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

David M. Versluis[1], Ruud Schoemaker [3], Ellen Looijesteijn[3], Daniël Muysken [4¤], Prescilla V. Jeurink [3], Marcel Paques[3], Jan M. W. Geurts[3], Roeland M. H. Merks[1,2*]

1 Institute of Biology, Leiden University, P.O. Box 9505, 2300 RA Leiden, The Netherlands
2 Mathematical Institute, Leiden University, P.O. Box 9512, 2300 RA Leiden, The Netherlands
3 FrieslandCampina, 3818LE, Amersfoort, the Netherlands.
4 Centrum Wiskunde & Informatica (CWI), 1098 XG Amsterdam, The Netherlands

¤Current Address: IBIVU – Center for Integrative Bioinformatics, Vrije Universiteit Amsterdam, Amsterdam 1081HV, The Netherlands
* r.m.h.merks@biology.leidenuniv.nl

Abstract

The human intestinal microbiome starts to form immediately after birth, and can greatly influence the health of the infant. During the first days facultative anaerobic species generally dominate, followed by a dominance of strictly anaerobic species, particularly *Bifidobacterium* species. An early transition to *Bifidobacterium* is associated with health benefits. To study the mechanisms of this transition and its hypothesised relation to oxygen, we introduce a multiscale mathematical model that considers metabolism, spatial bacterial population dynamics and resource sharing. Based on publicly available metabolic network data, the model predicts that differences in oxygen availability explain some of the observed individual variation in succession to anaerobic species. The model also predicts that anaerobic *Bifidobacterium* species become dominant through metabolizing lactose with a suboptimal yield, but a higher anaerobic growth rate than its competitors. The current work is the first step towards a more comprehensive understanding of the formation of a steady state adult colonic microbiota.

Introduction

The intestinal microbiota is key to human health and disease [60]. The composition of the adult microbiota differs from person to person, and it is partially established in the infant. The composition of the infant microbiota likely impacts infant health and development. Thus it is important to understand the factors that contribute to the formation of the infant microbiota.

The community types of the infant microbiota differs greatly from those of the adult microbiota. In the infant, three main community types are observed: (1) Proteobacteria-dominated, often Enterobacteriaceae such as *Escherichia coli*; (2) Actinobacteria-dominated, nearly always *Bifidobacterium* spp.; (3) Bacilli-dominated, which is less frequent [18]. None of these community types are primary community types in adults. They establish sequentially through a characteristic succession. In the first 24 to 48 hours after birth, the Proteobacteria and Bacilli clusters are the most common [18, 57]. In the following weeks Proteobacteria clusters are often replaced by Actinobacteria, whereas Bacilli clusters are succeeded by either Proteobacteria or Actinobacteria. Other anaerobic species are also found, but typically do not dominate. The main Actinobacteria in the infant gut are *Bifidobacterium* species [18]. The genus *Bifidobacterium* is of particular interest in the infant microbiota. It is not abundant in the adult gut because the presence of lactose, the most abundant glycan in human milk and most infant formulas, is essential for acquiring high levels of *Bifidobacterium* [25]. *Bifidobacterium* spp. is associated with a range of health benefits. For example, early succession to a *Bifidobacterium* spp. dominated microbiota in early infancy correlated with a reduced probability to be underweight at 18 months of age or to be overweight at 7 years of age.

The mechanisms underlying the ecology of the infant microbiota are still elusive. Shifts are attributed to developmental changes in the environment, such as weaning or the maturation of the gut, and to qualities of the microbial species themselves [18]. A possible trigger for the replacement of the main Proteobacteria, Enterobacteriaceae, facultative anaerobes, by the main Actinobacteria, *Bifidobacterium* spp, obligate anaerobes, is depletion of a hypothesized initial amount of oxygen [46]. Oxygen is known to diffuse into the gut lumen from the body, and be depleted by microbes, colonocytes and by non-biological chemical processes in the cecal contents [26]. The depletion of oxygen is considered to be consistent with a rapid consumption by the facultative aneaerobes [26]. However, it remains unclear to what extent oxygen is depleted created due to oxygen consumption of the bacteria, or due to non-microbial processes in the gut or consumption by colonocytes [26, 39]. Other potentially relevant mechanisms in the succession to *Bifidobacterium* spp. include the strength of indirect competition of *Bifidobacterium* spp., which can dominate the gut even in the absence of *Bifidobacterium*-specific nutrients [8].

