics*, vol. 11, no. 2, pp. 2–4, 2010.

[108] J. Zanghellini, D. E. Ruckerbauer, M. Hanscho, and C. Jungreuthmayer, "Elementary flux modes in a nutshell: Properties, calculation and applications," *Biotechnology Journal*, vol. 8, no. 9, pp. 1009–1016, 2013.

[109] A. Vazquez and Z. N. Oltvai, "Macromolecular crowding explains overflow metabolism in cells," *Scientific Reports*, vol. 6, pp. 1–7, 2016.

[110] A. Hoppe, S. Hoffmann, and H. G. Holzhütter, "Including metabolite concentrations into flux balance analysis: Thermodynamic realizability as a constraint on flux distributions in metabolic networks," *BMC Systems Biology*, vol. 1, pp. 1–12, 2007.

[111] K. Zhuang, G. N. Vemuri, and R. Mahadevan, "Economics of membrane occupancy and respiro-fermentation," *Molecular Systems Biology*, vol. 7, no. 500, pp. 1–9, 2011.

[112] D. H. de Groot, J. Lischke, R. Muolo, R. Planqué, F. J. Bruggeman, and B. Teusink, "The common message of constraint-based optimization approaches: overflow metabolism is caused by two growth-limiting constraints," *Cellular and Molecular Life Sciences*, vol. 77, no. 3, pp. 441–453, 2020.

[113] A. Varma and B. Ø. Palsson, "Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use," *Biotechnology*, vol. 12, no. October, pp. 994–998, 1994.

[114] J. L. Reed and B. Palsson, "Thirteen years of building constraint-based in silico models of Escherichia coli," *Journal of Bacteriology*, vol. 185, no. 9, pp. 2692–2699, 2003.

[115] S. Stolyar, S. Van Dien, K. L. Hillesland, N. Pinel, T. J. Lie, J. A. Leigh, and D. A. Stahl, "Metabolic modeling of a mutualistic microbial community," *Molecular Systems Biology*, vol. 3, no. 92, pp. 1–14, 2007.

[116] R. Mahadevan, J. S. Edwards, and F. J. Doyle, "Dynamic flux balance analysis of diauxic growth," *Biophysical Journal*, vol. 83, no. 3, pp. 1331–1340, 2002.

[117] C. Baroukh, R. Muñoz-Tamayo, J. P. Steyer, and O. Bernard, "DRUM: A new framework for metabolic modeling under non-balanced growth. Application to the carbon metabolism of unicellular microalgae," *PLoS ONE*, vol. 9, no. 8, 2014.

[118] K. Zhuang, M. Izallalen, P. Mouser, H. Richter, C. Risso, R. Mahadevan, and D. R. Lovley, "Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments," *ISME Journal*, vol. 5, no. 2, pp. 305–316, 2011.

[119] B. Allen, J. Gore, and M. A. Nowak, "Spatial dilemmas of diffusible public goods," *eLife*, vol. 2013, no. 2, pp. 1–11, 2013.

[120] J. L. Welch, Y. Hasegawa, N. P. McNulty, J. I. Gordon, and G. G. Borisy, "Spatial organization of a model 15-member human gut microbiota established in gnotobiotic mice," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 43, pp. E9105–E9114, 2017.

[121] W. R. Harcombe, W. J. Riehl, I. Dukovski, B. R. Granger, A. Betts, A. H. Lang, G. Bonilla, A. Kar, N. Leiby, P. Mehta, C. J. Marx, and D. Segrè, "Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics," *Cell Reports*, vol. 7, no. 4, pp. 1104–1115, 2014.

[122] M. J. A. van Hoek and R. M. H. Merks, "Emergence of microbial diversity due to cross-feeding interactions in a spatial model of gut microbial metabolism," *BMC Systems Biology*, vol. 11, no. 1, pp. 1–18, 2017.

[123] E. Bauer, J. Zimmermann, F. Baldini, I. Thiele, and C. Kaleta, "BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities," *PLoS Computational Biology*, vol. 13, no. 5, pp. 1–22, 2017.

[124] Chan, Friedman, Wu, and Maranas, "Predicting the Longitudinally and Radially Varying Microbial Metabolic Modeling," *Processes*, vol. 7, p. 394, 2019.

[125] P. L. Stark and A. Lee, "The microbial ecology of the large bowel of breast-fed and formula-fed infants during the first year of life," *Journal of Medical Microbiology*, vol. 15, no. 2, pp. 189–203, 1982.

[126] B. M. Dotinga, J. P. Mintzer, J. E. Moore, J. B. Hulscher, A. F. Bos, and E. M. Kooi, "Maturation of Intestinal Oxygenation: A Review of Mechanisms and Clinical Implications for Preterm Neonates," *Frontiers in Pediatrics*, vol. 8:354, 2020.

[127] E. S. Friedman, K. Bittinger, T. V. Esipova, L. Hou, L. Chau, J. Jiang, C. Mesaros, P. J. Lund, X. Liang, G. A. FitzGerald, M. Goulian, D. Lee, B. A. Garcia, I. A. Blair, S. A. Vinogradov, and G. D. Wu, "Microbes vs. chemistry in the origin of the anaerobic gut lumen," *Proceedings of the National Academy*

of Sciences of the United States of America, vol. 115, no. 16, pp. 4170–4175, 2018.

[128] C. Lay, C. W. Chu, R. W. Purbojati, E. Acerbi, D. I. Drautz-Moses, P. F. de Sessions, S. Jie, E. Ho, Y. J. Kok, X. Bi, S. Chen, S. Y. Mak, M. C. Chua, A. E. Goh, W. C. Chiang, R. Rao, S. Chaithongwongwatthana, N. Khemapech, V. Chongsrisawat, R. Martin, Y. Koh, S. R. Lohar, I. C. H. Tan, W. A. Anng, C. Jie, N. Bartke, K. Ben-Amor, I. B. Renes, F. Wong, G. Roeselers, Y. S. Ho, M. L. Hibberd, S. C. Schuster, and J. Knol, "A synbiotic intervention modulates meta-omics signatures of gut redox potential and acidity in elective caesarean born infants," *BMC Microbiology*, vol. 21, no. 1, pp. 1–17, 2021.

[129] N. van Best, M. W. Hornef, P. H. Savelkoul, and J. Penders, "On the origin of species: Factors shaping the establishment of infant's gut microbiota," *Birth Defects Research Part C - Embryo Today: Reviews*, vol. 105, no. 4, pp. 240–251, 2015.

[130] S. Magnúsdóttir and I. Thiele, "Modeling metabolism of the human gut microbiome," *Current Opinion in Biotechnology*, vol. 51, pp. 90–96, 2018.

[131] S. H. J. Chan, M. N. Simons, and C. D. Maranas, "SteadyCom: Predicting microbial abundances while ensuring community stability," *PLoS Computational Biology*, vol. 13, no. 5, pp. 1–25, 2017.

[132] I. Dukovski, D. Bajić, J. M. Chacón, M. Quintin, J. C. C. Vila, S. Sulheim, A. R. Pacheco, D. B. Bernstein, W. J. Riehl, K. S. Korolev, A. Sanchez, W. R. Harcombe, and D. Segrè, "A metabolic modeling platform for the computation of microbial ecosystems in time and space (COMETS)," *Nature Protocols*, vol. 16, pp. 5030–5082, nov 2021.

[133] Q. K. Beg, A. Vazquez, J. Ernst, M. A. De Menezes, Z. Bar-Joseph, A. L. Barabási, and Z. N. Oltvai, "Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 31, pp. 12663–12668, 2007.

[134] A. Monteagudo-Mera, J. C. Arthur, C. Jobin, T. Keku, J. M. Bruno-Barcena, and M. A. Azcarate-Peril, "High purity galacto-oligosaccharides enhance specific Bifidobacterium species and their metabolic activity in the mouse gut microbiome," *Beneficial Microbes*, vol. 7, no. 2, pp. 247–264, 2016.

[135] J. Valentin and C. Streffer, "Basic anatomical and physiological data for use in radiological protection: Reference values - ICRP Publication 89," *Annals of the ICRP*, vol. 32, no. 3-4, pp. 1–277, 2002.

[136] W. S. Davis, R. Parker Allen, B. E. Favara, and T. L. Slovis, "Neonatal small left colon syndrome," *American journal of Roentgenology*, vol. 120, no. 2, pp. 322–329, 1974.

[137] A. Swidsinski, V. Loening-Baucke, H. Verstraelen, S. Osowska, and Y. Doerffel, "Biostructure of Fecal Microbiota in Healthy Subjects and Patients With Chronic Idiopathic Diarrhea," *Gastroenterology*, vol. 135, no. 2, pp. 568–579, 2008.

[138] P. Y. Hong, J. A. Croix, E. Greenberg, H. R. Gaskins, and R. I. Mackie, "Pyrosequencing-based analysis of the mucosal microbiota in healthy individuals reveals ubiquitous bacterial groups and micro-heterogeneity," *PLoS ONE*, vol. 6, no. 9, 2011.

[139] D. Delsing, C. Lindner, H. van Dijck, E. Looijesteijn, E. S. Lima, M. Heerikhuizen, V. Agamennone, V. Triantis, and I. Bovee-Oudenhoven, "Gut microbiota modulating effects of GOS and 2'-FL studied by in vitro infant and toddler faecal fermentations," *ESPGHAN Conf. Abstr. N-P-006*, 2019.

[140] R. Schuetz, L. Kuepfer, and U. Sauer, "Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli," *Molecular Systems Biology*, vol. 3, no. 119, 2007.

[141] F. F. Rubaltelli and G. Largajolli, "Effect of Light Exposure on Gut Transit Time in Jaundiced Newborns," *Acta Pœdiatrica*, vol. 62, no. 2, pp. 146–148, 1973.

[142] J. S. Hyams, M. A. Geertsma, N. L. Etienne, and W. R. Treem, "Colonic hydrogen production in infants with colic," *The Journal of Pediatrics*, vol. 115, no. 4, pp. 592–594, 1989.

[143] S. M. Lee, G. P. Donaldson, Z. Mikulski, S. Boyajian, K. Ley, and S. K. Mazmanian, "Bacterial colonization factors control specificity and stability of the gut microbiota," *Nature*, vol. 501, no. 7467, pp. 426–429, 2013.

[144] M. J. van Hoek and R. M. H. Merks, "Redox balance is key to explaining full vs. partial switching to low-yield metabolism," *BMC Systems Biology*, vol. 6, 2012.

[145] W. de Vries and A. H. Stouthamer, "Fermentation of glucose, lactose, galactose, mannitol, and xylose by bifidobacteria.," *Journal of Bacteriology*, vol. 96, no. 2, pp. 472–478, 1968.

[146] S. F. Kotarski and A. A. Salyers, "Effect of long generation times on growth of Bacteroides thetaiotaomicron in carbohydrate-limited continuous culture," *Journal of Bacteriology*, vol. 146, no. 3, pp. 853–860, 1981.

