



Universiteit
Leiden
The Netherlands

Simulations of the infant gut microbiota suggest that complex ecological interactions regulate effects of human milk oligosaccharides on microbial mucin consumption

Versluis, D.M.; Wijtkamp, C.; Looijesteijn, E.; Geurts, J.M.W.; Merks R.M.H.

Citation

Versluis, D. M., Wijtkamp, C., Looijesteijn, E., & Geurts, J. M. W. (2024). Simulations of the infant gut microbiota suggest that complex ecological interactions regulate effects of human milk oligosaccharides on microbial mucin consumption. *Biorxiv*.
doi:10.1101/2024.07.15.603541

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/4176185>

Note: To cite this publication please use the final published version (if applicable).

RESEARCH PAPER

Simulations of the infant gut microbiota suggest that complex ecological interactions regulate effects of human milk oligosaccharides on microbial mucin consumption

David M. Versluis^a, Clair Wijtkamp^b, Ellen Looijesteijn^c, Jan M. W. Geurts^c, Roeland M. H. Merks^{a,b}

^a Institute of Biology, Faculty of Science, Leiden University, The Netherlands

^b Mathematical Institute, Faculty of Science, Leiden University, The Netherlands

^c FrieslandCampina, Amersfoort, The Netherlands

ARTICLE HISTORY

Compiled July 15, 2024

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 of human milk oligosaccharides (HMOs) in human milk. We hypothesize that HMOs protect mucin by stimulating non-mucin consuming bacteria. To understand the underlying mechanisms we developed a computational model that describes the metabolism and ecology of the infant gut microbiota. Model simulations suggest that extracellular digestion of the HMO 2'-fucosyllactose by the mucin-consumer *Bifidobacterium bifidum* may make this species vulnerable to competitors. The digestion products of HMOs become 'public goods' that can be consumed by competing species such as *Bacteroides vulgatus* instead. *Bifidobacterium longum*, which does not consume mucin or produce public goods, can then become dominant, despite growing less efficiently on HMOs in monocultures than *B. bifidum*. In conclusion, our model simulations suggest that, through complex ecological interactions, HMOs may help lower mucin consumption by stimulating the non-mucin consumer *B. longum* at the expense of the mucin consumer *B. bifidum*.

KEYWORDS

Infant gut microbiota; Microbial Ecology; Public Goods; Human milk oligosaccharides

1. Introduction

The interior of the gastrointestinal tract is covered by a layer of mucin which protects the gut against diseases such as colitis [1] and rotavirus infection [2]. The mucin layer consists largely of mucin glycoproteins. Some species of intestinal bacteria, such as *Bacteroides* spp. and *Bifidobacterium bifidum*, can consume these mucins [3,4]. As mucin protects the gut against infection, consumption of intestinal mucins by the residing microbiota potentially increases the risk of disease [5]. It is thought that breastfeeding shapes the infant gut microbiota in such a way that the microbiota consumes less mucin [6]. 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 [6].

As 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 [7]. They are not digested by the host, but exclusively by intestinal

bacteria, thereby shaping the microbiota [8]. About 200 HMO structures exist, each of which consists of a core of lactose with other sugars attached to it [9]. We will focus in particular on the potential effects of the HMO 2'-fucosyllactose (2'-FL), as it is the most abundant HMO in most human milk [10] and its digestion by the infant gut microbiota is well-characterized [11]. We also examine galacto-oligosaccharides (GOS), which have a similar structure and are consumed by the same bacteria, but are not HMOs [12]. GOS are frequently added to commercial infant formula [13]. 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 [14], and a high abundance of *B. longum* in the infant gut microbiota lowers total mucin consumption in breastfed infants [15]. Paradoxically, in *in vitro* experiments, the mucin consumer *Bifidobacterium bifidum* digests and takes up HMOs such as 2'-FL more efficiently than *B. longum* [11,16], but *B. bifidum* is much less abundant than *B. longum* in the infant gut [17,18]. 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 [19], but the species crucially differ in how they break down and transport HMOs into the cell [11]. *B. longum* strains that consume HMO, particularly *B. longum ssp. infantis*, import most HMOs, including 2'-FL, into the cell using ATP-dependent active transport [8]. 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 [7]. The import of lactose produced from 2'-FL does not require active transport [20]. 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 [21]. 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 [22]. The role of public goods in inter-species interaction has been studied in detail in many systems, in particular in budding yeast (*Saccharomyces cerevisiae*) [23]. Yeast cells can secrete the enzyme invertase that converts sucrose into fructose and glucose. With this enzyme yeast can grow in glucose-limited medium rich in sucrose. However, the cell that secretes invertase can only take up around 1% of the glucose produced, as 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 glucose produced by invertase from other cells. In yeast communities invertase producers and invertase non-producers typically co-exist. Co-existence is possible because they are playing a 'snowdrift game', in which whatever strategy is less common (producing or not producing invertase) has an advantage over the more common strategy [23]. The extracellular digestion of oligosaccharides by *B. bifidum* has also been considered as a form of public goods production [21], 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 populations may then explain why *B. bifidum* is less abundant than other bacteria *in vivo*. As in yeast, an advantage of the less common strategy (producing) over the more common strategy (non-producing) may explain why *B. bifidum* is often present at low abundance in the infant gut.

Concretely, we investigated (1) whether 2'-FL in milk can explain the reduced mucin consumption 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 that uses flux balance analysis (FBA) on genome-scale metabolic models (GEMs). Such models have previously been used to model dynamic bacterial interactions with spatial separation, such as in the gut microbiota [24–27]. We based the model on our previous microbiota models [26,28,29], which modelled microbial interactions in the adult and infant gut microbiota, including relevant host factors such as flow rate and the presence of initial oxy-

gen. 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 [26,28,29]. 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* as observed *in vitro* and that it predicts that:

- (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.

2. Results

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*, *Anaerobutyricum hallii* cannot grow on mucin in single culture, whereas it can grow on mucin in co-culture with *B. bifidum* [30]. *B. bifidum* digests mucin extracellularly, which leads to cross-feeding metabolites becoming available in the medium as public goods [30]. *A. hallii* can then grow on these cross-feeding metabolites [30]. We first investigated whether our computer model could reproduce this cross-feeding.

We simulated this two-species system by extending a multiscale model as previously described [28,29]. 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 FBA on GEMs. The growth rates of the bacteria are given by a biomass reaction whose rate is maximized using FBA. The default biomass reaction in the model uses ATP production as a proxy for biomass production. 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. 1A&B). We distributed all nutrients and metabolites uniformly across all lattice sites at the start of each timestep. We 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 that attempt to capture the same biological processes with different representations. In the first model variant, the 'control model', we assumed that bacterial populations have access to all metabolites they produce, even when they are produced extracellularly. We implement this by letting all reactions be represented by FBA. In this model 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, the 'public goods model', we assumed that the products of extracellular metabolism become available to other bacterial populations when they are produced. We implement this with a separate handling of public goods production from mucin and HMO. In this 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* did not grow, even when *B. bifidum* produced cross-feeding products such as GalNAc (Fig. 1C). However, in the 'public goods' model variant, the model predicted that *B. bifidum* also released galactose into the medium. *A. hallii* could

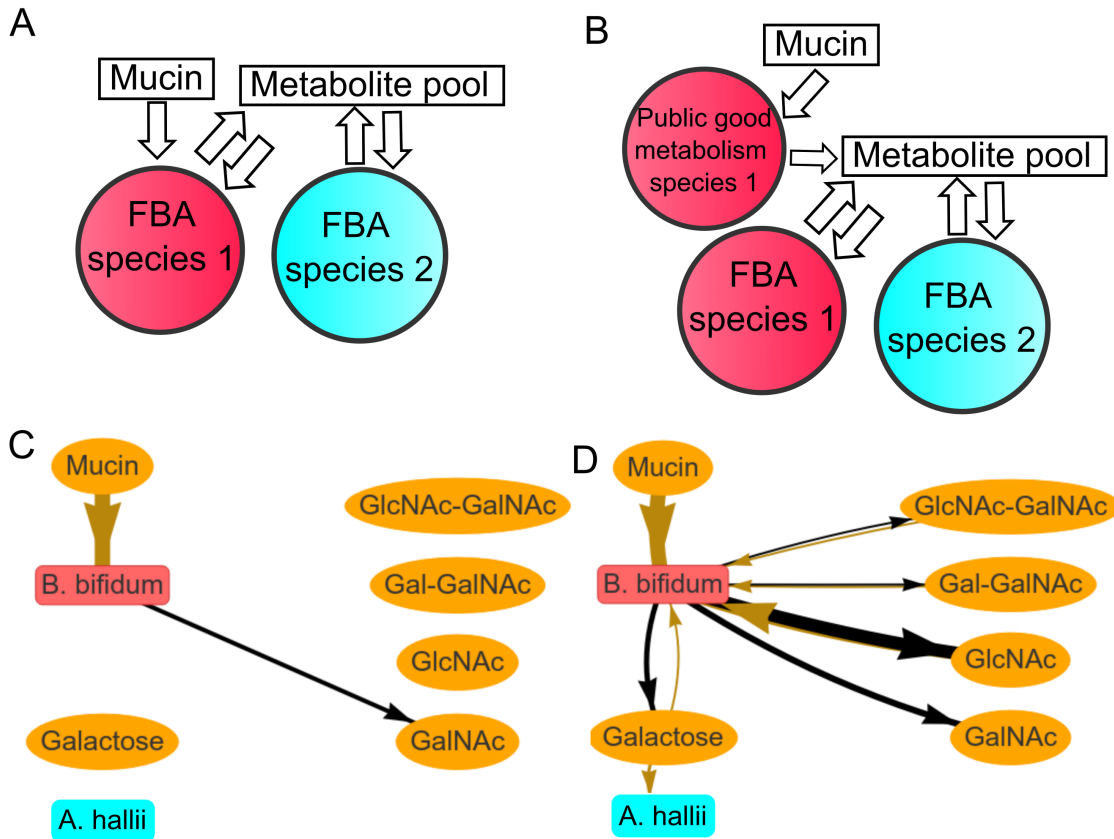


Figure 1. Model outline and effect of the public goods model

(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.

then grow in co-culture with *B. bifidum* by consuming the galactose (Fig. 1D), in agreement with experimental observations [30].

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.

2.2. Public goods model reproduces *in vivo* microbiota composition and mucin consumption

We next turned our attention to mucin consumption in the context of a more diverse infant gut 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 [6]. Secondly, *in vivo*, a breastfed infant microbiota dominated by *B. longum* consumes less mucin [15].