[19, 36]. Acetate and lactate produced by *Bifidobacterium* spp. also contribute to the acidification of the infant gut, which suppresses many pathogens [20]. These data suggest that *Bifidobacterium* spp. positively affect the infant host. A key question, therefore, is if and how *Bifidobacterium* can be promoted. The composition and further development of the infant microbiota is thought to be determined by many factors, including nutrition [4, 8, 67], i.e. human milk or infant formula.

To explore the key conditions and mechanisms for shifts to *Bifidobacterium*-rich community types requires insight into the underlying metabolic and ecological mechanisms [10]. Many previous models of the human microbiota make use of mathematical models to provide metabolic insight [43]. FBA works by mathematically analysing the flux of metabolites through a metabolic network [48]. This network is made up of reactions, each of which specifies input and output metabolites. FBA requires constraints on metabolic uptake and an objective function, which represents the metric that should be optimized, usually biomass production. Together with the constraint that there must be an internal steady state - there can be no net change of metabolites within the cell - these are used to calculate what reactions should be used to maximize the objective function. This set of reactions and the flux through them is the outcome of FBA. Earlier approaches have introduced spatial FBA models of metabolizing microbial communities [43] including spatial models suitable for the gut microbiota [6, 12, 32]. Steadycom focused on generating plausible stable human gut communities using a FBA framework to integrate bacterial models [12]. This framework required the assumption that the entire system is in a steady state. In contrast, BacArena [6] and our own approach focused on individual populations in a system that is not necessarily in a steady state [32]. This allows for the modelling of dynamic environments and populations. SteadyCom [12] and our own approach [32] were hampered by the lack of curated genome-scale models (GEMs) of the key species found in the human gut microbiota. BacArena incorporated seven species from the AGORA set [6]. AGORA includes 773 semi-automatically reconstructed GEMs of human gut bacteria, including all major infant species [42].

The infant colon presents several unique obstacles that preclude the use of these existing models created specifically for the adult gut. They do not cover the dynamic changes in populations during the initial colonisation and subsequent succession phases, which are distinctive features of the infant colon [18, 23]).

To explore the development of the gut microbiota and potential factors promoting succession by *Bifidobacterium* spp, here we present a computational model of microbial ecology in the infant gut based upon our previous model of the gut microbiota [32]. We represent the first three weeks of development of the infant microbiota in a spatial context. Metabolism is simulated on the level of local populations using flux balance analysis (FBA) with additional constraints representing metabolic trade-offs in the bacteria. Using this model, we find that the dominance of *Bifidobacterium* spp, the presence of other species, and the initial dominance of Enterobacteriacea can be explained through the initial presence and microbial depletion of oxygen, combined with their unique metabolisms.