[147] F. P. J. Martin, S. Moco, I. Montoliu, S. Collino, L. Da Silva, S. Rezzi, R. Prieto, M. Kussmann, J. Inostroza, and P. Steenhout, "Impact of breast-feeding and high-and low-protein formula on the metabolism and growth of infants from overweight and obese mothers," *Pediatric Research*, vol. 75, no. 4, pp. 535–543, 2014.

[148] C. Schwab, H. J. Ruscheweyh, V. Bunesova, V. T. Pham, N. Beerenwinkel, and C. Lacroix, "Trophic interactions of infant bifidobacteria and eubacterium hallii during L-fucose and fucosyllactose degradation," *Frontiers in Microbiology*, vol. 8, no. JAN, pp. 1–14, 2017.

[149] M. J. Wolin, S. Yerry, T. L. Miller, Y. Zhang, and S. Bank, "Changes in Production of Ethanol, Acids and H2 from Glucose by the Fecal Flora of a 16- to 158-d-Old Breast-Fed Infant," *The Journal of Nutrition*, vol. 128, no. 1, pp. 85–90, 1998.

[150] S. L. Bridgman, M. B. Azad, C. J. Field, A. M. Haqq, A. B. Becker, P. J. Mandhane, P. Subbarao, S. E. Turvey, M. R. Sears, J. A. Scott, D. S. Wishart, and A. L. Kozyrskyj, "Fecal Short-Chain Fatty Acid Variations by Breastfeeding Status in Infants at 4 Months: Differences in Relative versus Absolute Concentrations," *Frontiers in Nutrition*, vol. 4, no. April, 2017.

[151] W. de Vries and A. H. Stouthamer, "Pathway of glucose fermentation in relation to the taxonomy of bifidobacteria.," *Journal of bacteriology*, vol. 93, no. 2, pp. 574–576, 1967.

[152] L. Albenberg, T. V. Esipova, C. P. Judge, K. Bittinger, J. Chen, A. Laughlin, S. Grunberg, R. N. Baldassano, J. D. Lewis, H. Li, S. R. Thom, F. D. Bushman, S. A. Vinogradov, and G. D. Wu, "Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota," *Gastroenterology*, vol. 147, no. 5, pp. 1055–1063.e8, 2014.

[153] L. Zheng, C. J. Kelly, and S. P. Colgan, "Physiologic hypoxia and oxygen homeostasis in the healthy intestine. A review in the theme: Cellular responses to hypoxia," *American Journal of Physiology - Cell Physiology*, vol. 309, no. 6, pp. C350–C360, 2015.

[154] K. Sim, A. G. Shaw, P. Randell, M. J. Cox, Z. E. McClure, M. S. Li, M. Haddad, P. R. Langford, W. O. Cookson, M. F. Moffatt, and J. S. Kroll, "Dysbiosis anticipating necrotizing enterocolitis in very premature infants," *Clinical Infectious Diseases*, vol. 60, no. 3, pp. 389–397, 2015.

[155] M. Mori, T. Hwa, O. C. Martin, A. De Martino, and E. Marinari, "Constrained Allocation Flux Balance Analysis," *PLoS Computational Biology*, vol. 12, no. 6, pp. 1–24, 2016.

[156] R. Balakrishnan, R. T. Silva, T. Hwa, and J. Cremer, "Suboptimal resource allocation in changing environments constrains response and growth in bacteria," *Molecular Systems Biology*, vol. 17, no. 12, pp. 1–13, 2021.

[157] E. M. Stansbridge, V. Walker, M. A. Hall, S. L. Smith, M. R. Millar, C. Bacon, and S. Chen, "Effects of feeding premature infants with Lactobacillus GG on gut fermentation," *Archives of Disease in Childhood*, vol. 69, no. 5 SPEC NO, pp. 488–492, 1993.

[158] G. Falony, K. Lazidou, A. Verschaeren, S. Weckx, D. Maes, and L. De Vuyst, "In vitro kinetic analysis of fermentation of prebiotic inulin-type fructans by Bifidobacterium species reveals four different phenotypes," *Applied and Environmental Microbiology*, vol. 75, no. 2, pp. 454–461, 2009.

[159] J. Cremer, M. Arnoldini, and T. Hwa, "Effect of water flow and chemical environment on microbiota growth and composition in the human colon," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 25, pp. 6438–6443, 2017.

[160] J. Cremer, I. Segota, C. Y. Yang, M. Arnoldini, J. T. Sauls, Z. Zhang, E. Gutierreza, A. Groisman, and T. Hwa, "Effect of flow and peristaltic mixing on bacterial growth in a gut-like channel," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 41, pp. 11414–11419, 2016.

[161] S. K. Sarna, *Colonic Motility From Bench Side to Bedside*. San Rafael(CA): Morgan&Claypool Life Sciences, 2010.

[162] S. Nurko, "Motility of the Colon and Anorectum," *NeoReviews*, vol. 7, no. 1, pp. e34–e48, 2006.

[163] M. Juntunen, P. V. Kirjavainen, A. C. Ouwehand, S. J. Salminen, and E. Isolauri, "Adherence of probiotic bacteria to human intestinal mucus in healthy infants and during rotavirus infection," *Clinical and Diagnostic Laboratory Immunology*, vol. 8, no. 2, pp. 293–296, 2001.

[164] F. Turroni, C. Peano, D. A. Pass, E. Foroni, M. Severgnini, M. J. Claesson, C. Kerr, J. Hourihane, D. Murray, F. Fuligni, M. Gueimonde, A. Margolles, G. de Bellis, P. W. O'Toole, D. van Sinderen, J. R. Marchesi, and M. Ventura, "Diversity of bifidobacteria within the infant gut microbiota," *PLoS ONE*, vol. 7, no. 5, pp. 20–24, 2012.

[165] S. Magnúsdóttir, A. Heinken, R. M. T. Fleming, and I. Thiele, "Reply to "Challenges in modeling the human gut microbiome"," *Nature Biotechnology*, vol. 36, no. 8, pp. 686–691, 2018.

[166] R. Schuetz, L. Kuepfer, and U. Sauer, "Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli," *Molecular Systems Biology*, vol. 3, no. 119, 2007.

[167] B. Sánchez, C. G. De Los Reyes-Gavilán, and A. Margolles, "The F1F0-ATPase of Bifidobacterium animalis is involved in bile tolerance," *Environmental Microbiology*, vol. 8, no. 10, pp. 1825–1833, 2006.

[168] F. Krzewinski, C. Brassart, F. Gavini, and S. Bouquelet, "Characterization of the lactose transport system in the strain Bifidobacterium bifidum DSM 20082," *Current Microbiology*, vol. 32, no. 6, pp. 301–307, 1996.

[169] D. A. Sela, A. Adeuya, J. H. Kim, T. R. Whitehead, J. Chapman, D. S. Rokhsar, P. M. Richardson, A. Lapidus, J. B. German, N. P. Price, F. Chen, C. B. Lebrilla, and D. A. Mills, "The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome," *Proceedings of the National Academy of Sciences*, vol. 105, no. 48, pp. 18964–18969, 2008.

[170] C. Lessard, A. Cochu, J. D. Lemay, D. Roy, K. Vaillancourt, M. Frenette, S. Moineau, and C. Vadeboncoeur, "Phosphorylation of Streptococcus salivarius Lactose Permease (LacS) by HPr(His!P) and HPr(Ser-P)(His!P) and Effects on Growth," *Journal of Bacteriology*, vol. 185, no. 23, pp. 6764–6772, 2003.

[171] L. V. Holdeman, "Emendation of Bacteroidaceae and Butyrivibrio and Descriptions of Desulfornonas gen. nov. and Ten New Species in the Genera Desulfomonas, Butyrivibrio, Eubacterium, Clostridium, and Ruminococcus," *International Journal of Systematic Bacteriology*, vol. 26, no. 2, pp. 238–252, 1976.

[172] A. Flamholz, E. Noor, A. Bar-Even, and R. Milo, "EQuilibrator - The biochemical thermodynamics calculator," *Nucleic Acids Research*, vol. 40, no. D1, pp. 770–775, 2012.

[173] E. Noor, H. S. Haraldsdóttir, R. Milo, and R. M. Fleming, "Consistent Estimation of Gibbs Energy Using Component Contributions," *PLoS Computational Biology*, vol. 9, no. 7, 2013.

[174] P. S. Stewart, "Diffusion in biofilms," *Journal of Bacteriology*, vol. 185, no. 5, pp. 1485–1491, 2003.

[175] L. M. Grummer-Strawn, K. S. Scanlon, and S. B. Fein, "Infant feeding and feeding transitions during the first year of life," *Pediatrics*, vol. 122, no. SUPPL. 2, 2008.

[176] K. Kawasaki, "Diffusion constants near the critical point for time-dependent ising models. I," *Physical Review*, vol. 145, no. 1, pp. 224–230, 1966.

[177] N. Koumakis, C. Devailly, and W. C. Poon, "Motile bacteria in a critical fluid mixture," *Physical Review E*, vol. 97, jun 2018.

[178] P. O. Anderson and V. Valdés, "Variation of milk intake over time: Clinical and pharmacokinetic implications," *Breastfeeding Medicine*, vol. 10, no. 3, pp. 142–144, 2015.

[179] B. Gibson, D. J. Wilson, E. Feil, and A. Eyre-Walker, "The distribution of bacterial doubling times in the wild," *Proceedings of the Royal Society B: Biological Sciences*, vol. 285, no. 1880, 2018.

[180] F. Turroni, C. Milani, S. Duranti, G. A. Lugli, S. Bernasconi, A. Margolles, F. Di Pierro, D. Van Sinderen, and M. Ventura, "The infant gut microbiome as a microbial organ influencing host well-being," *Italian Journal of Pediatrics*, vol. 46, no. 1, pp. 1–13, 2020.

[181] S. Plöger, F. Stumpff, G. B. Penner, J. D. Schulzke, G. Gäbel, H. Martens, Z. Shen, D. Günzel, and J. R. Aschenbach, "Microbial butyrate and its role for barrier function in the gastrointestinal tract," *Annals of the New York Academy of Sciences*, vol. 1258, no. 1, pp. 52–59, 2012.

[182] X. Wu, Y. Wu, L. He, L. Wu, X. Wang, and Z. Liu, "Effects of the intestinal microbial metabolite butyrate on the development of colorectal cancer," *Journal of Cancer*, vol. 9, no. 14, pp. 2510–2517, 2018.

[183] T. Feehley, C. H. Plunkett, R. Bao, S. M. Choi Hong, E. Culleen, P. Belda-Ferre, E. Campbell, R. Aitoro, R. Nocerino, L. Paparo, J. Andrade, D. A. Antonopoulos, R. Berni Canani, and C. R. Nagler, "Healthy infants harbor intestinal bacteria that protect against food allergy," *Nature Medicine*, vol. 25, no. 3, pp. 448–453, 2019.