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 [28,29]. We extended this microbiota consortium with microbial reactions for mucin production and

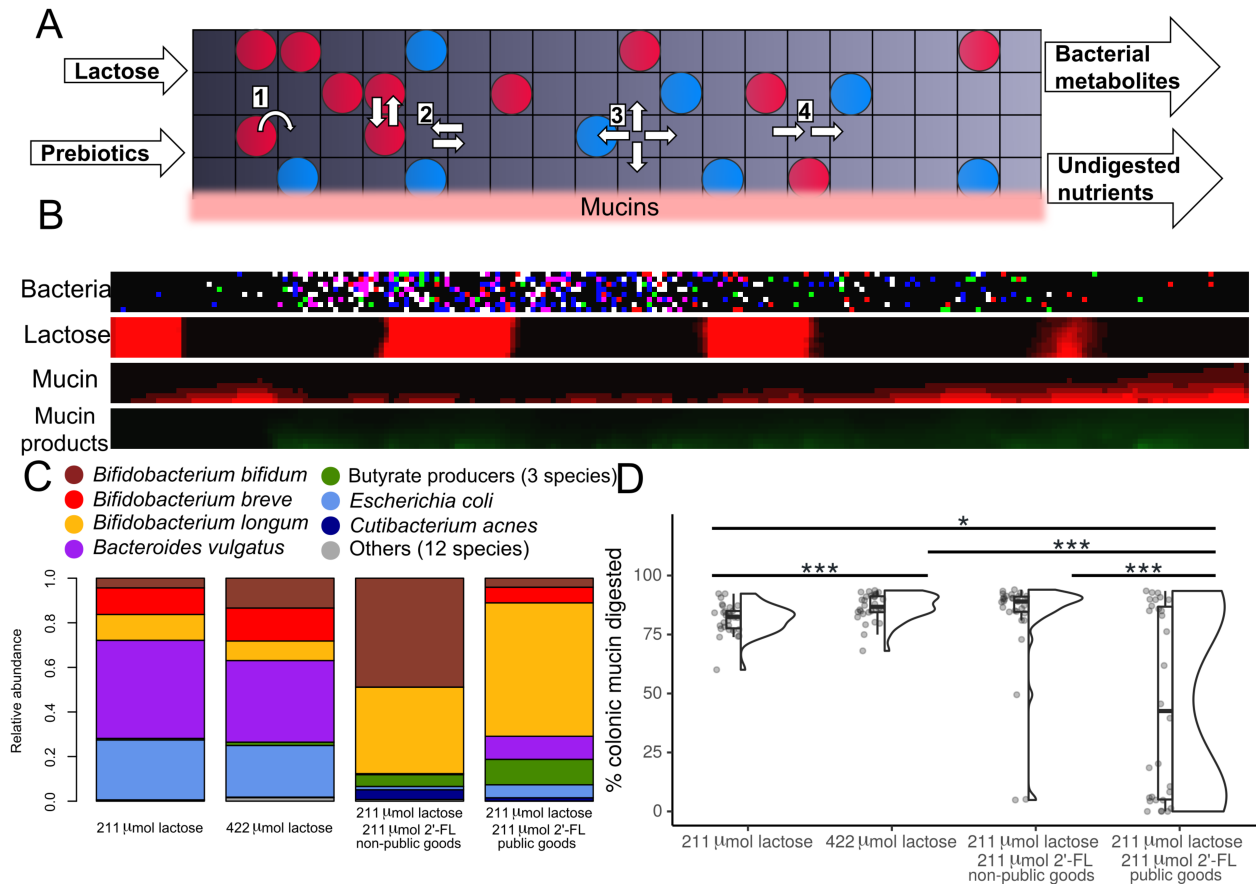


Figure 2. 2'-FL metabolism stimulates a microbiota that consumes less mucin and has dominant *B. longum* exclusively in the public goods model

(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 μ mol lactose, (2) 422 μ mol lactose, (3) 211 μ mol lactose and 211 μ mol 2'-FL without public goods 2'-FL metabolism, or 211 μ mol lactose and (4) 211 μ mol 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

consumption, and extracellular digestion of the HMO 2'-FL (see methods section 'Changes to genome-scale metabolic models'). Briefly, the spatial model represents the infant colon with a regular square lattice of 225×8 boxes of 2×2 mm, (Fig. 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. Mucin is produced at the upper boundaries of the system, to representing mucin production by the intestinal wall. The lower boundary represents the center of the intestinal lumen, the left boundary represents the entrance of the intestine, and the right boundary represents its exit.

Table 1. Species and subspecies included in the computational model. Color indicates color used in figures.

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

The model contains 21 bacterial species, each represented by a GEM (Table 1). The species were selected based on *in vivo* data [17], 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 variant alongside a non-public goods model variant as control. *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. *B. bifidum*, *Bacteroides vulgatus*, and *Parabacteroides distasonis* 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 (3 hours) (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 the conditions with 2'-FL (1 and 2), the simulations predicted a high abundance of all *Bifidobacterium* species, *Escherichia coli*, and *Bacteroides vulgatus*, (Fig. 2C) and a high mucin consumption (Fig. 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 [17,31,32] and high mucin consumption [6].

In the condition 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. 2C&D). In the condition with 2'-FL and public goods 2'-FL metabolism, we observed a dominance of *B. longum* and lower mucin consumption (Fig. 2C&D). Mucin consumption was not lower in all simulations, as the distribution was bimodal (Fig. 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 [17,18], and *in vitro* mucin consumption by the microbiota of breastfed infants is generally lower [6]. A high *B. longum* abundance is also associated with lower mucin consumption in breastfed infants [15].

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 hypothesise that this happens because 2'-FL stimulates *B. longum*, and that *B. longum* can suppress the mucin consumer *B. bifidum* through complex ecological interactions. We will further examine this in section '*B. bifidum* produces cross-feeding metabolites that may allow it to be exploited by other bacteria'.

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 analyzed 21 different parameter changes in total. We have collected these changes and the major outcomes in table 2. In the first 15 of the parameter changes we only changed the value of existing parameters. In the following 6 we also explored other relevant conditions, which we will now explain in some more detail. Firstly, we investigated the effects of replacing 2'-FL with GOS, a mixture of oligosaccharides that is commonly used in infant nutrition. To model GOS we implemented alternative digestion reactions in the GEMS. Notably, all *Bifidobacterium* species break down chains of four or more GOS sugars extracellularly [33–36]. 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 model is in methods section 'Changes to genome-scale metabolic models'. The second major change we investigated was a different mucin structure. In 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 [37,38]. To examine the consistency of our results for the non-fucosylated core-2 mucin we repeated the simulations, including the simulations with GOS, with a non-fucosylated core-2 mucin. A detailed description is in methods section 'Changes to genome-scale metabolic models'. Finally, to assess the effects of the simplifications we made in modelling biomass, we next tested four biomass proxies that different from our earlier assumption that ATP production rate determines biomass production: (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 used as building blocks for bacterial biomass in *E. coli in vitro* [39]. 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 we used [40].

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 the alternative biomass proxies. Proxies 2-4 led to unrealistic outcomes. Alternative biomass proxy 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. 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

Table 2. Parameter changes and outcomes

Parameter changed	Magnitude	Figure	Major differences in outcomes
Public goods production	×0.1	S1	-
	×0.5	S1	-
	×2	S1	-
Enzymatic constraint	×0.5	S2A	Butyrate producers extinct with 2'-FL
	×2	S2B	-
Growth & death probability	×0.1	S2C	Many others species are also abundant
	×10	S2D	-
Probability of random new populations	×0	S2E	-
	×10	S2F	-
Diffusion	×0.2	S2G	-
	×5	S2H	-
Initial oxygen	×0	S2I	Butyrate producers are also abundant without 2'-FL
	×10	S2J	-
Quantity of mucin released	×0	S2K	-
	×5	S2L	<i>Bifidobacterium</i> nearly extinct without 2'-FL
GOS instead of 2'-FL	N/A	S3	Higher abundance of <i>B. bifidum</i> and <i>B. breve</i> (p<0.001) Lower abundance of butyrate producers (p<0.001)
Mucin replaced with non-fucosylated mucin	N/A	S4	Similar outcomes
ATP + acetate as proxy for biomass	N/A	S5A&E	Similar outcomes, but <i>Lactobacillus gasseri</i> could never grow
Acetate as proxy for biomass	N/A	S5B&F	Microbiota consists nearly entirely of <i>B. longum</i>
ATP + pyruvate as proxy for biomass	N/A	S5C&G	<i>B. vulgatus</i> extinct
ATP + ketoglutarate as proxy for biomass	N/A	S5D&H	<i>B. vulgatus</i> extinct

of all GEMs [17,32]. We will discuss potential future improvements to the biomass reaction in the discussion section ‘Current and future extensions’.

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. 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 simulation without 2'-FL (Fig. 3A), of a simulation with public goods metabolism of 2'-FL and low mucin consumption (Fig. 3B) and of a simulation with public goods metabolism of 2'-FL and high mucin consumption (Fig. 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. 3A). The simulation with 2'-FL and low mucin consumption showed less diversity: it became dominated by 2'-FL consuming *B. longum* and some non-mucin consumers, which were predicted to cross-feed on lactate (Fig. 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. 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. 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*.