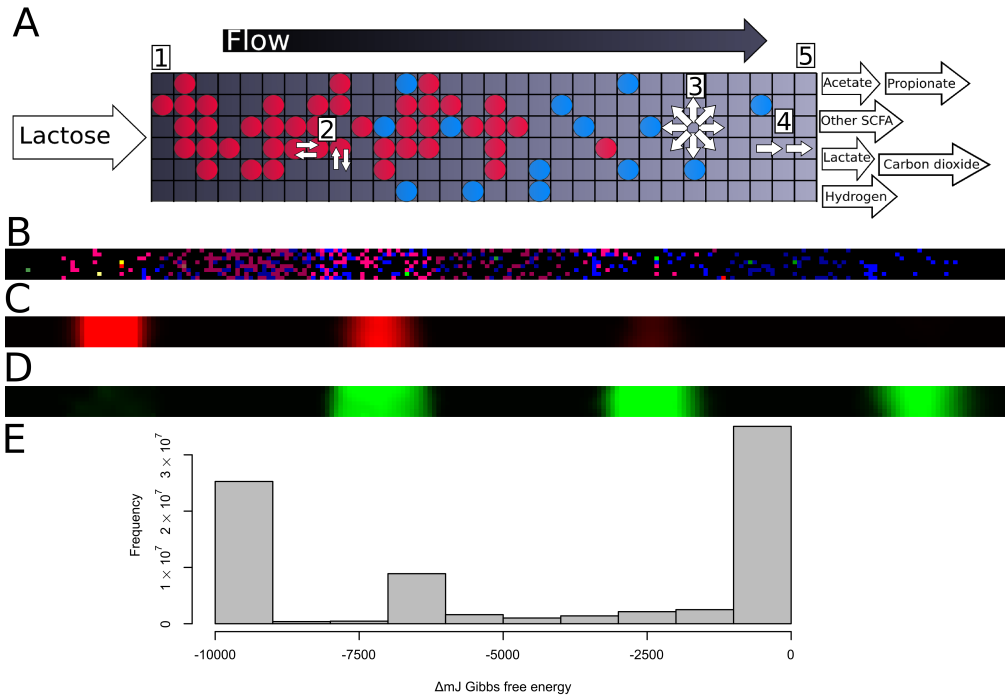

Fig 1. The multi-scale metabolic model
(A) Schematic of the model at a system level. Circles represent bacterial populations, color represents species. Flow through the tube is from left (proximal) to right (distal). Lactose is placed at the proximal side. Output metabolites are examples, and depend on bacterial metabolism. Lattice dimensions and ratio are schematic. (B) Screenshot of the bacterial layer of the model at a single timepoint. Color indicates species, brightness indicates growth rate on the current step. (C) Screenshot of lactose concentration of the model corresponding to B. Intensity indicates concentration. (D) Screenshot of lactate concentration of the model corresponding to B. Intensity indicates concentration. (E) Histogram displaying the distribution of change in Gibbs free energy for each non-zero FBA solution over 30 runs of the model.

Results

Model outline

The aim of our modeling approach was to provide insights into the dynamics and factors involved in the first 21 days of development of the infant microbiota. We focused particularly on succession in the newborn infant gut, and how this can be explained from the perspective of bacterial metabolisms.

The infant colon is represented by a regular square lattice populated with simulated bacterial populations (Fig.1). Each lattice site can contain metabolites and a single population of a single species. Metabolites enter the system on the proximal side in periodic meals every three hours(Fig. 1A-1). Bacterial populations diffuse through swaps of the content of adjacent lattice sites (Fig. 1A-2). Metabolites diffuse between adjacent lattice sites (Fig. 1A-3) and advect to more distal sites (Fig. 1A-4). Metabolites that leave the system on the distal side (Fig. 1A-5) are removed.

Bacterial populations can grow in their local lattice site depending on the available metabolites, and will divide when they grow too large. To calculate growth and metabolic output, we use FBA. FBA utilizes a reaction network and an objective function to calculate what set of reactions maximizes flux through an objective function. We calculate FBA solutions individually for each local population at each timepoint. Each species is represented by a GEM specific to that species. The FBA solutions depend on both the local metabolite concentrations in the model, which function as uptake reaction

bounds, and the GEM used. We set an objective function in our models requiring only ATP, and returning ADP and P_i. This makes the reaction only use energy, not mass. We also use a flux limit that caps the maximum amount of metabolic flux an FBA solution can have to a value relative to the local size of the population. This alters what FBA solution is reached, and also causes a minimal solution under many conditions.

From the AGORA dataset we selected 15 species, prevalent in the earliest infant gut microbiota [4, 16, 42] (Table 1).