[184] H. M. Hamer, D. Jonkers, K. Venema, S. Vanhoutvin, F. J. Troost, and R. J. Brummer, "Review article: The role of butyrate on colonic function," *Alimentary Pharmacology and Therapeutics*, vol. 27, no. 2, pp. 104–119, 2008.

[185] G. Moro, S. Arslanoglu, B. Stahl, J. Jelinek, U. Wahn, and G. Boehm, "A mixture of prebiotic oligosaccharides reduces the incidence of atopic dermatitis during the first six months of age," *Archives of Disease in Childhood*, vol. 91, no. 10, pp. 814–819, 2006.

[186] N. Sprenger, H. Odenwald, A. K. Kukkonen, M. Kuitunen, E. Savilahti, and C. Kunz, "FUT2-dependent breast milk oligosaccharides and allergy at 2 and 5 years of age in infants with high hereditary allergy risk," *European Journal of Nutrition*, vol. 56, no. 3, pp. 1293–1301, 2017.

[187] K. C. Goehring, B. J. Marriage, J. S. Oliver, J. A. Wilder, E. G. Barrett, and R. H. Buck, "Similar to those who are breastfed, infants fed a formula containing 2'-fucosyllactose have lower inflammatory cytokines in a randomized controlled trial," *Journal of Nutrition*, vol. 146, no. 12, pp. 2559–2566, 2016.

[188] P. Van den Abbeele, C. Duysburgh, E. Vazquez, J. Chow, R. Buck, and M. Marzorati, "2'-Fucosyllactose alters the composition and activity of gut microbiota from formula-fed infants receiving complementary feeding in a validated intestinal model," *Journal of Functional Foods*, vol. 61, no. July, p. 103484, 2019.

[189] S. A. Shetty, S. Zuffa, T. P. N. Bui, S. Aalvink, H. Smidt, and W. M. De Vos, "Reclassification of eubacterium hallii as Anaerobutyricum hallii gen. nov., comb. nov., and description of Anaerobutyricum soehngenii sp. nov., a butyrate

and propionate-producing bacterium from infant faeces," *International Journal of Systematic and Evolutionary Microbiology*, vol. 68, no. 12, pp. 3741–3746, 2018.

[190] M. J. Gnoth, C. Kunz, E. Kinne-Saffran, and S. Rudloff, "Human milk oligosaccharides are minimally digested in vitro," *Journal of Nutrition*, vol. 130, no. 12, pp. 3014–3020, 2000.

[191] Y. Song, C. Liu, and S. M. Finegold, "Bacteroides," in *Bergey's Manual of Systematics of Archaea and Bacteria*, pp. 1–24, Wiley, sep 2015.

[192] J. M. Macy, L. G. Ljungdahl, and G. Gottschalk, "Pathway of succinate and propionate formation in Bacteroides fragilis," *Journal of Bacteriology*, vol. 134, no. 1, pp. 84–91, 1978.

[193] C. F. Scholz and M. Kilian, "The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus propionibacterium to the proposed novel genera acidipropionibacterium gen. Nov., cutibacterium gen. nov. and pseudopropionibacterium gen. nov.," *International Journal of Systematic and Evolutionary Microbiology*, vol. 66, no. 11, pp. 4422–4432, 2016.

[194] V. T. Pham, C. Chassard, E. Rifa, C. Braegger, A. Geirnaert, V. N. Rocha Martin, and C. Lacroix, "Lactate Metabolism Is Strongly Modulated by Fecal Inoculum, pH, and Retention Time in PolyFermS Continuous Colonic Fermentation Models Mimicking Young Infant Proximal Colon," *mSystems*, vol. 4, no. 4, 2019.

[195] M. Goodfellow, P. Kämpfer, H.-J. Busse, M. E. Trujillo, K.-i. Suzuki, W. Ludwig, and W. B. Whitman, *Bergey's Manual of Systematic Bacteriology: Volume Five The Actinobacteria, Part A*. Springer, 2012.

[196] P. S. La Rosa, B. B. Warner, Y. Zhou, G. M. Weinstock, E. Sodergren, C. M. Hall-Moore, H. J. Stevens, W. E. Bennett, N. Shaikh, L. A. Linneman, J. A. Hoffmann, A. Hamvas, E. Deych, B. A. Shands, W. D. Shannon, and P. I. Tarr, "Patterned progression of bacterial populations in the premature infant gut," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 34, pp. 12522–12527, 2014.

[197] A. Mitchell, G. H. Romano, B. Groisman, A. Yona, E. Dekel, M. Kupiec, O. Dahan, and Y. Pilpel, "Adaptive prediction of environmental changes by microorganisms," *Nature*, vol. 460, no. 7252, pp. 220–224, 2009.

[198] L. Heirendt, S. Arreckx, T. Pfau, S. N. Mendoza, A. Richelle, A. Heinken, H. S. Haraldsdóttir, J. Wachowiak, S. M. Keating, V. Vlasov, S. Magnusdóttir, C. Y. Ng, G. Preciat, A. Žagare, S. H. J. Chan, M. K. Aurich, C. M. Clancy, J. Modamio, J. T. Sauls, A. Noronha, A. Bordbar, B. Cousins, D. C. El Assal, L. V. Valcarcel, I. Apaolaza, S. Ghaderi, M. Ahookhosh, M. Ben

Guebila, A. Kostromins, N. Sompairac, H. M. Le, D. Ma, Y. Sun, L. Wang, J. T. Yurkovich, M. A. P. Oliveira, P. T. Vuong, L. P. El Assal, I. Kuperstein, A. Zinovyev, H. S. Hinton, W. A. Bryant, F. J. Aragón Artacho, F. J. Planes, E. Stalidzans, A. Maass, S. Vempala, M. Hucka, M. A. Saunders, C. D. Maranas, N. E. Lewis, T. Sauter, B. Ø. Palsson, I. Thiele, and R. M. T. Fleming, "Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0," *Nature Protocols*, vol. 14, pp. 639–702, mar 2019.

[199] R. M. Fleming, I. Thiele, and H. P. Nasheuer, "Quantitative assignment of reaction directionality in constraint-based models of metabolism: Application to Escherichia coli," *Biophysical Chemistry*, vol. 145, no. 2-3, pp. 47–56, 2009.

[200] C. Rao, K. Z. Coyte, W. Bainter, R. S. Geha, C. R. Martin, and S. Rakoff-Nahoum, "Multi-kingdom ecological drivers of microbiota assembly in preterm infants," *Nature*, vol. 591, no. 7851, pp. 633–638, 2021.

[201] G. A. Dykes and J. W. Hastings, "Selection and fitness in bacteriocin-producing bacteria," *Proceedings of the Royal Society B: Biological Sciences*, vol. 264, no. 1382, pp. 683–687, 1997.

[202] S. Mills, F. Shanahan, C. Stanton, C. Hill, A. Coffey, and R. P. Ross, "Movers and shakers Influence of bacteriophages in shaping the mammalian gut microbiota," *Gut Microbes*, vol. 4, pp. 4–16, jan 2013.

[203] P. Mastromarino, D. Capobianco, G. Campagna, N. Laforgia, P. Drimaco, A. Dileone, and M. E. Baldassarre, "Correlation between lactoferrin and beneficial microbiota in breast milk and infant's feces," *BioMetals*, vol. 27, no. 5, pp. 1077–1086, 2014.

[204] P. Louis and H. J. Flint, "Formation of propionate and butyrate by the human colonic microbiota," *Environmental Microbiology*, vol. 19, no. 1, pp. 29–41, 2017.

[205] J. A. Vogt and T. M. S. Wolever, "Fecal Acetate Is Inversely Related to Acetate Absorption from the Human Rectum and Distal Colon," *American Society for Nutritional Sciences*, vol. 133, no. 10, pp. 3145–3148, 2003.

[206] S. S. Van Leeuwen, B. J. Kuipers, L. Dijkhuizen, and J. P. Kamerling, "1 H NMR analysis of the lactose/β-galactosidase-derived galacto-oligosaccharide components of Vivinal® GOS up to DP5," *Carbohydrate Research*, vol. 400, pp. 59–73, 2014.

[207] K. M. Van Laere, T. Abee, H. A. Schols, G. Beldman, and A. G. Voragen, "Characterization of a novel β-galactosidase from Bifidobacterium adolescentis DSM 20083 active towards transgalactooligosaccharides," *Applied and Environmental Microbiology*, vol. 66, no. 4, pp. 1379–1384, 2000.

[208] D. P. Clark, "The fermentation pathways of Escherichia coli," *FEMS Microbiology Letters*, vol. 63, no. 3, pp. 223–234, 1989.

[209] S. Parche, M. Beleut, E. Rezzonico, D. Jacobs, F. Arigoni, F. Titgemeyer, and I. Jankovic, "Lactose-over-Glucose Preference in Bifidobacterium longum NCC2705: glcP , Encoding a Glucose Transporter, Is Subject to Lactose Repression," *Journal of Bacteriology*, vol. 188, pp. 1260–1265, feb 2006.

[210] M. J. Pichler, C. Yamada, B. Shuoker, C. Alvarez-Silva, A. Gotoh, M. L. Leth, E. Schoof, T. Katoh, M. Sakanaka, T. Katayama, C. Jin, N. G. Karlsson, M. Arumugam, S. Fushinobu, and M. Abou Hachem, "Butyrate producing colonic Clostridiales metabolise human milk oligosaccharides and cross feed on mucin via conserved pathways," *Nature Communications*, vol. 11, no. 1, 2020.

[211] K. A. Hamed, P. R. Dormitzer, C. K. Su, and D. A. Relman, "Haemophilus parainfluenzae endocarditis: application of a molecular approach for identification of pathogenic bacterial species," *Clinical Infectious Diseases*, vol. 19, no. 4, pp. 677–683, 1994.

[212] C. C. Uranga, P. Arroyo, B. M. Duggan, W. H. Gerwick, and A. Edlund, "Commensal Oral Rothia mucilaginosa Produces Enterobactin, a Metal-Chelating Siderophore," *mSystems*, vol. 5, no. 2, 2020.

[213] E. T. Hillman, A. J. Kozik, C. A. Hooker, J. L. Burnett, Y. Heo, V. A. Kiesel, C. J. Nevins, J. M. Oshiro, M. M. Robins, R. D. Thakkar, S. T. Wu, and S. R. Lindemann, "Comparative genomics of the genus Roseburia reveals divergent biosynthetic pathways that may influence colonic competition among species," *Microbial Genomics*, vol. 6, no. 7, pp. 7–24, 2020.

[214] S. I. Yokoyama and T. Suzuki, "Isolation and characterization of a novel equol-producing bacterium from human feces," *Bioscience, Biotechnology and Biochemistry*, vol. 72, no. 10, pp. 2660–2666, 2008.