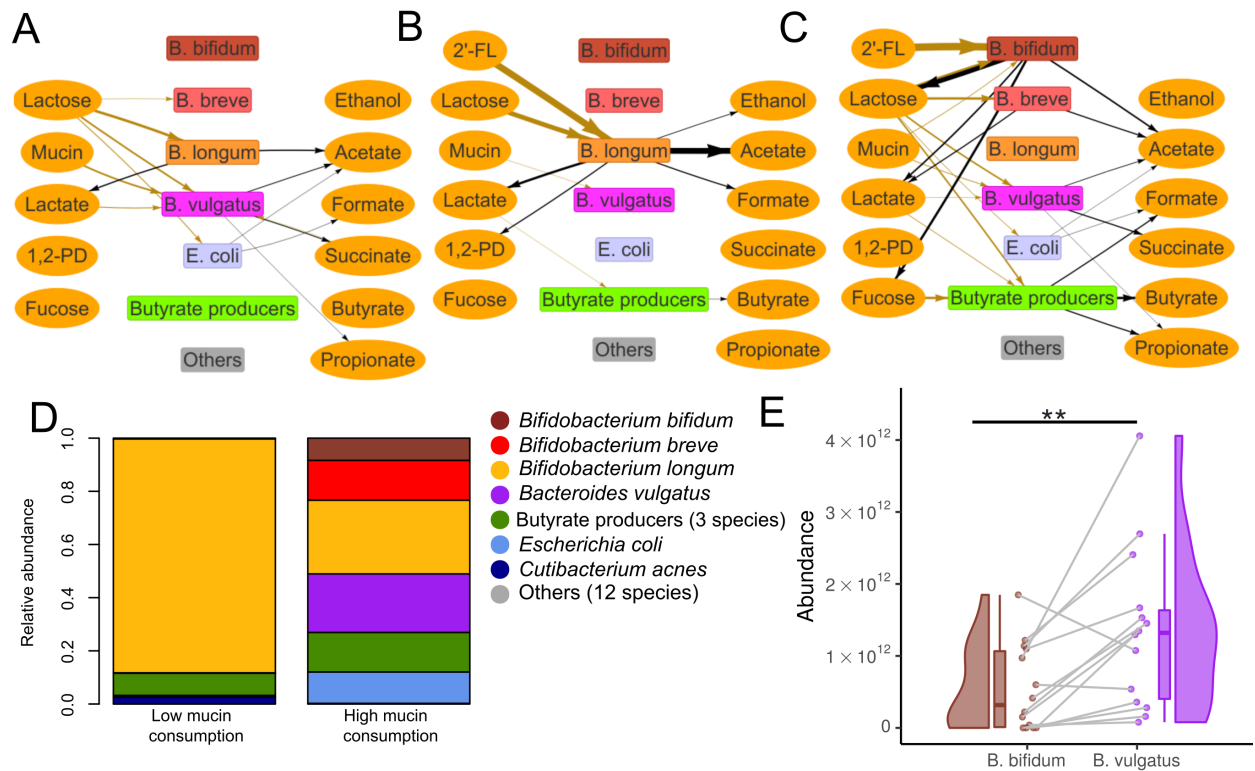


Figure 3. High consumption simulations show complex cross-feeding but higher abundance of *B. vulgatus* than *B. bifidum*

(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, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$

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 analyse 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. 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, within the simulations *B. bifidum* produced public goods that were consumed by other species, whereas *B. bifidum* was consistently less abundant than *B. vulgatus*, leading us to hypothesize that *B. vulgatus* exploited *B. bifidum*'s metabolism of 2'-FL.

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 goods model of 2'-FL metabolism. This would mean that *B. vulgatus* could not exploit *B. bifidum* without 2'-FL, nor with 2'-FL modelled without public goods metabolism.

To further investigate possible interactions between *B. vulgatus* and *B. bifidum* we studied a simplified model with only *B. vulgatus* and *B. bifidum* and no other species. We then performed new simulations using the same four conditions as those shown in Fig. 2. In simulations with only *B. bifidum* and *B. vulgatus*, both species became abundant in absence of 2'-FL (Fig. 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. 4A). However, with public goods 2'-FL metabolism *B. vulgatus* became much more abundant than *B. bifidum* (Fig. 4A), despite *B. vulgatus* not consuming 2'-FL itself. The network visualization showed that without public goods metabolism of 2'-FL, *B. bifidum* did not produce lactose as a public good (Fig. 4B). By contrast, with public goods metabolism of 2'-FL *B. vulgatus* fed on the lactose produced by *B. bifidum* from 2'-FL (Fig. 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 consumed mucin (Fig. 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 could perform public goods 2'-FL digestion we did not repeat the simulations with the public goods model of 2'-FL digestion. In absence of 2'-FL (conditions 1&2), *B. longum* and *B. vulgatus* were approximately equally abundant (Fig. 4E). In presence of 2'-FL *B. longum* became consistently more abundant than *B. vulgatus* (Fig. 4E). *B. longum* did not create lactose from 2'-FL like *B. bifidum* did (Fig. 4F). In contrast to *B. bifidum*, mucin consumption with *B. longum* and 2'-FL was lower compared to the simulations without 2'-FL ($p<0.01$, Fig. 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. 4F). We hypothesized that the consumption of lactate may have allowed *B. vulgatus* to survive in this limited consortium, but that this did not occur in the full consortium (Fig. 3D). We next added *Roseburia inulinivorans*, a lactate consumer from the full consortium, to see if it could outcompete *B. vulgatus* and so lower mucin consumption. *R. inulinivorans* indeed outcompeted *B. vulgatus* (Fig. 4E) and mucin consumption was indeed lower (Fig. 4G, $p<0.01$). To control for a possible effect of *R. inulinivorans* on mucin consumption without 2'-FL or without *B. longum* we performed several additional simulations. With *B. longum*, *B. vulgatus* and *R. inulinivorans* as species and without 2'-FL there was no reduction in mucin consumption compared to the same condition without

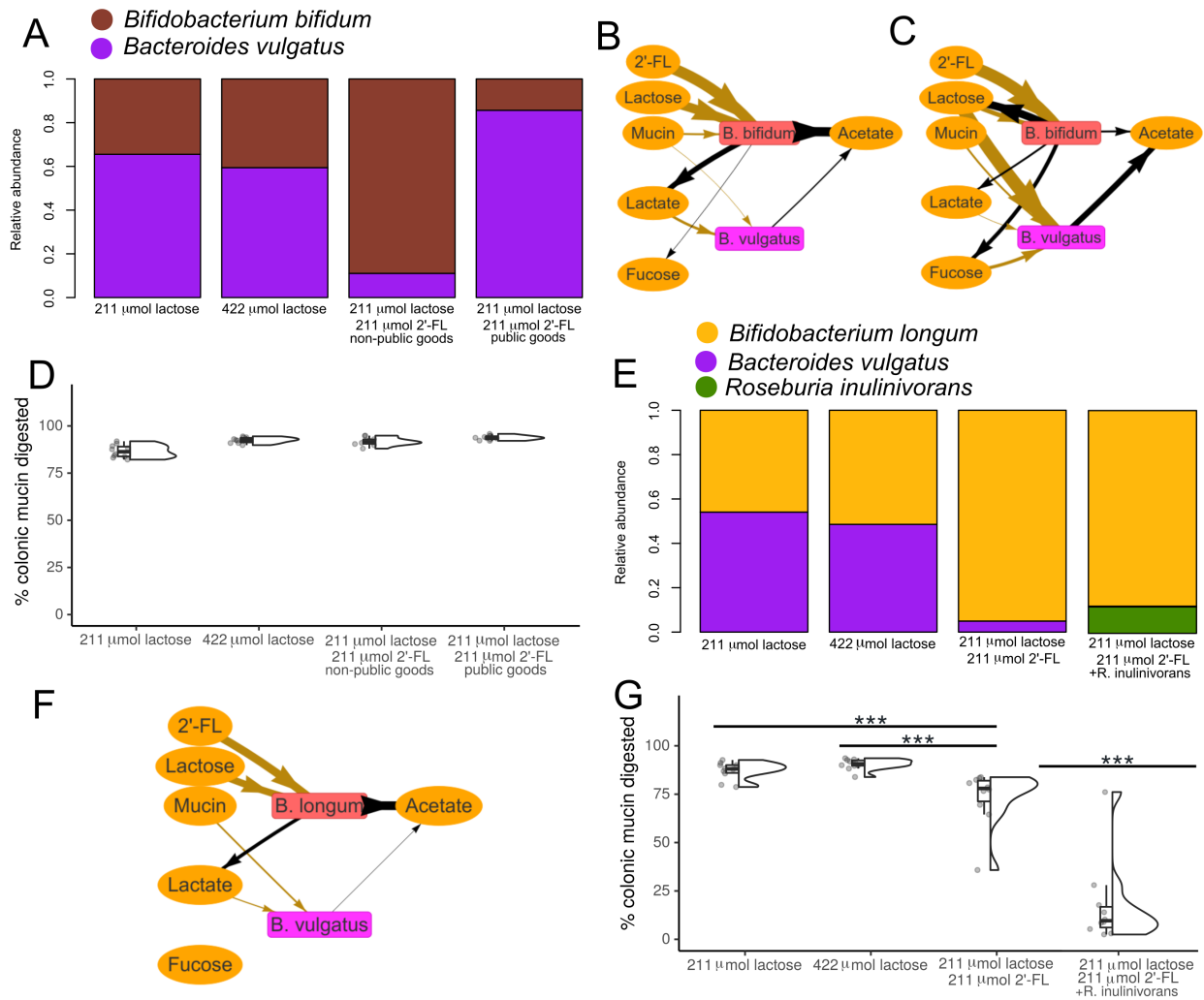


Figure 4. *B. bifidum*, but not *B. longum*, is exploited when digesting 2'-FL

(A) Average relative abundance of *B. bifidum* and *B. vulgatus*, with the following 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, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$

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 outcompete *B. vulgatus*. Because *B. longum* does not consume mucin, mucin consumption is lower when it outcompetes the mucin consumer *B. vulgatus*. *B. longum* metabolites can also feed a non-mucin consuming lactate consumer that can outcompete *B. vulgatus*, further lowering mucin consumption.

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 both the human milk oligosaccharide 2'-FL and galacto-oligosaccharides stimulate 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 by the microbiota [6]. The mechanism predicted by the model is as follows: oligosaccharides are consumed by *B. longum*, enabling it to outcompete mucin consuming species. *B. bifidum*, which consumes both mucin and oligosaccharides, will lose the competition against *B. longum* in a complex community, because of its extracellular metabolism of oligosaccharides. This extracellular metabolism produces intermediate digestion products which function as public goods, and allow competitors to consume them. *B. longum* digests oligosaccharides intracellularly and, therefore, does not produce public goods, and so is not sensitive to exploitation.

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. 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. 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* [30]. When the model included public goods metabolism of 2'-FL, the model predicted that *B. bifidum* produces lactose and fucose from 2'-FL which can be taken up by other species (Fig. 3C&4C). This is in line with *in vitro* observations [7,41]. Furthermore, the model predicted that 2'-FL will stimulate the growth of *B. longum* in the infant gut microbiota (Fig. 2C). This is a consequence of *B. longum* consuming 2'-FL intracellularly, thereby not producing public goods that could be used by competitors (Fig. 2C & 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 [18] or breastfeeding [31]. The model also predicted that the microbiota consumes less mucin in the presence of 2'-FL (Fig. 2D). In the model this is a result of the non-mucin consumer *B. longum* getting a competitive advantage over mucin consumers by 2'-FL (Fig. 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 [6]. 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. 3D). In agreement with this prediction, high abundance of *B. longum* is correlated with reduced mucin consumption *in vivo* in breastfed infants [15].

The simulations also predicted that in the presence of 2'-FL the public goods produced by *B. bifidum* are consumed by *B. vulgatus*, which prevents *B. bifidum* from becoming abundant. In agreement with this prediction, *B. bifidum* in the mouse gut has been shown to stimulate a, as of yet undetermined, species within the Bacteroidaceae, the family that contains *B. vulgatus* [42]. 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 [16,21]. These studies reported contradictory results: *B. breve* became less abundant than *B. bifidum* in a short-term *in vitro* experiment [16], 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 [21]. 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 [21]. However, in a culture of *B. bifidum* alone, a reduced flow rate only had a minor effect on the bacterial abundance [21]. 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 [21]. 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 [43]. 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.