Table 1. Species included in our model runs, all from the AGORA project [42]. Color indicates color used in figures

Species	Family	Anaerobic? [42]	Reactions
Bacteroides vulgatus	Bacteroidaceae	Yes	2474
Bifidobacterium breve	Bifidobacteriaceae	Yes	1987
Bifidobacterium longum ssp. *infantis*	Bifidobacteriaceae	Yes	1852
B. longum ssp. *longum* DJO10A	Bifidobacteriaceae	Yes	2043
Blautia hansenii	Lachnospiraceae	Yes	2055
Collinsella aerofaciens	Coriobacteriaceae	Yes	920
Dorea formicigenerans	Lachnospiraceae	Yes	2072
Parabacteroides distasonis	Porphyromonadaceae	Yes	2519
Ruminococcus gnavus	Lachnospiraceae	Yes	2225
Veillonella dispar	Veillonellaceae	Yes	1124
Enterococcus faecalis	Enterococcaceae	Facultative	1174
Enterobacter cloacae	Enterobacteriaceae	Facultative	1800
Escherichia coli SE11	Enterobacteriaceae	Facultative	2356
Lactobacillus gasseri	Lactobacillaceae	Facultative	1163
Streptococcus salivarius	Streptococcaceae	Facultative	1190

The parameters of the system were determined based on a number of sources including the typical length and diameter of the infant colon [13, 66] and an estimation of the colonic transit time [34, 54]; a complete description is in Section Parameters. No data was available for the flux constraint, the growth per unit of ATP, and the input of new bacterial populations. For these parameters we have performed a sensitivity analysis around the default parameter set shown in Table 2.

We initialize the system by placing many small populations of all species randomly across the lattice. We run the model in timesteps, each of which we define to be three minutes of real time. A timestep proceeds as follows: first the uptake bounds for each population are set according to the local concentration of metabolites. Water is always available. Each local population performs its metabolism, and the local metabolite concentrations are changed accordingly. Populations may now also divide, if they have grown too large for their lattice site. Populations and metabolites then diffuse. To mimic colonization with new species through feeding eachempty lattice site now has a small probability to be filled with a new bacterial population selected at random from all available species (Table 1)

In all experiments, a pulse of lactose is introduced every three hours to all lattice sites in the six most distal columns, to mimic feeding. Lactose is treated like any other metabolite. In some simulations, we initialize the system with oxygen spread evenly across all squares. Oxygen is not moved distally like other metabolites are, but it is diffused.

Figures 1 B-D show screenshots of a typical simulation. Many populations are present across the simulated gut (Fig. 1B, color code in Table 1). Lactose is introduced at the proximal side, advects distally and diffuses (Fig. 1C). It is taken up by the populations, which use it to grow (indicated by brightness in Fig. 1B), producing metabolites including lactate (Fig. 1D). Some metabolites can, in turn, be further metabolized to drive growth (Fig. 1B). In total, 723 extracellular metabolites are be described in the model, but only 17 of these are typically found in our simulations. Of these, the screenshots display only the input metabolite lactose and the produced metabolite lactate.

We recorded the metabolic and bacterial composition of the model at each timestep. We also recorded the metabolites flowing out of the distal end.

Checks for imbalances

To ensure the biochemical correctness of the FBA solutions we checked each in two ways for possible imbalances. We first checked the mass balance of hydrogen, oxygen, and carbon atoms exchanged with the environment by each bacterial population each step. Because the objective function does not remove any mass, and there are no other sinks, there should be no net change in atoms due to the calculated fluxes. This check uncovered two errors in glycogen metabolism in several metabolic models, that resulted in the energetically favorable destruction of intracellular protons. We replaced the reactions responsible for this erroneous energy source(S1 Table). We also checked each solution for thermodynamic plausibility, using a dataset of Gibbs free energy for common metabolites [24, 47]. Calculating the free energy of all inputs and outputs at each step allowed us to detect any potential FBA solutions that would erroneously increase the amount of free energy in the system. 99.9999% of all fluxes of Fig. 2 contained less free energy in the output than in the input (Fig. 1E). The remainder all had an amount of energy in the outputs equal to that of the inputs and very low growth rates, less than 0.001% of the average growth rate. The sum of these growth rates over 30 runs was less than 0.0002 population units. Additional checks showed that none of the models could have flux through their objective function while the uptake for all metabolites except water was set to 0. After applying these corrections we concluded that the system conserved mass and energy.