[215] P. Vos, G. Garrity, D. Jones, N. R. Krieg, W. Ludwig, F. A. Rainey, K.-H. Schleifer, and W. B. Whitman, *Bergey's manual of systematic bacteriology: Volume 3: The Firmicutes*, vol. 3. Springer Science & Business Media, 2011.

[216] M. Allen, D. Poggiali, K. Whitaker, T. R. Marshall, and R. A. Kievit, "Raincloud plots: A multi-platform tool for robust data visualization [version 1; peer review: 2 approved]," *Wellcome Open Research*, vol. 4, pp. 1–52, 2019.

[217] J. Fu, B. Wei, T. Wen, M. E. V. Johansson, X. Liu, E. Bradford, K. a. Thomsson, S. McGee, L. Mansour, M. Tong, J. M. McDaniel, T. J. Sferra, J. R. Turner, H. Chen, G. C. Hansson, J. Braun, and L. Xia, "Loss of intestinal core 1-derived O-glycans causes spontaneous colitis in mice.," *The Journal of clinical investigation*, vol. 121, pp. 1657–66, apr 2011.

[218] R. H. Yolken, C. Ojeh, I. A. Khatri, U. Sajjan, and J. F. Forstner, "Intestinal Mucins Inhibit Rotavirus Replication in an Oligosaccharide-Dependent Manner," *Journal of Infectious Diseases*, vol. 169, pp. 1002–1006, may 1994.

[219] A.-C. Midtvedt, B. Carlstedt-Duke, and T. Midtvedt, "Establishment of a Mucin-Degrading Intestinal Microflora During the First Two Years of Human Life," *Journal of Pediatric Gastroenterology and Nutrition*, vol. 18, pp. 321–326, apr 1994.

[220] M. Sakanaka, A. Gotoh, K. Yoshida, T. Odamaki, H. Koguchi, J. Z. Xiao, M. Kitaoka, and T. Katayama, "Varied pathways of infant gut-associated Bifidobacterium to assimilate human milk oligosaccharides: Prevalence of the gene set and its correlation with bifidobacteria-rich microbiota formation," *Nutrients*, vol. 12, no. 1, pp. 1–21, 2020.

[221] M. K. McGuire, C. L. Meehan, M. A. McGuire, J. E. Williams, J. Foster, D. W. Sellen, E. W. Kamau-Mbuthia, E. W. Kamundia, S. Mbugua, S. E. Moore, A. M. Prentice, L. J. Kvist, G. E. Otoo, S. L. Brooker, W. J. Price, B. Shafii, C. Placek, K. A. Lackey, B. Robertson, S. Manzano, L. Ruíz, J. M. Rodríguez, R. G. Pareja, and L. Bode, "What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically," *American Journal of Clinical Nutrition*, vol. 105, no. 5, pp. 1086–1100, 2017.

[222] D. Garrido, S. Ruiz-Moyano, D. G. Lemay, D. A. Sela, J. B. German, and D. A. Mills, "Comparative transcriptomics reveals key differences in the response to milk oligosaccharides of infant gut-associated bifidobacteria," *Scientific Reports*, vol. 5, no. May, pp. 1–18, 2015.

[223] D. Garrido, S. Ruiz-Moyano, R. Jimenez-Espinoza, H. J. Eom, D. E. Block, and D. A. Mills, "Utilization of galactooligosaccharides by Bifidobacterium longum subsp. infantis isolates," *Food Microbiology*, vol. 33, no. 2, pp. 262–270, 2013.

[224] M. N. Ojima, L. Jiang, A. A. Arzamasov, K. Yoshida, T. Odamaki, J. Xiao, A. Nakajima, M. Kitaoka, J. Hirose, T. Urashima, T. Katoh, A. Gotoh, D. van Sinderen, D. A. Rodionov, A. L. Osterman, M. Sakanaka, and T. Katayama, "Priority effects shape the structure of infant-type Bifidobacterium communities on human milk oligosaccharides," *ISME Journal*, vol. 16, no. 9, pp. 2265–2279, 2022.

[225] Z. T. Lewis, S. M. Totten, J. T. Smilowitz, M. Popovic, E. Parker, D. G. Lemay, M. L. Van Tassell, M. J. Miller, Y. S. Jin, J. B. German, C. B. Lebrilla, and D. A. Mills, "Maternal fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants," *Microbiome*, vol. 3, no. 1, pp. 15–17, 2015.

[226] P. Thomson, D. A. Medina, and D. Garrido, "Human milk oligosaccharides and infant gut bifidobacteria: Molecular strategies for their utilization," *Food Microbiology*, vol. 75, pp. 37–46, 2018.

[227] J. Cremer, A. Melbinger, K. Wienand, T. Henriquez, H. Jung, and E. Frey, "Cooperation in Microbial Populations: Theory and Experimental Model Systems," *Journal of Molecular Biology*, vol. 431, no. 23, pp. 4599–4644, 2019.

[228] H. Celiker and J. Gore, "Competition between species can stabilize public-goods cooperation within a species," *Molecular Systems Biology*, vol. 8, no. 621, pp. 1–9, 2012.

[229] D. M. Versluis, R. Schoemaker, E. Looijesteijn, J. M. W. Geurts, and R. M. H. Merks, "2'-Fucosyllactose helps butyrate producers outgrow competitors in infant gut microbiota simulations," *iScience*, vol. 27, 2024.

[230] P. L. Møller, F. Jørgensen, O. C. Hansen, S. M. Madsen, and P. Stougaard, "Intra- and Extracellular β-Galactosidases from Bifidobacterium bifidum and B. infantis: Molecular Cloning, Heterologous Expression, and Comparative Characterization," *Applied and Environmental Microbiology*, vol. 67, no. 5, pp. 2276–2283, 2001.

[231] M. O'Connell Motherway, M. Kinsella, G. F. Fitzgerald, and D. Van Sinderen, "Transcriptional and functional characterization of genetic elements involved in galacto-oligosaccharide utilization by Bifidobacterium breve UCC2003," *Microbial Biotechnology*, vol. 6, no. 1, pp. 67–79, 2013.

[232] R. J. Kelly, S. Rouquier, D. Giorgi, G. G. Lennon, and J. B. Lowe, "Sequence and expression of a candidate for the human Secretor blood group α(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme- inactivating nonsense mutation commonly correlates with the non-secretor phenotype," *Journal of Biological Chemistry*, vol. 270, no. 9, pp. 4640–4649, 1995.

[233] W. Holms, "The Central Metabolic Pathways of Escherichia coli: Relationship between Flux and Control at a Branch Point, Efficiency of Conversion to Biomass, and Excretion of Acetate," *Current Topics in Cellular Regulation*, pp. 69–105, 1986.

[234] A. L. Francl, J. L. Hoeflinger, and M. J. Miller, "Identification of lactose phosphotransferase systems in Lactobacillus gasseri ATCC 33323 required for lactose utilization," *Microbiology*, vol. 158, no. 4, pp. 944–952, 2012.

[235] A. Gotoh, T. Katoh, M. Sakanaka, Y. Ling, C. Yamada, S. Asakuma, T. Urashima, Y. Tomabechi, A. Katayama-Ikegami, S. Kurihara, K. Yamamoto, G. Harata, F. He, J. Hirose, M. Kitaoka, S. Okuda, and T. Katayama, "Sharing of human milk oligosaccharides degradants within bifidobacterial communities in faecal cultures supplemented with Bifidobacterium bifidum," *Scientific Reports*, vol. 8, no. 1, pp. 1–14, 2018.

[236] F. Turroni, C. Milani, S. Duranti, L. Mancabelli, M. Mangifesta, A. Viappiani, G. A. Lugli, C. Ferrario, L. Gioiosa, A. Ferrarini, J. Li, P. Palanza,

M. Delledonne, D. Van Sinderen, and M. Ventura, "Deciphering bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics approach," *ISME Journal*, vol. 10, no. 7, pp. 1656–1668, 2016.

[237] J. P. H. Wong, M. Fischer-Stettler, S. C. Zeeman, T. J. Battin, and A. Persat, "Fluid flow structures gut microbiota biofilm communities by distributing public goods," *Proceedings of the National Academy of Sciences*, vol. 120, p. 2017, jun 2023.

[238] R. F. Inglis, J. M. Biernaskie, A. Gardner, and R. Kümmerli, "Presence of a loner strain maintains cooperation and diversity in well-mixed bacterial communities," *Proceedings of the Royal Society B: Biological Sciences*, vol. 283, no. 1822, 2016.

[239] R. J. Lindsay, B. J. Pawlowska, and I. Gudelj, "Privatization of public goods can cause population decline," *Nature Ecology and Evolution*, vol. 3, no. 8, pp. 1206–1216, 2019.

[240] A. M. Feist and B. O. Palsson, "The biomass objective function," *Current Opinion in Microbiology*, vol. 13, no. 3, pp. 344–349, 2010.

[241] C. Belzer, L. W. Chia, S. Aalvink, B. Chamlagain, V. Piironen, J. Knol, and W. M. de Vos, "Microbial metabolic networks at the mucus layer lead to diet-independent butyrate and vitamin B12 production by intestinal symbionts," *mBio*, vol. 8, no. 5, 2017.

[242] J. M. Hascoët, C. Hubert, F. Rochat, H. Legagneur, S. Gaga, S. Emady-Azar, and P. G. Steenhout, "Effect of formula composition on the development of infant gut microbiota," *Journal of Pediatric Gastroenterology and Nutrition*, vol. 52, no. 6, pp. 756–762, 2011.

[243] A. P. Corfield, "The interaction of the gut microbiota with the mucus barrier in health and disease in human," *Microorganisms*, vol. 6, no. 3, pp. 18–26, 2018.

[244] S. Thurl, M. Munzert, J. Henker, G. Boehm, B. Mller-Werner, J. Jelinek, and B. Stahl, "Variation of human milk oligosaccharides in relation to milk groups and lactational periods," *British Journal of Nutrition*, vol. 104, no. 9, pp. 1261–1271, 2010.

[245] T. Shiba, Y. Aiba, H. Ishikawa, A. Ushiyama, A. Takagi, T. Mine, and Y. Koga, "The suppressive effect of bifidobacteria on Bacteroides vulgatus, a putative pathogenic microbe in inflammatory bowel disease," *Microbiology and Immunology*, vol. 47, no. 6, pp. 371–378, 2003.

[246] Y. Belkaid and T. W. Hand, "Role of the microbiota in immunity and inflammation," *Cell*, vol. 157, no. 1, pp. 121–141, 2014.

[247] E. Klement, R. V. Cohen, J. Boxman, A. Joseph, and S. Reif, "Breastfeeding and risk of inflammatory bowel disease: a systematic review with meta-analysis," *The American Journal of Clinical Nutrition*, vol. 80, pp. 1342–1352, nov 2004.

[248] Y. Kang, H. Park, B. H. Choe, and B. Kang, "The Role and Function of Mucins and Its Relationship to Inflammatory Bowel Disease," *Frontiers in Medicine*, vol. 9, no. May, pp. 1–7, 2022.