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 have been previously studied *in vitro* in budding yeast (*S. cerevisiae*) [23]. Some yeast cells break down sucrose into glucose and fructose extracellularly with the enzyme invertase, but lose 99% of the glucose produced to diffusion, which then functions as a public good [23]. This leads to a dynamic community of invertase producers and non-producers [23,44]. 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 [23]. In the yeast system, the ratio of producers to non-producers depends on the cost associated with production [23]. High cost or low efficiency of production leads to a smaller fraction of producers [23]. 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 [23]. 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 [23]. In yeast, the cells expend energy to create invertase, which is modelled by directly decreasing their growth [23]. The yeast system predicts that producers are more abundant than non-producers when public goods production has low or no cost [23]. 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 [11]. Furthermore, *B. bifidum* is not able to consume fucose like *B. vulgatus* can [45]. An additional explicit metabolic cost for *B. bifidum*, as the producer has in the yeast model [23], may further reduce its predicted 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. A similar ecological role has been coined the 'loner' type in *in vitro* cultures of *Pseudomonas aeruginosa* strains [46], where 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 loners can obtain more substrate. The producers outcompete the loners because they can obtain even more substrate, as extracellular metabolism is more efficient [46]. 'Loners' have also been created through genetic modification in an *in vitro* budding yeast community [47]. As expected, yeast loners outcompete non-producers, but in contrast to *P. aeruginosa*, 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. 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. 3B&D). Thus, neither the yeast nor the *P. aeruginosa* 'loner' types 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. S1). 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.

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 [17,32]. The model simulations with either GOS or 2'-FL also predicted a high abundance of *Bifidobacterium*, which matches *in vivo* data [18,31,48]. However, various discrepancies exist between the model predictions and *in vivo* data. Firstly, the model predicts a gut microbiota composition typically consisting of only a few species (Fig. 2C), whereas *in vivo* measurements have shown that the gut microbiota consists of dozens of species [32]. Most of these species have a low abundance *in vivo*, but they may still influence the dynamics of the system as a whole [17,49]. 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. 2C), in contrast with *in vivo* data where such dominance is sometimes seen (e.g. in 18% of infants [50]). Although it is unclear why this is seen *in vivo*, there is a positive correlation with a lower gestational age and a defective mucin barrier [51,52]. The lower gestational age in particular may lead to a different initial composition, such as a composition without *Bifidobacterium*. 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 [38]. It is unclear why different mucin fucosylation leads to different *Bifidobacterium* abundances [38]. Thus, it is also unclear what additions could be made to the model to improve these predictions.

3.4. Current and future extensions

A potential problem with the dynamic FBA modelling approaches that we use is their difficulty in reproducing *in vitro* results on relative growth rates of different combinations of bacteria [53,54]. Inaccurate predictions have been attributed to a combination of factors, primarily insufficient curation of reactions in GEMs, incorrect biomass reactions, and a lack of condition-specific constraints [53–56]. We have addressed these issues as follows. To address the insufficient curation of many GEMs, we created an additional curation pipeline that included checks for mass-balance and the absence of ATP production without substrate. We also performed extensive comparison with previous *in vitro* observation and corrected the presence or absence of important carbon uptake reactions (Methods section ‘Changes to genome-scale metabolic models’). As part of the model we imposed an enzymatic constraint on the FBA solutions (Methods section ‘Flux balance analysis approach’), which improves metabolic predictions by correctly predicting overflow metabolism in many situations where classic FBA would not [57]. We paid particular attention to the curation of the crucial *Bifidobacterium* GEMs, and ensured that the simulations correctly reproduced the substrate concentration dependent bifid shunt pathway [19,28]. To address the uncertain and arbitrary nature of biomass reactions included in the GEMs, our model uses a biomass reaction that only requires ATP, which simplifies the system and allows us to focus on curating the carbon metabolism of the bacteria. However, it does limit the scope of the predictions in our model, as we cannot incorporate interactions through, e.g. differences in amino acid availability, which may be important [58,59]. Finally, we addressed condition-specific constraints. We selected the simulated nutrient inputs in the model to be relevant to the infant gut situation, and constructed the ‘public goods’ system to better simulate extracellular metabolism. We have shown that the ‘public goods’ system allows the model to reproduce *in vitro* results that it otherwise would not capture (Fig. 1). However, as we discussed previously, there are a number of discrepancies between our model predictions and existing *in vitro* and *in vivo* data. We want to emphasize that it remains uncertain whether our model predictions are correct, and that *in vitro* and *in vivo* validation are essential.

The discrepancies between the model outcomes and existing *in vitro* and *in vivo* data are likely the result of the model’s incomplete and simplified representation of many aspects of the infant gut and the infant gut microbiota. We will now discuss several aspects in which the model is incomplete, and may be expanded in the future. 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 [17]. We also do not include any viruses, fungi or archaea, which may play a significant role in metabolic interactions in the infant gut [60]. Future versions of the model may include a larger consortium of species. On a metabolic level, FBA requires many assumptions regarding bacterial metabolism, growth rate and biomass production [61,62]. To get insight into how the model predictions depended on our metabolic assumptions we performed further parameters variation after generating our main results. We examined the effect of different biomass reactions on the model predictions (Fig. S5) 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 selected in the different FBA solutions 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, many of these these cellular components 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 [63] and the type and concentration of

protein in infant formula influences microbiota composition [58,59]. 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 composition 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 and higher complexity of mucins compared to the mucin structures used in our model [4,64]. 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 [65]. We also modelled GOS, as it is a common component of infant formulas [66]. 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 [11]. 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 [67,68]. These factors could also influence interactions and competition between bacteria in the infant gut, and thus influence the predictions of the model.

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 [69], as both disorders have been associated with damage to the gastrointestinal mucin layer [70]. Strategies to prevent mucin loss, like nutritional interventions with 2'-FL or GOS as described in this paper, may be important in this light. Of course, other effects of breastfeeding, such as immune system modulation [71,72], may also play a role.

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 [73], 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. Methods

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 [26] and of the infant microbiota in particular [28,29]. 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. 2A). Each lattice site can contain a single population of a single species (Section 4.2, 'Species composition'). Each species is represented by a genome-scale metabolic model, to which we have made some improvements to (Section 4.3, 'Changes to genome-scale metabolic models'). Each lattice site can contain any number of nutrients and metabolites

in any concentration. Metabolism determines growth (Fig. 2A-1) and is calculated for each population using flux balance analysis (FBA, section 4.4, ‘Flux balance analysis approach’). We have developed both a non-spatial (section 4.5, ‘Non-spatial model’) and a spatial version of the model (section 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. 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. 2A-4). Nutrients are input into the system at regular intervals to allow for metabolism (Section 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 4.8, ‘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 4.9, ‘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.

Table 3. 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

4.2. Species composition

We selected the list of species to represent in the model (Table 1) from [17], 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 [74]. We then added an additional *Bifidobacterium breve* and *Bifidobacterium bifidum* GEM to represent the diversity of *Bifidobacterium* species in the infant gut [75]. We also added a GEM of *Roseburia inulinivorans*, as in our previous model [29]. *Roseburia* spp. have been shown to be prevalent butyrate producing bacteria in infants in other studies [76]. The populations of *Roseburia inulinivorans*, *Eubacterium hallii*, and *Clostridium butyricum* are pooled and listed as ‘Butyrate producers’ in the visualisations.

4.3. Changes to genome-scale metabolic models

We applied various changes and additions to the GEMs, as in our previous version of the model [29]. A full list of changed and added reactions is in table S1. We will highlight the more notable changes here. We disabled anaerobic L-lactate uptake for the *Bifidobacterium*

and *E. coli* GEMs [77,78] and added a lactose symporter to *Anaerobutyricum hallii* [30], all *Bifidobacterium* GEMs [79], *Roseburia inulinivorans* [80], *Haemophilus parainfluenzae* [81], and *Rothia mucilaginosa* [82]. We also added galactose metabolism to *R. inulinivorans* [83] and *R. mucilaginosa* [82]. 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 [84] and digests it to lactose and fucose inside the cell. *B. bifidum* breaks down 2'-FL to lactose and fucose using an extracellular fucosidase [7]. *B. longum* and *B. breve* import DP3 fractions of GOS intracellularly, and digest them to monosaccharides using beta-galactosidases [85]. *B. longum* and *B. breve* break down fractions longer than DP3 extracellularly using glycoside hydrolases [35,36]. *B. bifidum* digests all fractions extracellularly to lactose and galactose using extracellular beta-galactosidases [34]. We also added a fucosidase reaction to *B. bifidum* to allow it to remove fucose groups from mucin in line with the available literature [86].

4.4. Flux balance analysis approach

We use a modified version of FBA with an enzymatic constraint [61,87], as in previous model versions [28]. 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,

$$S \cdot \vec{f} = 0, \quad (1)$$

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 [61] 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,

$$\vec{F}_{in} \leq \vec{F}_{ub}, \quad (2)$$

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} :

$$\vec{F}_{ub}(\vec{x}, t) = \frac{\vec{c}(\vec{x}, t)}{B(\vec{x}, t)}, \quad (3)$$

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 [28]. \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 [28,87]. The enzymatic constraint represents the maximum, total amount of flux that can be performed per cell in each population:

$$\sum \vec{f} \leq a. \quad (4)$$

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 $F_{in}^{\vec{x}}(\vec{x}, t)$ and $F_{out}^{\vec{x}}(\vec{x}, t)$, 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.

4.5. Non-Spatial model

For the simulations of Fig. 1 we used a non-spatial version of the model to represent well-mixed *in vitro* conditions. All nutrients and metabolites are distributed uniformly 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,

$$\frac{d\vec{c}(\vec{x}, t)}{dt} = F_{out}^{\vec{x}}(\vec{x}, t)B(\vec{x}, t) - F_{in}^{\vec{x}}(\vec{x}, t)B(\vec{x}, t), \quad (5)$$

where $F_{out}^{\vec{x}}(\vec{x}, t)$ 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.

4.6. Spatial model

For all simulations after those of Fig. 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:

$$\frac{d\vec{c}(\vec{x}, t)}{dt} = F_{out}^{\vec{x}}(\vec{x}, t)B(\vec{x}, t) - F_{in}^{\vec{x}}(\vec{x}, t)B(\vec{x}, t) + \frac{D}{L^2} \sum_{\vec{i} \in NB(\vec{x})} (\vec{c}(\vec{i}, t) - \vec{c}(\vec{x}, t)), \quad (6)$$

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 $6.3 \cdot 10^5 \text{ cm}^2/\text{s}$. 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 [88,89]. 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.