Model predicts *Bifidobacterium* dominance through metabolism under anaerobic conditions

After establishing mass and energy conservation, we first studied the behavior of the mathematical model for anaerobic conditions. Fig. 2 summarizes the results of 30 simulations (part of a simulation is recorded in S2 Video). After 21 simulated days (10080 timesteps) *Bifidobacterium* spp. had become the most abundant in all runs (Fig. 2A), a typical situation found in infants of a corresponding age [4, 19]. Also in agreement with *in vivo* infant data, the model predicted the presence of resident populations of *Escherichia coli*. Our model also often predicted considerable levels of *Blautia hansenii*, which is only occasionally present in the infant colon [4]. We will investigate the ecological role of this species further in section 'Non-*Bifidobacterium* species benefit from lactate consumption'. Other species experienced an initial growth, but declined after day 2. These species had abundances either lower or not significantly higher than their initial abundance after 21 simulated days ($p>0.05$, two sample t-test).

We next analysed the metabolites found in the simulated fecal output over the last two days of the simulations. We measured metabolites diffusing or advecting out of the distal end of the gut every step, summing over every 60 timesteps. We found considerable amounts of acetate, formate, lactate, hydrogen, ethanol, and succinate (Fig. 2B). All other metabolites (other than water) were present in much smaller quantities. The high abundance of acetate and lower abundance of lactate agree well with the fecal composition of three month old formula-fed infants [45], and is characteristic of microbiota dominated by *Bifidobacterium* spp. Acetate and lactate are produced by *Bifidobacterium* spp in the characteristic bifid shunt, which produces acetate and lactate in a 3:2 ratio [15]. Formate, hydrogen, ethanol and succinate are also all reported in infant feces [9, 25, 59, 64].

Emergent crossfeeding

The high abundance of *Bifidobacterium* spp. and the growth of *E. coli* and *B. hansenii* (Fig. 2A) in conjunction with the high abundance of secreted metabolites (Fig. 2B) suggested that production of metabolites by *Bifidobacterium* spp. might support additional species. To further analyse such crossfeeding interactions we next analysed the spatial distribution of the primary Bifidobacterial metabolites. From day 19 to 21, lactose was present mostly proximally (Fig. 2C), lactate was most present in the middle of the colon (Fig. 2D), and acetate was most present distally (Fig. 2E). To analyse the role of *Bifidobacterium* spp. in shaping this metabolic pattern, we analysed the location of populations from day 19 to 21 (Fig. 2F). *Bifidobacterium* spp was consistently located proximally,