[249] I. Azagra-Boronat, M. Massot-Cladera, J. Mayneris-Perxachs, K. Knipping, B. Van't Land, S. Tims, B. Stahl, J. Garssen, À. Franch, M. Castell, M. J. Rodríguez-Lagunas, and F. J. Pérez-Cano, "Immunomodulatory and Prebiotic Effects of 2'-Fucosyllactose in Suckling Rats," *Frontiers in immunology*, vol. 10, no. July, p. 1773, 2019.

[250] C. Agostoni, E. Riva, and M. Giovannini, "Dietary fiber in weaning foods of young children," *Pediatrics*, vol. 96, no. 5 II SUPPL., pp. 1002–1005, 1995.

[251] M. Sakanaka, M. E. Hansen, A. Gotoh, T. Katoh, K. Yoshida, T. Odamaki, H. Yachi, Y. Sugiyama, S. Kurihara, J. Hirose, T. Urashima, J. zhong Xiao, M. Kitaoka, S. Fukiya, A. Yokota, L. L. Leggio, M. A. Hachem, and T. Katayama, "Evolutionary adaptation in fucosyllactose uptake systems supports bifidobacteria-infant symbiosis," *Science Advances*, vol. 5, no. 8, pp. 1–16, 2019.

[252] H. Ashida, A. Miyake, M. Kiyohara, J. Wada, E. Yoshida, H. Kumagai, T. Katayama, and K. Yamamoto, "Two distinct α-L-fucosidases from Bifidobacterium bifidum are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates," *Glycobiology*, vol. 19, no. 9, pp. 1010–1017, 2009.

[253] D. Szatmári, P. Sárkány, B. Kocsis, T. Nagy, A. Miseta, S. Barkó, B. Longauer, R. C. Robinson, and M. Nyitrai, "Intracellular ion concentrations and cation-dependent remodelling of bacterial MreB assemblies," *Scientific Reports*, vol. 10, no. 1, pp. 1–13, 2020.

[254] J. L. Slonczewski, M. Fujisawa, M. Dopson, and T. A. Krulwich, "Cytoplasmic pH Measurement and Homeostasis in Bacteria and Archaea," *Advances in Microbial Physiology*, vol. 55, 2009.

[255] H. Siegumfeldt, K. Björn Rechinger, and M. Jakobsen, "Use of fluorescence ratio imaging for intracellular pH determination of individual bacterial cells in mixed cultures," *Microbiology*, vol. 145, no. 7, pp. 1703–1709, 1999.

[256] W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing.* Cambridge university press, 2007.

[257] N. D. Rivera-Chaparro, M. Cohen-Wolkowiez, and R. G. Greenberg, "Dosing antibiotics in neonates: Review of the pharmacokinetic data," *Future Microbiology*, vol. 12, no. 11, pp. 1001–1016, 2017.

[258] A. H. Kim, M. P. Hogarty, V. C. Harris, and M. T. Baldridge, "The Complex Interactions Between Rotavirus and the Gut Microbiota," *Frontiers in Cellular and Infection Microbiology*, vol. 10, jan 2021.

[259] M. Pop, A. W. Walker, J. Paulson, B. Lindsay, M. Antonio, M. A. Hossain, J. Oundo, B. Tamboura, V. Mai, I. Astrovskaya, H. C. Bravo, R. Rance, M. Stares, M. M. Levine, S. Panchalingam, K. Kotloff, U. N. Ikumapayi, C. Ebruke, M. Adeyemi, D. Ahmed, F. Ahmed, M. T. Alam, R. Amin, S. Siddiqui, J. B. Ochieng, E. Ouma, J. Juma, E. Mailu, R. Omore, J. G. Morris, R. F. Breiman, D. Saha, J. Parkhill, J. P. Nataro, and O. C. Stine, "Diarrhea in young children from low-income countries leads to large-scale alterations in wintestinal microbiota composition," *Genome Biology*, vol. 15, no. 6, pp. 1–12, 2014.

[260] E. Goksör, B. Alm, R. Pettersson, P. Möllborg, L. Erdes, N. Åberg, and G. Wennergren, "Early fish introduction and neonatal antibiotics affect the risk of asthma into school age," *Pediatric Allergy and Immunology*, vol. 24, no. 4, pp. 339–344, 2013.

[261] B. C. Oosterloo, R. M. van Elburg, N. B. Rutten, C. M. Bunkers, C. E. Crijns, C. B. Meijssen, J. H. Oudshoorn, G. T. Rijkers, C. K. van der Ent, and A. M. Vlieger, "Wheezing and infantile colic are associated with neonatal antibiotic treatment," *Pediatric Allergy and Immunology*, vol. 29, no. 2, pp. 151–158, 2018.

[262] S. Salvatore, M. E. Baldassarre, A. Di Mauro, N. Laforgia, S. Tafuri, F. P. Bianchi, E. Dattoli, L. Morando, L. Pensabene, F. Meneghin, D. Dilillo, V. Mancini, V. Talarico, F. Tandoi, G. Zuccotti, and M. Agosti, "Neonatal Antibiotics and Prematurity Are Associated with an Increased Risk of Functional Gastrointestinal Disorders in the First Year of Life," *Journal of Pediatrics*, vol. 212, pp. 44–51, 2019.

[263] M. F. Endika, D. J. Barnett, C. E. Klostermann, H. A. Schols, I. C. Arts, J. Penders, A. Nauta, H. Smidt, and K. Venema, "Microbiota-dependent influence of prebiotics on the resilience of infant gut microbiota to amoxicillin/clavulanate perturbation in an in vitro colon model," *Frontiers in Microbiology*, vol. 14, no. May, 2023.

[264] M. Reyman, M. A. van Houten, R. L. Watson, M. L. J. Chu, K. Arp, W. J. de Waal, I. Schiering, F. B. Plötz, R. J. Willems, W. van Schaik, E. A. Sanders, and D. Bogaert, "Effects of early-life antibiotics on the developing infant gut microbiome and resistome: a randomized trial," *Nature Communications*, vol. 13, no. 1, 2022.

[265] D. L. Dai, C. Petersen, C. Hoskinson, K. L. Del Bel, A. B. Becker, T. J. Moraes, P. J. Mandhane, B. B. Finlay, E. Simons, A. L. Kozyrskyj, D. M. Patrick, P. Subbarao, L. Bode, M. B. Azad, and S. E. Turvey, "Breastfeeding enrichment of B. longum subsp. infantis mitigates the effect of antibiotics on the microbiota and childhood asthma risk," *Med*, vol. 4, no. 2, pp. 92–112.e5, 2023.

[266] M. K. Gibson, T. S. Crofts, and G. Dantas, "Antibiotics and the developing infant gut microbiota and resistome," *Current Opinion in Microbiology*, vol. 27, pp. 51–56, 2015.

[267] M. F. Laursen, M. I. Bahl, and T. R. Licht, "Settlers of our inner surface-factors shaping the gut microbiota from birth to toddlerhood," *FEMS Microbiology Reviews*, vol. 45, no. 4, pp. 1–14, 2021.

[268] A. T. Reese, E. H. Cho, B. Klitzman, S. P. Nichols, N. A. Wisniewski, M. M. Villa, H. K. Durand, S. Jiang, F. S. Midani, S. N. Nimmagadda, T. M. O'connell, J. P. Wright, M. A. Deshusses, and L. A. David, "Antibiotic-induced changes in the microbiota disrupt redox dynamics in the gut," *eLife*, vol. 7, pp. 1–22, 2018.

[269] S. E. Winter, M. G. Winter, M. N. Xavier, P. Thiennimitr, V. Poon, A. M. Keestra, R. C. Laughlin, G. Gomez, J. Wu, S. D. Lawhon, I. E. Popova, S. J. Parikh, L. G. Adams, R. M. Tsolis, V. J. Stewart, and A. J. Bäumler, "Host-derived nitrate boosts growth of E. coli in the inflamed gut," *Science*, vol. 339, no. 6120, pp. 708–711, 2013.

[270] M. Yassour, T. Vatanen, H. Siljander, A. M. Hämäläinen, T. Härkönen, S. J. Ryhänen, E. A. Franzosa, H. Vlamakis, C. Huttenhower, D. Gevers, E. S. Lander, M. Knip, and R. J. Xavier, "Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability," *Science Translational Medicine*, vol. 8, no. 343, 2016.

[271] A. J. Eherer and J. S. Fordtran, "Fecal osmotic gap and pH in experimental diarrhea of various causes," *Gastroenterology*, vol. 103, no. 2, pp. 545–551, 1992.

[272] K. Shimizu, I. Seiki, Y. Goto, and T. Murata, "Measurement of the intestinal pH in mice under various conditions reveals alkalization induced by antibiotics," *Antibiotics*, vol. 10, no. 2, pp. 1–10, 2021.

[273] K. Faust, F. Bauchinger, B. Laroche, S. de Buyl, L. Lahti, A. D. Washburne, D. Gonze, and S. Widder, "Signatures of ecological processes in microbial community time series," *Microbiome*, vol. 6, no. 1, pp. 1–13, 2018.

[274] I. Ormaasen, K. Rudi, D. B. Diep, and L. Snipen, "Metagenome-mining indicates an association between bacteriocin presence and strain diversity in the infant gut," *BMC Genomics*, vol. 24, no. 1, pp. 1–12, 2023.

[275] K. L. Fields and S. E. Luria, "Effects of Colicins E1 and K on Cellular Metabolism," *Journal of Bacteriology*, vol. 97, pp. 64–77, jan 1969.

[276] T. L. Czárán, R. F. Hoekstra, and L. Pagie, "Chemical warfare between microbes promotes biodiversity," *Proceedings of the National Academy of Sciences*, vol. 99, pp. 786–790, jan 2002.

[277] D. Podlesny and W. F. Fricke, "Strain inheritance and neonatal gut microbiota development: A meta-analysis," *International Journal of Medical Microbiology*, vol. 311, no. 3, p. 151483, 2021.

[278] S. Shimamura, F. Abe, N. Ishibashi, H. Miyakawa, T. Yaeshima, T. Araya, and M. Tomita, "Relationship Between Oxygen Sensitivity and Oxygen Metabolism of Bifidobacterium Species," *Journal of Dairy Science*, vol. 75, no. 12, pp. 3296–3306, 1992.

[279] A. Goelzer, V. Fromion, and G. Scorletti, "Cell design in bacteria as a convex optimization problem," *Automatica*, vol. 47, no. 6, pp. 1210–1218, 2011.

[280] D. Rios Garza, D. Gonze, H. Zafeiropoulos, B. Liu, and K. Faust, "Metabolic models of human gut microbiota: Advances and challenges," *Cell Systems*, vol. 14, no. 2, pp. 109–121, 2023.

[281] S. Labarthe, B. Polizzi, T. Phan, T. Goudon, M. Ribot, and B. Laroche, "A mathematical model to investigate the key drivers of the biogeography of the colon microbiota," *Journal of Theoretical Biology*, vol. 462, pp. 552–581, 2019.