4.7. 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 [90]. Food intake contains 211 μmol of lactose (‘lcts’ in the VMH database) by default, a concentration in line with human milk [91], assuming 98% host uptake of carbohydrates before reaching the colon [25]. 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 are inserted as separate fractions of DP3, DP4, or DP5 based on analysis of the composition of Vivinal-GOS [33]. 64% is DP3, 28% is DP4 and 8% is DP5. Water (‘h2o’ in the VMH database) is provided in unlimited quantities.

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:

$$\vec{F}_{in}(\vec{x}, t) \leq \frac{\vec{c}(\vec{x}, t)}{B(\vec{x}, t)}, \quad (7)$$

The change in concentration of metabolites is still determined by eq. 5 in the non-spatial model and eq. 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. 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 Flux balance analysis approach does.

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. 6) and the growth rate $g(\vec{x}, t)$ to the local population size, as follows:

$$\frac{dB(\vec{x}, t)}{dt} = B(\vec{x}, t)g(\vec{x}, t). \quad (8)$$

After initialisation, new bacterial populations can be created through reproduction (1) or introduction of new populations (2). To mimic reproduction (1), each population of at least $1 \cdot 10^{10}$ bacteria (Table 3) 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 populations (2) for each empty lattice site, with a probability of 0.00005, a species is selected at random, with equal probability for each species, and a population of this species is introduced into the lattice site. We initialize these populations at the same population size (B) as the initial populations in the model (Table 3). 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 [92], also used previously [26,28]. 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 $6.3 \cdot 10^5 \text{cm}^2/\text{s}$, 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.

4.10. 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* [93], *Cutibacterium acnes* [94], *Eggerthella* sp. YY7918 [95], and *Gemella morbillorum* [96]. 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 [97]. 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 [98–100]. 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,

$$l = \sum_i F(i) \cdot E(i). \quad (9)$$

In the simulations of Fig 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.

4.11. Parameters

Parameters of the system are listed in table 3. We estimate that the infant colon has a volume of 90ml [101,102]. 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} [103]. Values for free parameters were estimated and evaluated in the sensitivity analysis(Fig. S1&Fig. S2).

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 [104]. Diffusion of metabolites was implemented using the Forward Euler method. The model is based on our own earlier models of the gut microbiota [26,28,29]. 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) [97]. 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 [105].

5. Supplemental material

S1 Table

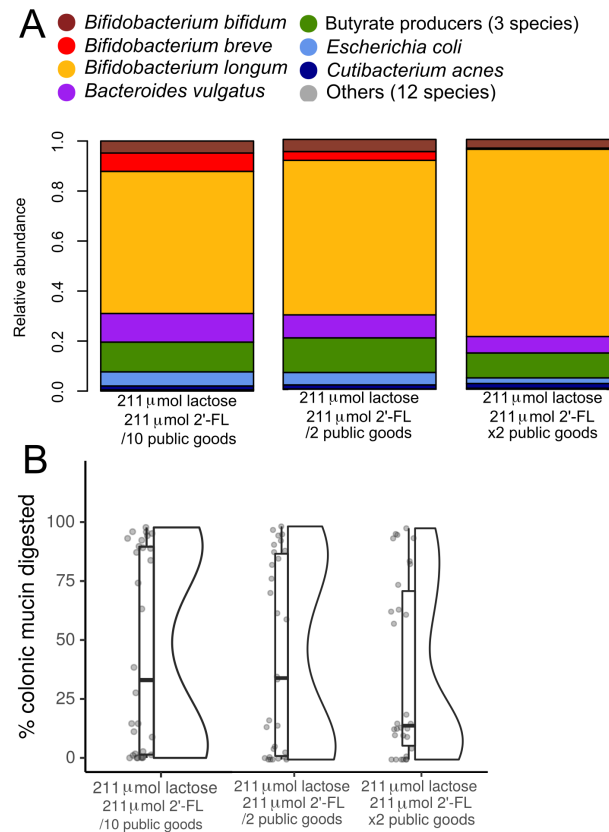
S1table.csv

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

S2 Table

S2table.csv

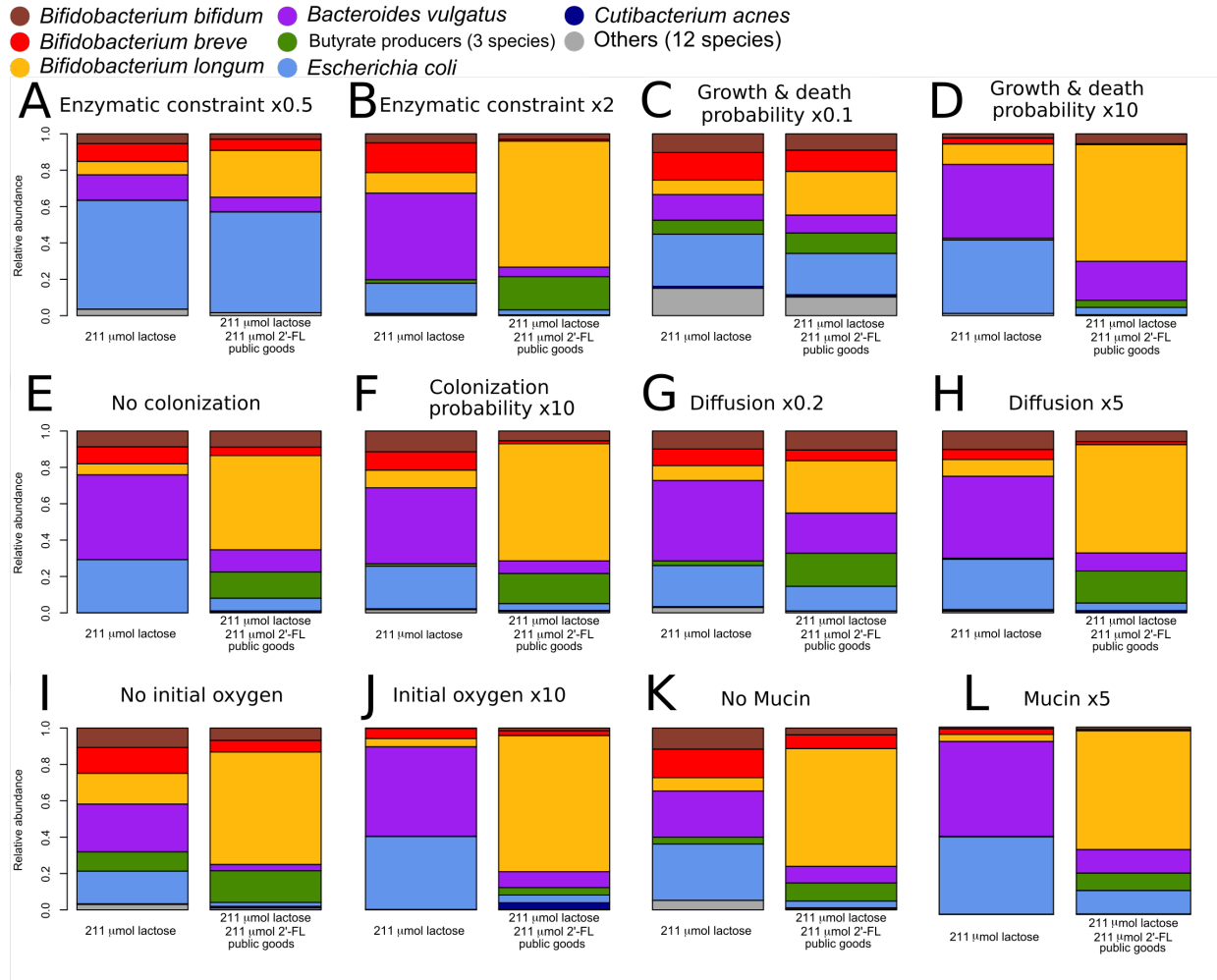
A table of mucin reactions used for public goods metabolism as a .csv file.



S1 Figure

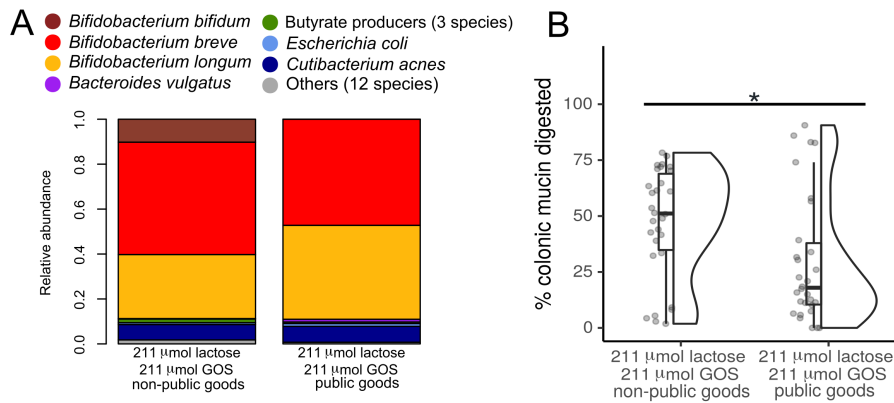
(A) Average relative abundance of bacterial species in the condition with 2'-FL with either the public good production rate reduced to a tenth, halved, or increased to double. $n=30$ for each condition.