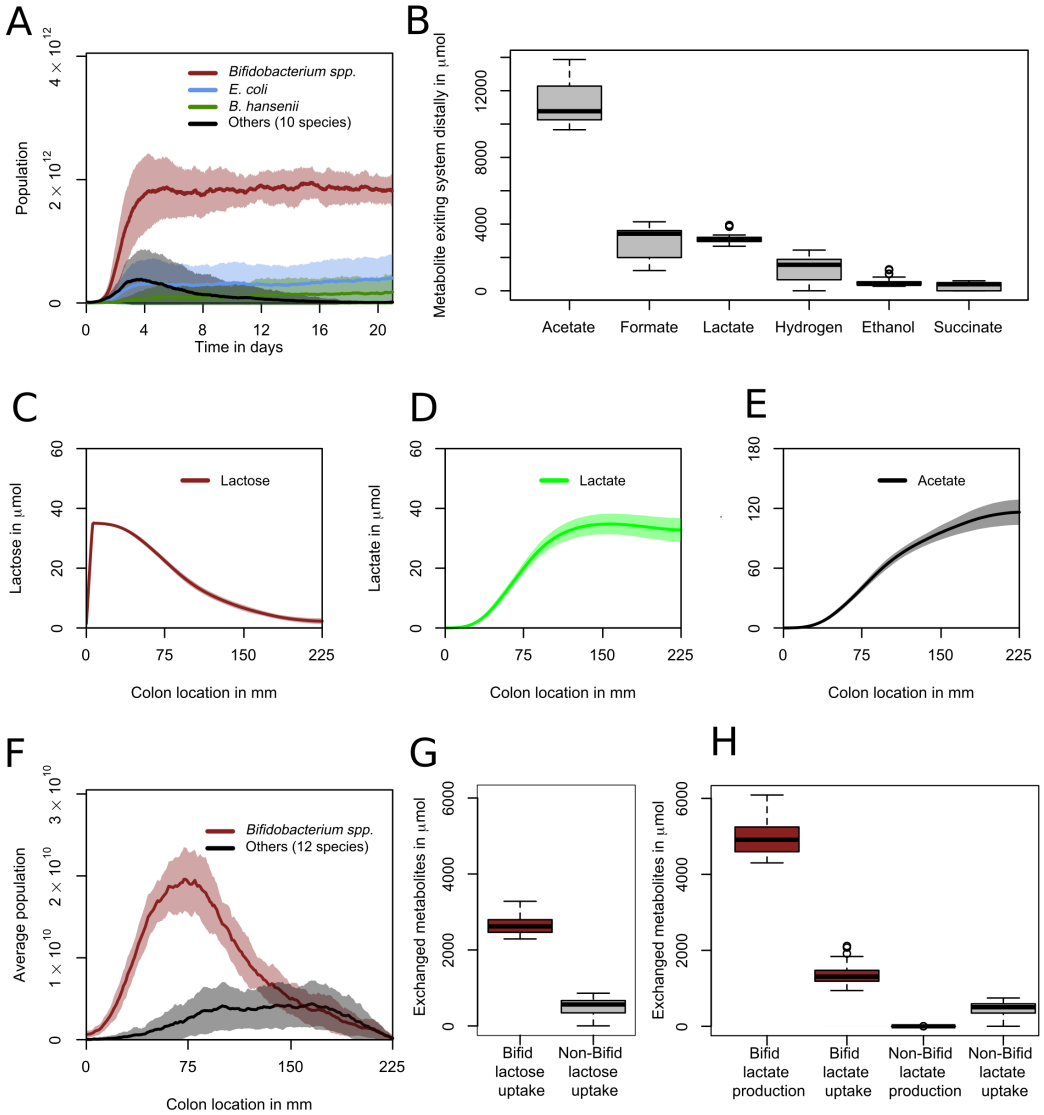

Fig 2. *Bifidobacterium* spp become dominant anaerobically and mostly proximally with a unique metabolic profile

(A) Abundances for grouped species over 21 days. One day is 480 timesteps. Shaded area indicates one standard deviation. n=30. (B) Distribution of metabolites exiting the system distally over the last two days (960 timesteps) of the simulation. n=30. (C) Spatial distribution of lactose over the last two days (960 timesteps) of the simulation. Shaded area indicates one standard deviation, measured over the runs. n=30. (D) Spatial distribution of lactate over the last two days (960 timesteps) of the simulation.Shaded area indicates one standard deviation, measured over the runs. n=30. (E) Spatial distribution of acetate over the last two days (960 timesteps) of the simulation.Shaded area indicates one standard deviation, measured over the runs. n=30. (F)Abundance per location for grouped species over the last two days (960 timesteps). Shaded area indicates one standard deviation. n=30. (G) Metabolism of lactose by grouped species. All data from last two days. n=30. (H) Metabolism of lactate by grouped species. All data from last two days. n=30.

whereas the other species were consistently located more distally. These results suggested that separated metabolic niches had formed in the simulated colon, potentially supporting cross-feeding interactions. To analyze potential crossfeeding interactions we summed the net flux of lactose and lactate for *Bifidobacterium* and for all non-*Bifidobacterium* species grouped together. *Bifidobacterium* spp was the largest consumer of lactose (Fig. 2G) and both the largest producer and consumer of lactate (Fig. 2H). However, the non-*Bifidobacterium* groups consumed more lactate relative to their lactose consumption. From this, we hypothesized that *Bifidobacterium*'s primary niche is the consumption of lactose, while one or more other niches exist for the other species around lactose or lactate consumption. Combined with the observation (Fig. 2D) that lactate concentration is lower in the distal colon, we can conclude that the population at the distal end of the simulated colon consists at least partially of cross-feeders.