[282] M. Arnoldini, J. Cremer, and T. Hwa, "Bacterial growth, flow, and mixing shape human gut microbiota density and composition," *Gut microbes*, vol. 9, no. 6, pp. 559–566, 2018.

[283] D. Labavić, C. Loverdo, and A. F. Bitbol, "Hydrodynamic flow and concentration gradients in the gut enhance neutral bacterial diversity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 119, no. 1, 2022.

[284] T. Jiang, F. L. Suarez, M. D. Levitt, S. E. Nelson, and E. E. Ziegler, "Gas production by feces of infants," *Journal of Pediatric Gastroenterology and Nutrition*, vol. 32, no. 5, pp. 534–541, 2001.

[285] A. L. Morrow, A. J. Lagomarcino, K. R. Schibler, D. H. Taft, Z. Yu, B. Wang, M. Altaye, M. Wagner, D. Gevers, D. V. Ward, M. A. Kennedy, C. Huttenhower, and D. S. Newburg, "Early microbial and metabolomic signatures predict later onset of necrotizing enterocolitis in preterm infants," *Microbiome*, vol. 1, no. 1, pp. 1–16, 2013.

[286] Z. T. Yu, C. Chen, and D. S. Newburg, "Utilization of major fucosylated and sialylated human milk oligosaccharides by isolated human gut microbes," *Glycobiology*, vol. 23, no. 11, pp. 1281–1292, 2013.

[287] A. Heinken, J. Hertel, G. Acharya, D. A. Ravcheev, M. Nyga, O. E. Okpala, M. Hogan, S. Magnúsdóttir, F. Martinelli, B. Nap, G. Preciat, J. N. Edirisinghe, C. S. Henry, R. M. Fleming, and I. Thiele, "Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine," *Nature Biotechnology*, 2023.

[288] M. E. Beber, M. G. Gollub, D. Mozaffari, K. M. Shebek, A. I. Flamholz, R. Milo, and E. Noor, "EQuilibrator 3.0: A database solution for thermodynamic constant estimation," *Nucleic Acids Research*, vol. 50, no. D1, pp. D603–D609, 2022.

[289] K. Faust and J. Raes, "Rules of the game for microbiota," *Nature*, vol. 534, no. 7606, pp. 182–183, 2016.

[290] M. Ghoul and S. Mitri, "The Ecology and Evolution of Microbial Competition," *Trends in Microbiology*, vol. 24, no. 10, pp. 833–845, 2016.

[291] A. Bashan, T. E. Gibson, J. Friedman, V. J. Carey, S. T. Weiss, E. L. Hohmann, and Y. Y. Liu, "Universality of human microbial dynamics," *Nature*, vol. 534, no. 7606, pp. 259–262, 2016.

[292] S. Shibasaki and S. Mitri, "A spatially structured mathematical model of the gut microbiome reveals factors that increase community stability," *iScience*, vol. 26, no. 9, p. 107499, 2023.

[293] A. Marantos, N. Mitarai, and K. Sneppen, "From kill the winner to eliminate the winner in open phage-bacteria systems," *PLoS Computational Biology*, vol. 18, no. 8, pp. 1–15, 2022.

[294] D. Bajić, J. C. Vila, Z. D. Blount, and A. Sánchez, "On the deformability of an empirical fitness landscape by microbial evolution," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, no. 44, pp. 11286–11291, 2018.

[295] B. Van Dijk, J. Meijer, T. D. Cuypers, and P. Hogeweg, "Trusting the hand that feeds: Microbes evolve to anticipate a serial transfer protocol as individuals or collectives," *BMC Evolutionary Biology*, vol. 19, no. 1, pp. 1–18, 2019.

[296] S. A. Becker, A. M. Feist, M. L. Mo, G. Hannum, B. Palsson, and M. J. Herrgard, "Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox," *Nature Protocols*, vol. 2, no. 3, pp. 727–738, 2007.

[297] A. R. Zomorrodi and C. D. Maranas, "OptCom: A multi-level optimization framework for the metabolic modeling and analysis of microbial communities," *PLoS Computational Biology*, vol. 8, no. 2, 2012.

[298] J. Chen, J. A. Gomez, K. Höffner, P. Phalak, P. I. Barton, and M. A. Henson, "Spatiotemporal modeling of microbial metabolism," *BMC Systems Biology*, vol. 10, no. 1, pp. 1–13, 2016.

[299] P. Phalak, J. Chen, R. P. Carlson, and M. A. Henson, "Metabolic modeling of a chronic wound biofilm consortium predicts spatial partitioning of bacterial species," *BMC Systems Biology*, vol. 10, no. 1, pp. 1–20, 2016.

[300] B. Teusink, A. Wiersma, D. Molenaar, C. Francke, W. M. De Vos, R. J. Siezen, and E. J. Smid, "Analysis of growth of Lactobacillus plantarum WCFS1 on a complex medium using a genome-scale metabolic model," *Journal of Biological Chemistry*, vol. 281, no. 52, pp. 40041–40048, 2006.

[301] D. R. Garza, M. C. Van Verk, M. A. Huynen, and B. E. Dutilh, "Towards predicting the environmental metabolome from metagenomics with a mechanistic model," *Nature Microbiology*, vol. 3, no. 4, pp. 456–460, 2018.

[302] C. Diener, S. M. Gibbons, and O. Resendis-Antonio, "MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota," *mSystems*, vol. 5, no. 1, 2020.

[303] A. Noronha, J. Modamio, Y. Jarosz, E. Guerard, N. Sompairac, G. Preciat, A. D. Daníelsdóttir, M. Krecke, D. Merten, H. S. Haraldsdóttir, A. Heinken, L. Heirendt, S. Magnúsdóttir, D. A. Ravcheev, S. Sahoo, P. Gawron, L. Friscioni, B. Garcia, M. Prendergast, A. Puente, M. Rodrigues, A. Roy, M. Rouquaya, L. Wiltgen, A. Žagare, E. John, M. Krueger, I. Kuperstein, A. Zinovyev, R. Schneider, R. M. Fleming, and I. Thiele, "The Virtual Metabolic Human database: Integrating human and gut microbiome metabolism with nutrition and disease," *Nucleic Acids Research*, vol. 47, no. D1, pp. D614–D624, 2019.

[304] D. Popp and F. Centler, "μBialSim: Constraint-Based Dynamic Simulation of Complex Microbiomes," *Frontiers in Bioengineering and Biotechnology*, vol. 8, no. June, 2020.

[305] E. Karimian and E. Motamedian, "ACBM: An Integrated Agent and Constraint Based Modeling Framework for Simulation of Microbial Communities," *Scientific Reports*, vol. 10, no. 1, pp. 1–10, 2020.

[306] B. Yu, I. Dukovski, D. Kong, J. Bobrow, A. Ostrinskaya, D. Segrè, and T. Thorsen, "Experiments and simulations on short chain fatty acid production in a colonic bacterial community," *bioRxiv*, 2018.

[307] A. R. Pacheco, M. Moel, and D. Segrè, "Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems.," *Nature communications*, vol. 10, no. 1, p. 103, 2019.

[308] J. J. Faith, J. L. Guruge, M. Charbonneau, S. Subramanian, H. Seedorf, A. L. Goodman, J. C. Clemente, R. Knight, A. C. Heath, R. L. Leibel, M. Rosenbaum, and J. I. Gordon, "The long-term stability of the human gut microbiota," *Science*, vol. 341, no. 6141, 2013.

[309] M. A. Henson and P. Phalak, "Suboptimal community growth mediated through metabolite crossfeeding promotes species diversity in the gut microbiota," *PLoS Computational Biology*, vol. 14, no. 10, pp. 1–21, 2018.

[310] Q. Garrigues, E. Apper, S. Chastant, and H. Mila, "Gut microbiota development in the growing dog: A dynamic process influenced by maternal, environmental and host factors," *Frontiers in Veterinary Science*, vol. 9, 2022.

[311] R. B. Guevarra, J. H. Lee, S. H. Lee, M. J. Seok, D. W. Kim, B. N. Kang, T. J. Johnson, R. E. Isaacson, and H. B. Kim, "Piglet gut microbial shifts early in life: Causes and effects," *Journal of Animal Science and Biotechnology*, vol. 10, no. 1, pp. 1–10, 2019.

[312] C. Zhang, J. Chen, Q. Wu, B. Xu, and Z. Huang, "The Gut Microbiota of Young Asian Elephants with Different Milk-Containing Diets," *Animals*, vol. 13, no. 5, 2023.

[313] J. Kambe, Y. Sasaki, R. Inoue, S. Tomonaga, T. Kinjo, G. Watanabe, W. Jin, and K. Nagaoka, "Analysis of infant microbiota composition and the relationship with breast milk components in the asian elephant (Elephas maximus) at the zoo," *Journal of Veterinary Medical Science*, vol. 82, no. 7, pp. 983–989, 2020.

[314] Y. Shigeno, H. Zhang, T. Banno, K. Usuda, T. Nochi, R. Inoue, G. Watanabe, W. Jin, Y. Benno, and K. Nagaoka, "Gut microbiota development in mice is affected by hydrogen peroxide produced from amino acid metabolism during lactation," *FASEB Journal*, vol. 33, no. 3, pp. 3343–3352, 2019.

[315] C. Kunz, S. Rudloff, W. Schad, and D. Braun, "Lactose-derived oligosaccharides in the milk of elephants: comparison with human milk," *British Journal of Nutrition*, vol. 82, pp. 391–399, nov 1999.

[316] M. Messer and K. R. Kerry, "Milk Carbohydrates of the Echidna and the Platypus," *Science*, vol. 180, pp. 201–203, apr 1973.

[317] M. A. O'Malley, "The nineteenth century roots of 'everything is everywhere'," *Nature Reviews Microbiology*, vol. 5, no. 8, pp. 647–651, 2007.

[318] J. Walter, A. M. Armet, B. B. Finlay, and F. Shanahan, "Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents," *Cell*, vol. 180, no. 2, pp. 221–232, 2020.

[319] N. D. Aluthge, W. A. Tom, A. C. Bartenslager, T. E. Burkey, P. S. Miller, K. D. Heath, C. Kreikemeier-Bower, H. Kittana, R. J. Schmaltz, A. E. Ramer-Tait, and S. C. Fernando, "Differential longitudinal establishment of human fecal bacterial communities in germ-free porcine and murine models," *Communications Biology*, vol. 3, no. 1, pp. 1–14, 2020.

[320] G. Trujillo-de Santiago, M. J. Lobo-Zegers, S. L. Montes-Fonseca, Y. S. Zhang, and M. M. Alvarez, "Gut-microbiota-on-a-chip: an enabling field for physiological research," *Microphysiological Systems*, vol. 1, no. 27, pp. 1–1, 2018.