(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

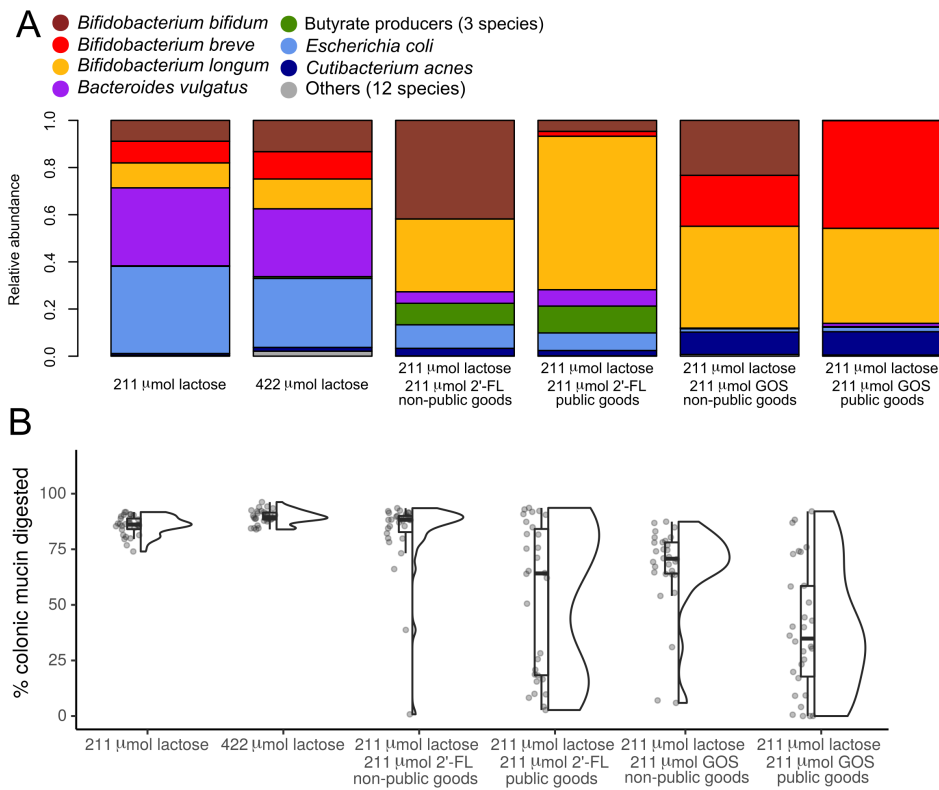
(A to J) Average relative abundance of bacterial species in the conditions with only lactose, or with 2'-FL and public goods 2'-FL metabolism, at the end of 21 days, with the following alteration from the baseline of Fig. 2A: (A) Enzymatic constraint loosened by a factor of 2, to 4 μmol flux per timestep 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 $1.26 \cdot 10^{-6} \text{ cm}^2/\text{s}$ (H) Diffusion of both metabolites and bacteria increased by a factor of 5 to $3.15 \cdot 10^{-5} \text{ cm}^2/\text{s}$ (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.

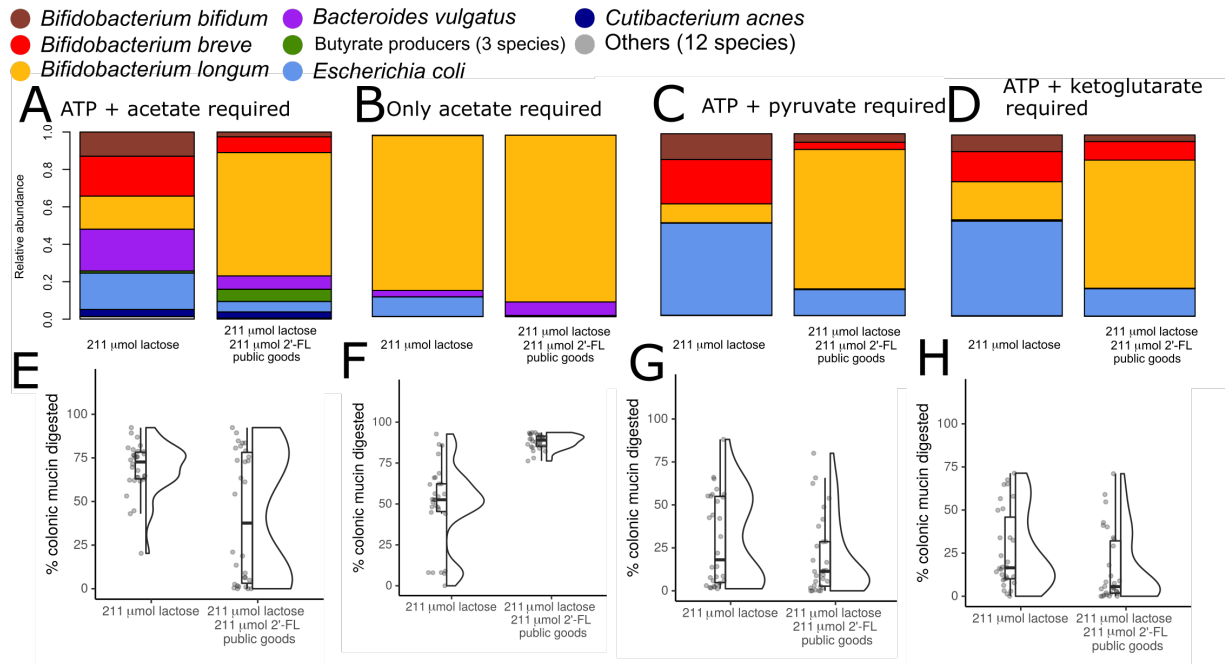
(B) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 60 timesteps of the model, per condition. n=30 for each condition



S4 Figure

(A) Average relative abundance of bacterial species with non-fucosylated mucin instead of fucosylated mucin in each condition from Fig. 2 and Fig. S4 at the end of 21 days. $n=30$ for each condition.

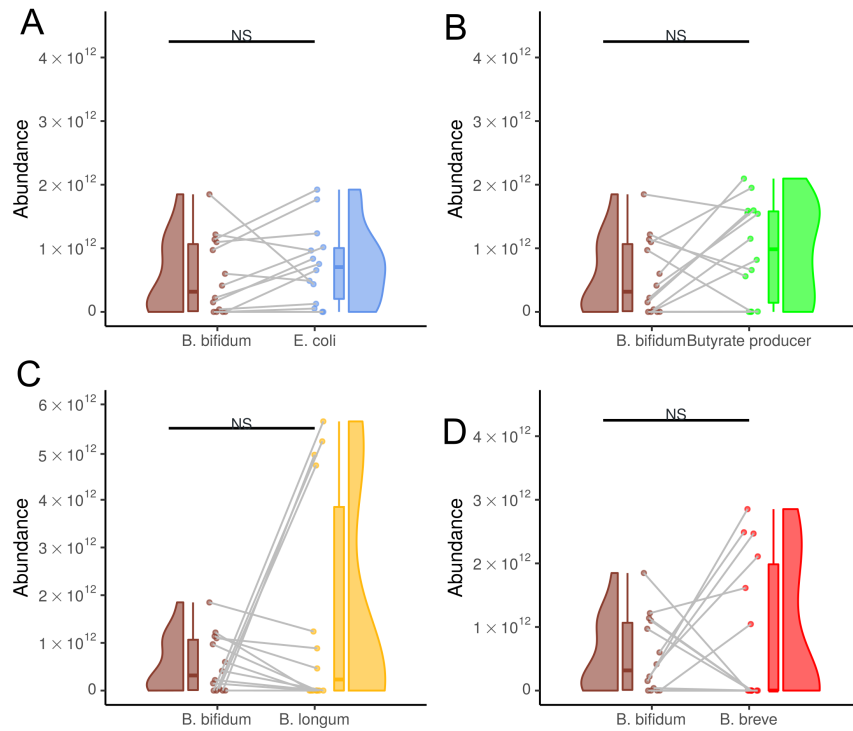
(B) Amount of colonic mucin digested by the microbiota as a percentage of total mucin released into the gut over the final 60 timesteps of the model, per condition. $n=30$ for each condition



S5 Figure

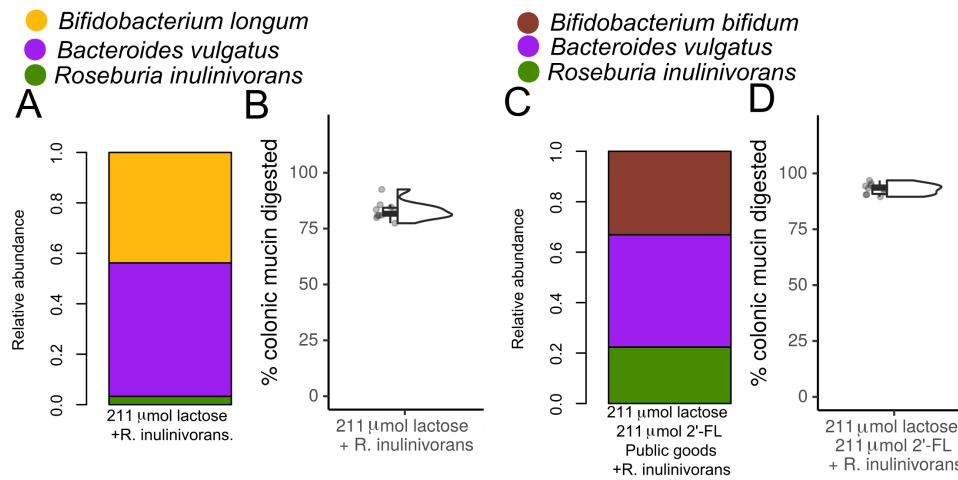
(A-D) Average relative abundance of bacterial species at the end of 21 days with the following alternative biomass reactions: (A) ATP + acetate (B) acetate (C) ATP + pyruvate (D) ATP + ketoglutarate n=30 for each condition.

(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

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 analysed 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.

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.

References

- [1] Fu J, Wei B, Wen T, et al. Loss of intestinal core 1-derived O-glycans causes spontaneous colitis in mice. *The Journal of clinical investigation*. 2011 apr;121(4):1657–66. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21383503>.
- [2] Yolken RH, Ojeh C, Khatri IA, et al. Intestinal Mucins Inhibit Rotavirus Replication in an Oligosaccharide-Dependent Manner. *Journal of Infectious Diseases*. 1994 may;169(5):1002–1006. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/169.5.1002>.
- [3] Bunesova V, Lacroix C, Schwab C. Fucosyllactose and L-fucose utilization of infant *Bifidobacterium longum* and *Bifidobacterium kashiwanohense*. *BMC Microbiology*. 2016;16(1):1–12. Available from: <http://dx.doi.org/10.1186/s12866-016-0867-4>.
- [4] Tailford LE, Crost EH, Kavanaugh D, et al. Mucin glycan foraging in the human gut microbiome. *Frontiers in Genetics*. 2015 mar;6(FEB). Available from: <http://www.frontiersin.org/Nutrigenomics/10.3389/fgene.2015.00081/abstract>.
- [5] Desai MS, Seekatz AM, Koropatkin NM, et al. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. *Cell*. 2016;167(5):1339–1353.e21.
- [6] Midtvedt AC, Carlstedt-Duke B, Midtvedt T. Establishment of a Mucin-Degrading Intestinal Microflora During the First Two Years of Human Life. *Journal of Pediatric Gastroenterology and Nutrition*. 1994 apr;18(3):321–326. Available from: <http://journals.lww.com/00005176-199404000-00012>.
- [7] Sakanaka M, Gotoh A, Yoshida K, et al. 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*. 2020;12(1):1–21.
- [8] Sela DA, Adeuya A, Kim JH, et al. The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proceedings of the National Academy of Sciences*. 2008;105(48):18964–18969. Available from: www.pnas.org/cgi/content/full/.
- [9] Bode L. Human milk oligosaccharides: Every baby needs a sugar mama. *Glycobiology*. 2012; 22(9):1147–1162.
- [10] McGuire MK, Meehan CL, McGuire MA, et al. What’s normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically. *American Journal of Clinical Nutrition*. 2017;105(5):1086–1100.
- [11] Garrido D, Ruiz-Moyano S, Lemay DG, et al. Comparative transcriptomics reveals key differences in the response to milk oligosaccharides of infant gut-associated bifidobacteria. *Scientific Reports*. 2015;5(May):1–18. Available from: <http://dx.doi.org/10.1038/srep13517>.
- [12] Garrido D, Ruiz-Moyano S, Jimenez-Espinoza R, et al. Utilization of galactooligosaccharides by *Bifidobacterium longum* subsp. *infantis* isolates. *Food Microbiology*. 2013;33(2):262–270. Available from: <http://dx.doi.org/10.1016/j.fm.2012.10.003>.
- [13] Moro G, Arslanoglu S, Stahl B, et al. A mixture of prebiotic oligosaccharides reduces the incidence