Metabolic specialisation on lactose is concentration dependent

Next we asked why *Bifidobacterium* is so strongly present proximally(Fig. 2F). We analysed the ATP production of the three most prevalent species in the experiment of Fig. 2 for varying concentrations of lactose (Fig. 3 A). To ensure they only grew on lactose we performed separate simulations of a single timestep. At high lactose concentrations *Bifidobacterium* had a higher ATP production per timestep than any other species (Fig. 3 A). At lower lactose concentrations, *E. coli* and *B. hansenii* produce ATP at a faster rate than *Bifidobacterium*. This is a result of the populations being limited in how much metabolism they can perform per timestep. Given unlimited lactose, *Bifidobacterium* spp. uses fewer reactions in its central carbon metabolism than the other species in Fig. 3A (31 for *Bifidobacterium* spp., 70 for *E. coli*, and 35 for *B. hansenii*). This allows it to metabolize more lactose and produce more ATP per timestep within the flux limit, despite its lower yield per unit of lactose. This metabolism is the experimentally observed bifid shunt. [15].

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

To determine whether the bifid shunt is essential to *Bifidobacterium* dominance in the simulation, we disabled the bifid shunt by blocking all flux through fructose-6-phosphate phosphoketolase and ran 30 experiments. Fructose-6-phosphate phosphoketolase is unique to and necessary for the bifid shunt, and characteristic of *Bifidobacterium* as a genus [14]. In the absence of fructose-6-phosphate phosphoketolase *Bifidobacterium* cannot metabolize lactose through the usual metabolic route, but instead uses a route involving phosphofructokinase. This route produces less ATP on the same concentration of lactose. *Bifidobacterium* was dominant in none of these thirty simulations (Fig. 3 B). Instead, *E. coli*, *B. hansenii* and *Bacteroides vulgatis* were dominant. This indicates that the bifid shunt is crucial to the dominance of *Bifidobacterium* in our model.

B. hansenii but not *E. coli* depends on crossfeeding

We next analyzed how, alongside the dominance of *Bifidobacterium*, other species could stably persist in the model. We observed that the non-*Bifidobacterium* species did not occur as strongly proximally as *Bifidobacterium* spp. (Fig. 2 F). More distal parts of the colon have a lower lactose concentration compared to the proximal part (Fig. 2C), and in the presence of low concentrations of lactose *some non-Bifidobacterium species such as* E. coli produces more ATP per timestep than *Bifidobacterium* spp. (Fig. 3 A). We therefore hypothesized that these non-*Bifidobacterium* species outcompete *Bifidobacterium* on low-concentration lactose in the model. To test this, we disabled the uptake of lactose by non-*Bifidobacterium* species. To rule out the possibility that the initial populations of non-*Bifidobacterium* species would die out before *Bifidobacterium* had produced sufficiently high-levels of crossfeeding metabolites, we ran 30 simulations in which lactose consumption was disabled after 10 days. This led to a sharp population decrease of non-*Bifidobacterium* species from this point onwards (Fig. 3 C). *E. coli* had a near-zero abundance at the end of the simulations with lactose disabled. *B. hansenii* also had a significantly smaller population compared to the baseline