[321] N. Ashammakhi, R. Nasiri, N. R. de Barros, P. Tebon, J. Thakor, M. Goudie, A. Shamloo, M. G. Martin, and A. Khademhosseni, "Gut-on-a-chip: Current progress and future opportunities," *Biomaterials*, vol. 255, no. June, 2020.

Summary

A complex community of microbes develops in the infant gut shortly after birth. We call this community the infant gut microbiota. The microbiota can influence the health of the infant, which makes the composition and function of the infant gut microbiota an important topic to study. It is not possible to directly study the development of the microbiota inside the infant, so we are limited to information from fecal samples and laboratory experiments. These can provide information on the composition of bacteria that leave the infant gut, and on the potential interactions between these bacteria, but we cannot assume that these observations are the same in the infant gut microbiota as it exists in the infant. Because it is so difficult to study, the processes and mechanisms that shape the microbiota also remain unclear. One technique that can provide insight into unseen processes such as these is mathematical modelling. Mathematical models can generate hypotheses and predictions about the inner workings of a system, such as the infant gut microbiota. In this thesis, we develop mathematical models to predict how the microbiota in the infant gut grows and interacts, and how it responds to oligosaccharides, a type of complex sugar.

The infant gut microbiota is composed of many species of microbes, including bacteria, archaea, and fungi. In this thesis we focus on the bacteria, which make up most of the microbiota. Previous studies have used the genomes of gut bacteria to predict what metabolic reactions they can do. These lists of metabolic reactions can be used to try to predict what metabolism a bacteria has under different circumstances, i.e. what metabolic reactions it uses. We have curated and refined these lists of reactions to better match the available data on how bacteria act in real life. The mathematical model we developed then predicts what metabolic reactions a bacterium uses in a particular environment with a technique called flux balance analysis. This lets the model make predictions both on how bacteria are influenced

by the environment and on how they influence the environment. In the model bacterial populations can influence their environment by taking up nutrients from it, by depositing other substances into it, and by growing. By applying this influence to the environment and repeating the technique, the model can make predictions for how the environment changes over time, and how the abundance and metabolism of bacteria changes over time. The model uses a grid to represent the gut environment, so that each bacterial population, nutrient, and metabolite has a particular location in the gut. This way it takes into account that different bacterial species may occur in different parts of the infant gut, and how the availability of nutrients may differ throughout the gut because of the influence of bacteria.

In this thesis, we use and further develop this model to make predictions on how various changes to the environment, such as the presence of oxygen, oligosaccharides, or antibiotic disturbances, influence the infant gut microbiota, their metabolism, and ultimately the infant.

In chapter 1 we first give an overview of the typical composition of the infant gut microbiota, and how this can change over time. We then describe our modelling approach in general, and flux balance analysis, the technique we use to model microbial metabolism, in particular. We explain how we use this technique to model not only a single bacterium but an entire microbiota. We also discuss how we model the potential variation over space and time in the infant gut microbiota.

In chapter 2 we introduce the mathematical model and its computational implementation, and use it to examine the effect of oxygen on the simulated infant gut bacteria. We show how the model can reproduce important parts of the metabolism of *Bifidobacterium* bacteria, and that it correctly predicts the major groups in the infant gut microbiota, but not the smaller groups. We also show that by introducing oxygen into the initial environment the model can reproduce the succession from *Escherichia coli* bacteria to *Bifidobacterium* bacteria that is observed in the real infant gut microbiota. The model can explain this through the different metabolisms of these species: while oxygen is abundant *E. coli* grows faster, but as oxygen depletes *Bifidobacterium* bacteria grow faster.

In chapter 3 we extend the model to also include different prebiotic oligosaccharides. Prebiotic oligosaccharides are complex sugars that cannot be digested by the infant, but can be digested by *Bifidobacterium* bacteria. These oligosaccharides are present in human milk, and also often added to infant formula. With the model we examine how oligosaccharides may affect not only the *Bifidobacterium* bacteria that

consume them, but also butyrate producing bacteria that consume the products of *Bifidobacterium* bacteria. These butyrate producing bacteria may play an important role in infant health. The model predicts that, in isolation, butyrate producing bacteria can co-exist with *Bifidobacterium* bacteria with or without oligosaccharides. However, the model predicts that in a complex ecosystem, like you might find in the infant gut, butyrate producing bacteria are outcompeted by other bacteria in most conditions. What bacterial species are effective competitors depends on whether oligosaccharides are present, and on what kind of oligosaccharide is present. The model predicts that the prebiotic milk oligosaccharide 2'-fucosyllactose can indirectly stimulate butyrate production within the infant gut microbiota, allowing them to outgrow their competitors. It can do so because *Bifidobacterium* produces propane-1,2-diol from 2'-fucosyllactose, which is consumed most efficiently by butyrate producers. These model predictions should be tested in laboratory experiments.

In chapter 4 we further extend the model with intestinal mucin. Mucin is produced by the gut wall, and consists largely of oligosaccharides that can be consumed by some bacteria. We also extend the model with extracellular digestion of oligosaccharides from both mucin and prebiotics. We aim to explain why the gut microbiota of breastfed infants consumes less mucin than the microbiota of non-breastfed infants, because the consumption of mucin may have negative health effects for the infant. The model predicts that the milk oligosaccharide 2'-fucosyllactose stimulates non-mucin consuming *Bifidobacterium* bacteria more than mucin-consuming *Bifidobacterium* bacteria. It further predicts that this happens because the mucin-consuming *Bifidobacterium* bacteria use extracellular digestion, and this causes the sugars they produce to be 'stolen' by other bacteria. This lets the non-mucin consuming *Bifidobacterium* bacteria outcompete the mucin consumers. We link these observations to the literature on 'public goods' that is available in the field of ecology.

In chapter 5 we examine how an antibiotic disturbance may affect the infant gut microbiota. We simulate an antibiotic disturbance in the model by increasing the probability for bacterial populations to die off on the simulated day 8 and 9 after birth, and then observe how the microbiota recovers. The model reproduces the reduced diversity of an antibiotic disturbance observed in the infant microbiota. The model further predicts that stronger disturbance of the microbiota by antibiotics that is observed in infants, compared to adults, may be partially explained by the antibiotics disturbing the natural succession that occurs in the early infant gut microbiota. Finally, the model predicts that prebiotic oligosaccharides in nutrition may lead to

a microbiota that returns to the original state more consistently after a disturbance, because the oligosaccharides consistently stimulate *Bifidobacterium* bacteria. However, the model does not reproduce the effect of a disturbance on the abundance of major bacterial groups correctly under most conditions. We give suggestions on how the model may be expanded to give a more complete view of the potential effects of antibiotic disturbances, and potentially function as a framework for future investigations.

Finally, in chapter 6 we place the model in a wider perspective. We first discuss limitations and future opportunities for the model, in particular how future versions may incorporate competition or cooperation between bacteria through antibiotics or acidification. We then compare our modelling approach with other approaches, and finally we discuss the direction of infant gut microbiota research in general. We predict that the microbiota of non-human mammals may provide further insights into the role of milk oligosaccharides in shaping the gut microbiota, and we also predict that a further standardisation of microbial modelling methods can lead to a better exchange of results between different studies.

Samenvatting

Een complexe gemeenschap van microben ontwikkelt zich in de babydarm kort na de geboorte. We noemen deze gemeenschap de babydarmmicrobiota. De microbiota kan de gezondheid van de baby beïnvloeden, en dat maakt de samenstelling en functie van de babydarmmicrobiota een belangrijk onderwerp om te bestuderen. Het is niet mogelijk om de ontwikkeling van de microbiota in de baby direct te bestuderen, dus de informatie die we hebben is beperkt tot poepmonsters en labexperimenten. Deze kunnen nuttige informatie geven over de samenstelling van bacteriën die de babydarm verlaten, en over de potentiële interacties tussen deze bacteriën, maar we kunnen er niet van uit gaan dat deze observaties hetzelfde zijn in de baby. Hierdoor is het onduidelijk welke processen en mechanismen belangrijk zijn om de microbiota te vormen. Een techniek die meer inzicht kan bieden in ongeziene processen zoals deze is wiskundig modelleren. Wiskundige modellen kunnen hypothesen en voorspellingen genereren over de mechanismen binnen een systeem, zoals de babydarmmicrobiota. In dit proefschrift ontwikkelen we wiskundige modellen die kunnen voorspellen hoe de microbiota in de babydarm groeit en interacteert, en hoe het reageert op oligosacchariden, een soort complex suiker.

De babydarmmicrobiota bestaat uit vele soorten microben, waaronder bacteriën, archaea, en schimmels. In dit proefschrift focussen we op de bacteriën, die het grootste deel van de microbiota vormen. Eerdere studies hebben het genoom van darmbacteriën gebruikt om te voorspellen welke metabole reacties ze kunnen gebruiken. Deze lijsten van metabole reacties kunnen gebruikt worden om te proberen te voorspellen wat voor metabolisme een bacterie heeft onder verschillende omstandigheden, i.e. wat voor metabole reacties het gebruikt. We hebben deze lijsten van reacties opgeschoond en verfijnd om beter overeen te komen met de beschikbare data over hoe bacteriën zich echt gedragen. Het wiskundige model dat we ontwikkeld hebben

voorspelt vervolgens welke metabole reacties een bacterie in een bepaalde omgeving gebruikt met een techniek die fluxbalansanalyse heet. Hierdoor kan het model voorspellingen maken zowel over hoe bacteriën beïnvloed worden door de omgeving, en hoe ze hun omgeving beïnvloeden. In het model kunnen bacteriële populaties de omgeving beïnvloeden door voedingsstoffen op te nemen, andere stoffen er in uit te scheiden en door te groeien. Door deze invloed toe te passen op de omgeving en de techniek te herhalen kan het model voorspellingen maken over hoe de omgeving over de tijd verandert, en hoe de grootte van de bacteriepopulaties en hun metabolisme verandert over de tijd. Het model gebruikt een raster om de omgeving te vertegenwoordigen, zodat elke bacteriepopulatie, voedingsstof, en restproduct een specifieke locatie heeft in de darm. Op deze manier kan het model er rekening mee houden dat verschillende bacteriesoorten in verschillende delen van de darm kunnen voorkomen, en hoe de beschikbaarheid van voedingsstoffen kan verschillen door de darm door de invloed van bacteriën.

In dit proefschrift gebruiken we dit model om voorspellingen te maken over hoe verschillende veranderingen in de omgeving, zoals de aanwezigheid van zuurstof, oligosachariden, of antibiotica, van invloed kunnen zijn op de babydarmmicrobiota, hun metabolisme, en uiteindelijk op de baby.

In hoofdstuk 1 geven we eerst een overzicht van de typische samenstelling van de babydarmmicrobiota, en hoe deze door de tijd heen kan veranderen. We beschrijven vervolgens onze modelaanpak en in het bijzonder fluxbalansanalyse, de methode die we gebruiken gebruikt om microbieel metabolisme te modelleren. We leggen uit hoe we deze techniek gebruiken om niet enkel individuele bacteriën te simuleren, maar ook de wisselwerking tussen bacteriën in een microbiota. We bespreken ook hoe we het de ruimtelijke variatie en de variatie over de tijd in de babydarmmicrobiota meenemen in het model.