- of atopic dermatitis during the first six months of age. *Archives of Disease in Childhood*. 2006; 91(10):814–819.
- [14] Underwood MA, German JB, Lebrilla CB, et al. Bifidobacterium longum subspecies infantis: Champion colonizer of the infant gut. *Pediatric Research*. 2015;77(1):229–235.
- [15] Karav S, Casaburi G, Frese SA. Reduced colonic mucin degradation in breastfed infants colonized by Bifidobacterium longum subsp. infantis EVC001. *FEBS Open Bio*. 2018;8(10):1649–1657.
- [16] Ojima MN, Jiang L, Arzamasov AA, et al. Priority effects shape the structure of infant-type Bifidobacterium communities on human milk oligosaccharides. *ISME Journal*. 2022;16(9):2265–2279.
- [17] Bäckhed F, Roswall J, Peng Y, et al. Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host and Microbe*. 2015;17(5):690–703.
- [18] Lewis ZT, Totten SM, Smilowitz JT, et al. Maternal fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants. *Microbiome*. 2015;3(1):15–17.
- [19] De Vuyst L, Moens F, Selak M, et al. Summer Meeting 2013: Growth and physiology of bifidobacteria. *Journal of Applied Microbiology*. 2014;116(3):477–491.
- [20] Thomson P, Medina DA, Garrido D. Human milk oligosaccharides and infant gut bifidobacteria: Molecular strategies for their utilization. *Food Microbiology*. 2018;75:37–46. Available from: <https://doi.org/10.1016/j.fm.2017.09.001>.
- [21] Centanni M, Ferguson SA, Sims IM, et al. 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*. 2019;85(7):1–17.
- [22] Cremer J, Melbinger A, Wienand K, et al. Cooperation in Microbial Populations: Theory and Experimental Model Systems. *Journal of Molecular Biology*. 2019;431(23):4599–4644. Available from: <https://doi.org/10.1016/j.jmb.2019.09.023>.
- [23] Gore J, Youk H, Van Oudenaarden A. Snowdrift game dynamics and facultative cheating in yeast. *Nature*. 2009;459(7244):253–256.
- [24] Dukovski I, Bajić D, Chacón JM, et al. A metabolic modeling platform for the computation of microbial ecosystems in time and space (COMETS). *Nature Protocols*. 2021 nov;16(11):5030–5082. Available from: <https://www.nature.com/articles/s41596-021-00593-3>.
- [25] Chan SHJ, Simons MN, Maranas CD. SteadyCom: Predicting microbial abundances while ensuring community stability. *PLoS Computational Biology*. 2017;13(5):1–25.
- [26] van Hoek MJA, Merks RMH. Emergence of microbial diversity due to cross-feeding interactions in a spatial model of gut microbial metabolism. *BMC Systems Biology*. 2017;11(1):1–18.
- [27] Bauer E, Zimmermann J, Baldini F, et al. BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLoS Computational Biology*. 2017;13(5):1–22.
- [28] Versluis DM, Schoemaker R, Looijesteijn E, et al. A Multiscale Spatiotemporal Model Including a Switch from Aerobic to Anaerobic Metabolism Reproduces Succession in the Early Infant Gut Microbiota. *mSystems*. 2022 oct;7(5). Available from: <https://journals.asm.org/doi/10.1128/mSystems.00446-22>.
- [29] Versluis DM, Schoemaker R, Looijesteijn E, et al. 2'-Fucosyllactose helps butyrate producers outgrow competitors in infant gut microbiota simulations. *iScience*. 2024;27. Available from: <https://doi.org/10.1016/j.isci.2024.109085>.
- [30] Bunesova V, Lacroix C, Schwab C. Mucin Cross-Feeding of Infant Bifidobacteria and Eubacterium hallii. *Microbial Ecology*. 2018;75(1):228–238.
- [31] Berger B, Porta N, Foata F, et al. Linking human milk oligosaccharides, infant fecal community types, and later risk to require antibiotics. *mBio*. 2020;11(2):1–18.
- [32] Tsukuda N, Yahagi K, Hara T, et al. Key bacterial taxa and metabolic pathways affecting gut short-chain fatty acid profiles in early life. *The ISME Journal*. 2021 sep;15(9):2574–2590. Available from: <http://dx.doi.org/10.1038/s41396-021-00937-7>.
- [33] Van Leeuwen SS, Kuipers BJ, Dijkhuizen L, et al. 1 H NMR analysis of the lactose/ β -galactosidase-derived galacto-oligosaccharide components of Vivinal® GOS up to DP5. *Carbohydrate Research*. 2014;400:59–73. Available from: <http://dx.doi.org/10.1016/j.carres.2014.08.012>.
- [34] Møller PL, Jørgensen F, Hansen OC, et al. Intra- and Extracellular β -Galactosidases from Bifidobacterium bifidum and B. infantis: Molecular Cloning, Heterologous Expression, and Comparative Characterization. *Applied and Environmental Microbiology*. 2001;67(5):2276–2283.
- [35] Hinz SW, Pastink MI, Van Den Broek LA, et al. Bifidobacterium longum endogalactanase liberates galactotriose from type I galactans. *Applied and Environmental Microbiology*. 2005;71(9):5501–5510.

- [36] O'Connell Motherway M, Kinsella M, Fitzgerald GF, et al. Transcriptional and functional characterization of genetic elements involved in galacto-oligosaccharide utilization by *Bifidobacterium breve* UCC2003. *Microbial Biotechnology*. 2013;6(1):67–79.
- [37] Kelly RJ, Rouquier S, Giorgi D, et al. Sequence and expression of a candidate for the human Secretor blood group $\alpha(1,2)$ fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *Journal of Biological Chemistry*. 1995;270(9):4640–4649. Available from: <http://dx.doi.org/10.1074/jbc.270.9.4640>.
- [38] Thorman AW, Adkins G, Conrey SC, et al. Gut Microbiome Composition and Metabolic Capacity Differ by FUT2 Secretor Status in Exclusively Breastfed Infants. *Nutrients*. 2023;15(2):1–13.
- [39] Holms W. 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*. 1986;:69–105 Available from: <http://dx.doi.org/10.1016/B978-0-12-152828-7.50004-4>.
- [40] Magnúsdóttir S, Thiele I. Modeling metabolism of the human gut microbiome. *Current Opinion in Biotechnology*. 2018;51:90–96.
- [41] Gotoh A, Katoh T, Sakanaka M, et al. Sharing of human milk oligosaccharides degradants within bifidobacterial communities in faecal cultures supplemented with *Bifidobacterium bifidum*. *Scientific Reports*. 2018;8(1):1–14. Available from: <http://dx.doi.org/10.1038/s41598-018-32080-3>.
- [42] Turrone F, Milani C, Duranti S, et al. Deciphering bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics approach. *ISME Journal*. 2016;10(7):1656–1668.
- [43] Wong JPH, Fischer-Stettler M, Zeeman SC, et al. Fluid flow structures gut microbiota biofilm communities by distributing public goods. *Proceedings of the National Academy of Sciences*. 2023 Jun;120(25):2017. Available from: <https://pnas.org/doi/10.1073/pnas.2217577120>.
- [44] Celiker H, Gore J. Competition between species can stabilize public-goods cooperation within a species. *Molecular Systems Biology*. 2012;8(621):1–9.
- [45] Garber JM, Nothaft H, Pluvinage B, et al. The gastrointestinal pathogen *Campylobacter jejuni* metabolizes sugars with potential help from commensal *Bacteroides vulgatus*. *Communications Biology*. 2020;3(1):1–11.
- [46] Inglis RF, Biernaskie JM, Gardner A, et al. Presence of a loner strain maintains cooperation and diversity in well-mixed bacterial communities. *Proceedings of the Royal Society B: Biological Sciences*. 2016;283(1822).
- [47] Lindsay RJ, Pawlowska BJ, Gudelj I. Privatization of public goods can cause population decline. *Nature Ecology and Evolution*. 2019;3(8):1206–1216. Available from: <http://dx.doi.org/10.1038/s41559-019-0944-9>.
- [48] Sierra C, Bernal MJ, Blasco J, et al. Prebiotic effect during the first year of life in healthy infants fed formula containing GOS as the only prebiotic: a multicentre, randomised, double-blind and placebo-controlled trial. *European Journal of Nutrition*. 2014;54(1):89–99.
- [49] Pham VT, Lacroix C, Braegger CP, et al. Lactate-utilizing community is associated with gut microbiota dysbiosis in colicky infants. *Scientific Reports*. 2017;7(1):1–13. Available from: <http://dx.doi.org/10.1038/s41598-017-11509-1>.
- [50] Dogra S, Sakwinska O, Soh SE, et al. Rate of establishing the gut microbiota in infancy has consequences for future health. *Gut Microbes*. 2015;6(5):321–325.
- [51] La Rosa PS, Warner BB, Zhou Y, et al. Patterned progression of bacterial populations in the premature infant gut. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111(34):12522–12527.
- [52] Dogra S, Sakwinska O, Soh SEE, et al. Dynamics of infant gut microbiota are influenced by delivery mode and gestational duration and are associated with subsequent adiposity. *mBio*. 2015;6(1):1–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25650398>.
- [53] Joseph C, Zafeiropoulos H, Bernaerts K, et al. Predicting microbial interactions with approaches based on flux balance analysis: an evaluation. *BMC Bioinformatics*. 2024;25(1):1–24. Available from: <https://doi.org/10.1186/s12859-024-05651-7>.
- [54] Scott WT, Benito-Vaquerizo S, Zimmermann J, et al. A structured evaluation of genome-scale constraint-based modeling tools for microbial consortia. *PLoS Computational Biology*. 2023;19(8 August):1–32.
- [55] Jansma J, El Aidy S. Understanding the host-microbe interactions using metabolic modeling. *Microbiome*. 2021;9(1):1–14.