In hoofdstuk 2 introduceren we ons wiskundige model en de computationele implementatie, en gebruiken we het om het mogelijke effect van zuurstof op en gesimuleerde babydarmmicrobiota te bestuderen. We laten zien hoe het model belangrijke delen van het metabolisme van *Bifidobacterium*-bacteriën kan reproduceren, en dat het correct de grote groepen in de babydarmmicrobiota kan voorspellen, maar niet de kleinere groepen. We laten ook zien dat, als we zuurstof in de omgeving aan het begin van de simulaties introduceren, het model de opvolging van *Escherichia coli*-bacteriën door *Bifidobacterium*-bacteriën kan reproduceren. Deze opvolging wordt ook waargenomen in de echte babydarmmicrobiota. Het model kan dit verklaren

door de verschillende metabolismes van deze soorten: als er veel zuurstof is groeien de *E. coli*-bacteriën sneller, maar als dit opraakt groeien de *Bifidobacterium*-bacteriën sneller.

In hoofdstuk 3 breiden we het model uit met verschillende prebiotische oligosachariden. Prebiotische oligosacchariden zijn complexe suikers die niet verteerd kunnen worden door de baby, maar wel door *Bifidobacterium*-bacteriën. Deze oligosachariden zijn aanwezig in menselijke melk, en worden ook vaak toegevoegd aan babyvoeding. We bestuderen hoe oligosachariden in het model niet enkel de *Bifidobacterium*-bacteriën die ze consumeren beïnvloeden, maar ook butyraatproducerende bacteriën die de restproducten van *Bifidobacterium*-bacteriën consumeren. Deze butyraatproducerende bacteriën spelen mogelijk een belangrijke rol in de gezondheid van de baby. Het model voorspelt dat, in isolatie, butyraatproducerende bacteriën kunnen samenleven met *Bifidobacterium*-bacteriën met of zonder oligosachariden. Het model voorspelt echter dat in een complex ecosysteem, zoals bestaat in de babydarm, butyraatproducerende bacteriën de competitie verliezen van andere bacteriën onder de meeste omstandigheden. Welke bacteriële soorten de competitie kunnen winnen van butyraatproducerende bacteriën hangt af van de aanwezigheid van oligosachariden, en van het type oligosachariden. Het model voorspelt dat het prebiotische melk-oligosacharide 2'-fucosyllactose indirect butyraatproductie kan stimuleren binnen de babydarmmicrobiota, waardoor de butyraatproducenten de competitie kunnen winnen. Dit kan gebeuren omdat *Bifidobacterium*-bacteriën propaan-1,2-diol produceren uit 2'-fucosyllactose, en omdat butyraatproducerende bacteriën efficiënter propaan-1,2-diol consumeren dan andere bacteriën. Deze modelvoorspellingen zouden getest moeten worden in labexperimenten.

In hoofdstuk 4 breiden we het model verder uit met darmmucine. Mucine wordt geproduceerd door de darmwand, en bestaat grotendeels uit oligosachariden die geconsumeerd kunnen worden door sommige bacteriën. We breiden het model ook uit met de extracellulaire vertering van oligosachariden uit zowel mucine als prebiotica. We proberen te verklaren waarom de darmmicrobiota van baby's die borstvoeding krijgen minder mucine consumeert dan de darmmicrobiota van baby's die geen borstvoeding krijgen, omdat de consumptie van mucine mogelijk negatieve gezondheidseffecten heeft voor de baby. Het model voorspelt dat het melk-oligosacharide 2'-fucosyllactose de *Bifidobacterium*-bacteriën die geen mucine consumeren meer stimuleert dan *Bifidobacterium*-bacteriën bacteriën die wel mucine consumeren. Daarnaast voorspelt het dat dit gebeurt doordat de mucine consumerende *Bifidobac-*

terium-bacteriën extracellulaire vertering gebruiken, waardoor de suikers die ze produceren 'gestolen' worden door andere bacteriën. We linken deze observaties aan de bestaande ecologische literatuur over collectieve goederen.

In hoofdstuk 5 kijken we naar de mogelijke effecten van een verstoring door antibiotica op de babydarmmicrobiota. We simuleren een verstoring met antibiotica door de kans dat bacteriële populaties uit het model verwijderd worden te verhogen op de gesimuleerde dag 8 en 9 na de geboorte, om vervolgens te observeren hoe de microbiota zich herstelt. Het model reproduceert de verminderde diversiteit die na een verstoring met antibiotica plaatsvindt. Daarnaast voorspelt het model dat de sterkere verstoring van de microbiota door antibiotica bij baby's, ten opzichte van volwassenen, deels verklaard kan worden doordat de antibiotica de natuurlijk opvolging in de vroege babydarmmicrobiota verstoort. Tot slot voorspelt het model dat prebiotische oligosacchariden in de voeding kunnen leiden tot een microbiota die sneller naar de oorspronkelijke staat terugkeert na een verstoring, doordat de oligosacchariden consistent de *Bifidobacterium*-bacteriën stimuleren. Het model voorspelt echter in de meeste gevallen niet correct de verschuiving in voorkomen van de belangrijke bacteriegroepen na een verstoring. We geven suggesties over hoe het model uitgebreid kan worden om een completer beeld te geven van de potentiële effecten van verstoringen door antibiotica, en mogelijk als raamwerk kan dienen voor toekomstig onderzoek.

Tot slot plaatsen we in hoofdstuk 6 het model in een breder perspectief. Eerst bespreken we de limitaties en toekomstige mogelijkheden voor het model, in het bijzonder hoe toekomstige versies competitie of samenwerking tussen bacteriën via antibiotica of verzuring kunnen modelleren. Vervolgens vergelijken we onze modelaanpak met andere modellen, en we bespreken de richting van het onderzoek naar de babydarmmicrobiota in het algemeen. We voorspellen dat de microbiota van niet-menselijke zoogdieren verder inzicht kan geven in de rol van melk-oligosacharide in het vormen van de darmmicrobiota. Daarnaast voorspellen we dat een verdere standaardisering van microbiële modelleringsmethoden kan leiden tot een betere uitwisseling van resultaten tussen studies.

Publications

D. M. Versluis, R. Schoemaker, E. Looijesteijn, D. Muysken, P. V. Jeurink, M. Paques, J. M. W. Geurts, and R. M. H. Merks, "A Multiscale Spatiotemporal Model Including a Switch from Aerobic to Anaerobic Metabolism Reproduces Succession in the Early Infant Gut Microbiota," *mSystems*, vol. 7, 2022.
D. M. Versluis, R. Schoemaker, E. Looijesteijn, J. M. W. Geurts, and R. M. H. Merks, "2'-Fucosyllactose helps butyrate producers outgrow competitors in infant gut microbiota simulations," *iScience*, vol. 27, 2024.

Curriculum Vitae

David Versluis was born on September 9, 1996 in Gorinchem, the Netherlands. He attended the basisschool Koningin Wilhelmina in Leerdam, and in 2011 he obtained his VWO Diploma with the profiles "Natuur en Techniek" en "Natuur en Gezondheid" at the Heerenlanden College in Leerdam. He followed the bachelor Biology at Utrecht University, with a focus on plant sciences, microbial interactions, and (paleo)ecology, graduating in 2014 with a thesis on the effects of geography on migrations and competition during the glacial periods.
He then switched topics and followed a pre-master in computer science and the master program Artificial Intelligence at Utrecht University. He graduated from this in 2018 with a thesis on solving the vehicle routing program using a model inspired by the slime mold *Physarum polycephalum*.
Combining his previous interests he started his PhD research on 'Multiscale modelling of the infant gut microbiota' in 2019 in the Multiscale Mathematical Biology group of prof. dr. Roeland Merks at Leiden University, funded by FrieslandCampina. In the course of this PhD he has presented posters and presentations at many national and international conferences, workshops, and seminars. He has also assisted the Mathematical multiscale biology course and given lectures on multiscale modelling in general and metabolic modelling in particular. He is now carrying out postdoctoral work funded by FrieslandCampina at Leiden University and at the university of Cincinnati, to further improve and validate the models presented in this thesis.

Acknowledgements

Here at the conclusion of my PhD I'd like to thank all the people that supported me throughout the years, and without whose advice and solace this project would never have been completed. Firstly I want to thank my promotor Roeland Merks, for believing the project could be completed, even when I didn't, and for his enthusiasm at getting me to talk more widely about my research. I also want to thank my second promotor Herman Spaink for his support, for the frequent well-wishes in the hallway, and for all the food and drink he generously prepared.

For their help with writing, coding, thinking, relaxing, and supporting me during everything else a PhD entails, many thanks to my officemates, past and present: Leonie, Sandro, Roy, Erika, Tessa and Joaquin. Also many thanks for the same to the others from multiscale mathematical biology: Daipeng, Koen, Martijn and Daphne. For her brief but very useful insights at a crucial time I want to thank Bastienne from the bioinformatics center. I also want to thank my students, Yang and Clair, for their own useful questions about our methods, and wish them all the best in their current and future careers.

For their help with analysis, proofreading, and planning our research, and for bringing us in contact with the people from Cincinnati, I want to thank the people from FrieslandCampina, former and current: Prescilla, Jan, Ellen, Andre, Ruud, and everyone behind the scenes. I want to thank our collaborators from Cincinnati, Alexander, David, and Ardythe, for their enthusiasm about our methods, their eagerness to test them, and their hospitality.

For the many cups of coffee and conversations, relevant or not, I want to thank Gabriel and Erin in particular, and all the other colleagues from Animal Sciences and Health, who carefully listened even when the topic of my presentations was far removed from their expertise: Guus, Alex, Saskia, Saloni, Patrick, Monica, Salome,

Wendy, Vincent, Wanbin, Jie, Jiajun, Gangyin, Arwin, Hildert, Mayra, Li, Bei, Bjørn, Natasha, Ulrike, Ewa, Marcel, Annemarie, and Coert. For their suggestions, both for careers and dinners, I want to thank the people (formerly) from biophysics who joined me for coffee or at the lunch table during the most difficult parts of my PhD: Andreea, Rik, Jeremy and Esmée.
I also want to thank my dear friends, particularly Kayley, Peter, Daan, Robin, Maarten, and Jonathan, and everyone else who listened to my stressed out stories about timeplans and infant feces. I want to thank my family for their unwavering support, and for bearing with me for all this time: my parents Wilke and Jan Willem and my sisters Eva and Laura, as well as oma Bep, who would have loved to see this thesis finished, and all the rest of the Versluis and Olbertijn families.
Finally, I wish to thank all the unseen and anonymous support staff that worked behind the scenes at the university during my time here.