- [56] Marinos G, Kaleta C, Waschina S. Defining the nutritional input for genome-scale metabolic models: A roadmap. *PLoS ONE*. 2020;15(8):1–17. Available from: <http://dx.doi.org/10.1371/journal.pone.0236890>.
- [57] de Groot DH, Lischke J, Muolo R, et al. The common message of constraint-based optimization approaches: overflow metabolism is caused by two growth-limiting constraints. *Cellular and Molecular Life Sciences*. 2020;77(3):441–453. Available from: <https://doi.org/10.1007/s00018-019-03380-2>.
- [58] Martin FPJ, Moco S, Montoliu I, et al. Impact of breast-feeding and high-and low-protein formula on the metabolism and growth of infants from overweight and obese mothers. *Pediatric Research*. 2014;75(4):535–543.
- [59] Hascoët JM, Hubert C, Rochat F, et al. Effect of formula composition on the development of infant gut microbiota. *Journal of Pediatric Gastroenterology and Nutrition*. 2011;52(6):756–762.
- [60] Rao C, Coyte KZ, Bainter W, et al. Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature*. 2021;591(7851):633–638. Available from: <http://dx.doi.org/10.1038/s41586-021-03241-8>.
- [61] Orth JD, Thiele I, Palsson BO. What is flux balance analysis? *Nature Biotechnology*. 2010; 28(3):245–248. Available from: <http://dx.doi.org/10.1038/nbt.1614>.
- [62] Feist AM, Palsson BO. The biomass objective function. *Current Opinion in Microbiology*. 2010; 13(3):344–349. Available from: <http://dx.doi.org/10.1016/j.mib.2010.03.003>.
- [63] Belzer C, Chia LW, Aalvink S, et al. Microbial metabolic networks at the mucus layer lead to diet-independent butyrate and vitamin B12 production by intestinal symbionts. *mBio*. 2017;8(5).
- [64] Corfield AP. The interaction of the gut microbiota with the mucus barrier in health and disease in human. *Microorganisms*. 2018;6(3):18–26.
- [65] Thurl S, Munzert M, Henker J, et al. Variation of human milk oligosaccharides in relation to milk groups and lactational periods. *British Journal of Nutrition*. 2010;104(9):1261–1271.
- [66] Boehm G, Stahl B, Jelinek J, et al. Prebiotic carbohydrates in human milk and formulas. *Acta Paediatrica, International Journal of Paediatrics, Supplement*. 2005;94(449):18–21.
- [67] Shiba T, Aiba Y, Ishikawa H, et al. The suppressive effect of bifidobacteria on *Bacteroides vulgatus*, a putative pathogenic microbe in inflammatory bowel disease. *Microbiology and Immunology*. 2003; 47(6):371–378.
- [68] Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. *Cell*. 2014;157(1):121–141. Available from: <http://dx.doi.org/10.1016/j.cell.2014.03.011>.
- [69] Klement E, Cohen RV, Boxman J, et al. Breastfeeding and risk of inflammatory bowel disease: a systematic review with meta-analysis. *The American Journal of Clinical Nutrition*. 2004 nov;80(5):1342–1352. Available from: <https://academic.oup.com/ajcn/article/82/2/486/4863019>.
- [70] Kang Y, Park H, Choe BH, et al. The Role and Function of Mucins and Its Relationship to Inflammatory Bowel Disease. *Frontiers in Medicine*. 2022;9(May):1–7.
- [71] Goehring KC, Marriage BJ, Oliver JS, et al. 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*. 2016;146(12):2559–2566.
- [72] Azagra-Boronat I, Massot-Cladera M, Mayneris-Perxachs J, et al. Immunomodulatory and Prebiotic Effects of 2'-Fucosyllactose in Suckling Rats. *Frontiers in immunology*. 2019;10(July):1773.
- [73] Agostoni C, Riva E, Giovannini M. Dietary fiber in weaning foods of young children. *Pediatrics*. 1995;96(5 II SUPPL.):1002–1005.
- [74] Magnúsdóttir S, Heinken A, Kutt L, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nature Biotechnology*. 2017;35(1):81–89. Available from: <http://dx.doi.org/10.1038/nbt.3703>.
- [75] Turrone F, Peano C, Pass DA, et al. Diversity of bifidobacteria within the infant gut microbiota. *PLoS ONE*. 2012;7(5):20–24.
- [76] Appert O, Garcia AR, Frei R, et al. Initial butyrate producers during infant gut microbiota development are endospore formers. *Environmental Microbiology*. 2020;22(9):3909–3921.
- [77] Falony G, Lazidou K, Verschaeren A, et al. In vitro kinetic analysis of fermentation of prebiotic inulin-type fructans by *Bifidobacterium* species reveals four different phenotypes. *Applied and Environmental Microbiology*. 2009;75(2):454–461.
- [78] Clark DP. The fermentation pathways of *Escherichia coli*. *FEMS Microbiology Letters*. 1989; 63(3):223–234.
- [79] Parche S, Beleut M, Rezzonico E, et al. Lactose-over-Glucose Preference in *Bifidobacterium longum*

- NCC2705: *glcP*, Encoding a Glucose Transporter, Is Subject to Lactose Repression. *Journal of Bacteriology*. 2006 feb;188(4):1260–1265. Available from: <https://journals.asm.org/doi/10.1128/JB.188.4.1260-1265.2006>.
- [80] Pichler MJ, Yamada C, Shuoker B, et al. Butyrate producing colonic Clostridiales metabolise human milk oligosaccharides and cross feed on mucin via conserved pathways. *Nature Communications*. 2020;11(1). Available from: <http://dx.doi.org/10.1038/s41467-020-17075-x>.
- [81] Hamed KA, Dormitzer PR, Su CK, et al. *Haemophilus parainfluenzae* endocarditis: application of a molecular approach for identification of pathogenic bacterial species. *Clinical Infectious Diseases*. 1994;19(4):677–683.
- [82] Uranga CC, Arroyo P, Duggan BM, et al. Commensal Oral *Rothia mucilaginosa* Produces Enterobactin, a Metal-Chelating Siderophore. *mSystems*. 2020;5(2).
- [83] Hillman ET, Kozik AJ, Hooker CA, et al. Comparative genomics of the genus *Roseburia* reveals divergent biosynthetic pathways that may influence colonic competition among species. *Microbial Genomics*. 2020;6(7):7–24.
- [84] Sakanaka M, Hansen ME, Gotoh A, et al. Evolutionary adaptation in fucosyllactose uptake systems supports bifidobacteria-infant symbiosis. *Science Advances*. 2019;5(8):1–16.
- [85] Viborg AH, Katayama T, Abou Hachem M, et al. Distinct substrate specificities of three glycoside hydrolase family 42 β -galactosidases from *Bifidobacterium longum* subsp. *infantis* ATCC 15697. *Glycobiology*. 2014;24(2):208–216.
- [86] Ashida H, Miyake A, Kiyohara M, et al. Two distinct α -L-fucosidases from *Bifidobacterium bifidum* are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates. *Glycobiology*. 2009;19(9):1010–1017.
- [87] Majewski RA, Domach MM. Simple constrained-optimization view of acetate overflow in *E. coli*. *Biotechnology and Bioengineering*. 1990;35(7):732–738.
- [88] Rubaltelli FF, Larga jolli G. Effect of Light Exposure on Gut Transit Time in Jaundiced Newborns. *Acta Pædiatrica*. 1973;62(2):146–148.
- [89] Hyams JS, Geertsma MA, Etienne NL, et al. Colonic hydrogen production in infants with colic. *The Journal of Pediatrics*. 1989;115(4):592–594.
- [90] Grummer-Strawn LM, Scanlon KS, Fein SB. Infant feeding and feeding transitions during the first year of life. *Pediatrics*. 2008;122(SUPPL. 2).
- [91] Ballard O, Morrow AL. Human Milk Composition. Nutrients and Bioactive Factors. *Pediatric Clinics of North America*. 2013;60(1):49–74. Available from: <http://dx.doi.org/10.1016/j.pcl.2012.10.002>.
- [92] Kawasaki K. Diffusion constants near the critical point for time-dependent ising models. I. *Physical Review*. 1966;145(1):224–230.
- [93] Rogosa M. the Genus *Veillonella*. I. General Cultural, Ecological, and Biochemical Considerations. *Journal of Bacteriology*. 1964;87:162–170.
- [94] Goodfellow M, Kämpfer P, Busse HJ, et al. *Bergey's Manual of Systematic Bacteriology: Volume Five The Actinobacteria, Part A*. Springer; 2012.
- [95] Yokoyama SI, Suzuki T. Isolation and characterization of a novel equol-producing bacterium from human feces. *Bioscience, Biotechnology and Biochemistry*. 2008;72(10):2660–2666.
- [96] Vos P, Garrity G, Jones D, et al. *Bergey's manual of systematic bacteriology: Volume 3: The Firmicutes*. Vol. 3. Springer Science & Business Media; 2011.
- [97] Noor E, Haraldsdóttir HS, Milo R, et al. Consistent Estimation of Gibbs Energy Using Component Contributions. *PLoS Computational Biology*. 2013;9(7).
- [98] Szatmári D, Sárkány P, Kocsis B, et al. Intracellular ion concentrations and cation-dependent remodelling of bacterial MreB assemblies. *Scientific Reports*. 2020;10(1):1–13. Available from: <https://doi.org/10.1038/s41598-020-68960-w>.
- [99] Slonczewski JL, Fujisawa M, Dopson M, et al. Cytoplasmic pH Measurement and Homeostasis in Bacteria and Archaea. *Advances in Microbial Physiology*. 2009;55. Available from: [http://dx.doi.org/10.1016/S0065-2911\(09\)05501-5](http://dx.doi.org/10.1016/S0065-2911(09)05501-5).
- [100] Siegumfeldt H, Björn Rechinger K, Jakobsen M. Use of fluorescence ratio imaging for intracellular pH determination of individual bacterial cells in mixed cultures. *Microbiology*. 1999;145(7):1703–1709.
- [101] Valentin J, Streffer C. Basic anatomical and physiological data for use in radiological protection: Reference values - ICRP Publication 89. *Annals of the ICRP*. 2002;32(3-4):1–277.
- [102] Davis WS, Parker Allen R, Favara BE, et al. Neonatal small left colon syndrome. *American journal of Roentgenology*. 1974;120(2):322–329.

- [103] Palmer C, Bik EM, DiGiulio DB, et al. Development of the human infant intestinal microbiota. *PLoS Biology*. 2007;5(7):1556–1573.
- [104] Press WH. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press; 2007.
- [105] Allen M, Poggiali D, Whitaker K, et al. Raincloud plots: A multi-platform tool for robust data visualization [version 1; peer review: 2 approved]. *Wellcome Open Research*. 2019;4:1–52.