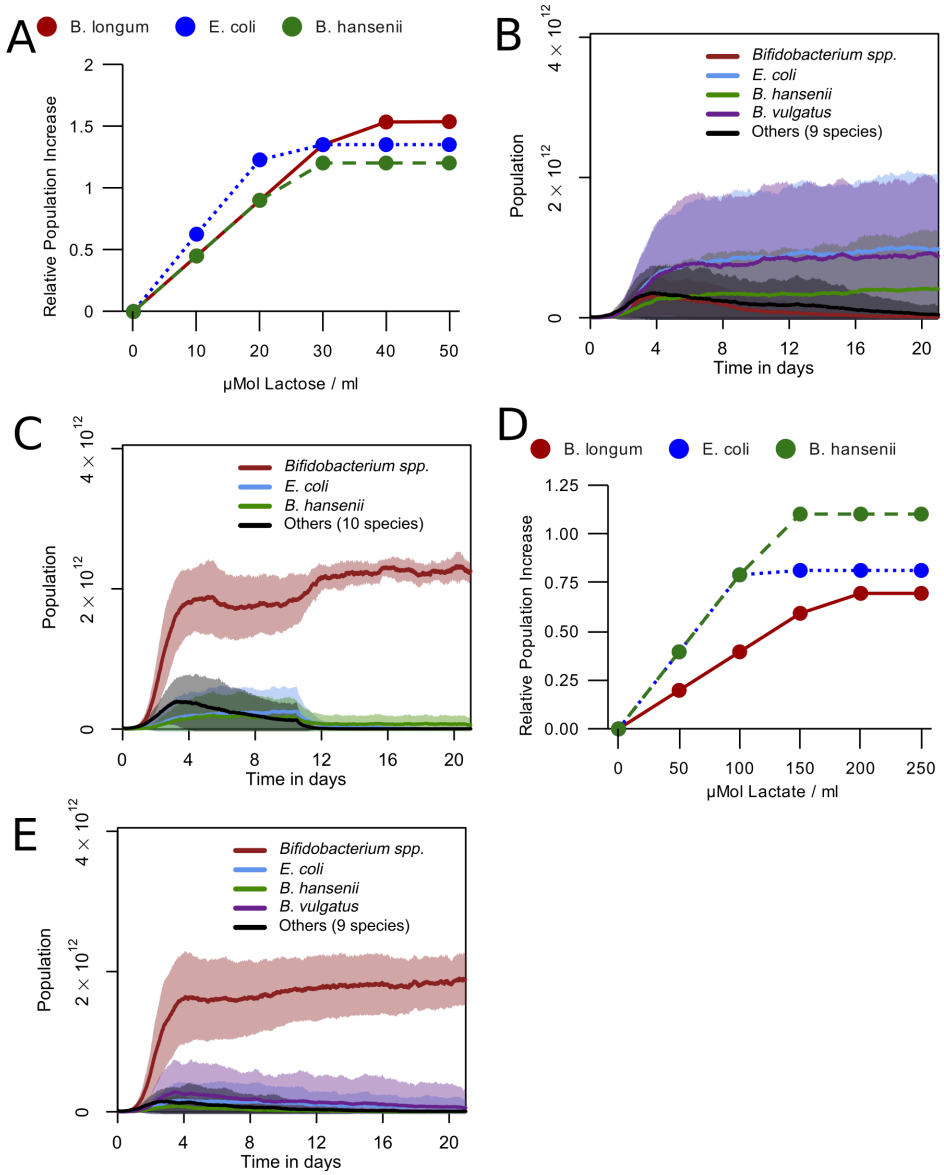

Fig 3. *Bifidobacterium* and *E. coli* require lactose, but *B. hansenii* requires lactate
(A) Growth (through ATP production) per timestep by lactose concentration for some major bacterial species using our FBA with flux limit method. Water uptake was unbound. (B) Abundances for grouped species over 21 days. The bifid shunt was disabled for all *Bifidobacterium* species. One day is 480 timesteps. Shaded area indicates one standard deviation. n=30. (C) Abundances for grouped species over the full length as in B, but with lactose disabled for non-*Bifidobacterium* species. n=30 (D) Abundances for grouped species over the full length of the simulations as in B, but with lactose disabled for non-*Bifidobacterium* species from 4800 steps (10 days) onwards. n=30 (E) Growth (through ATP production) per timestep by lactate concentration for some major bacterial species using our FBA with flux limit method. Water uptake was unbound. (F) Abundances for grouped species, with lactate consumption disabled for all species. One day is 480 timesteps. Shaded area indicates one standard deviation. n=30.

($p<0.01$, two sample t-test), but still had a presence (Fig. 3 C). This indicates that *B. hansenii*, in contrast to the other non-*Bifidobacterium* species, is capable of sustaining itself on a different substrate than lactose.

Non-*Bifidobacterium* species benefit from lactate consumption

To further analyse how *B. hansenii* could grow in the absence of lactose, we investigated the substrates used by non-*Bifidobacterium* species in the model. Lactate is the second most consumed carbohydrate in the model, after lactose. It is primarily produced by *Bifidobacterium* and taken up by most species in our model (Fig. 2D). *Bifidobacterium* takes up lactate in our model and converts it into pyruvate catalyzed by lactate dehydrogenase. *Bifidobacterium* then further converts pyruvate to acetate. *E. coli* and *B. hansenii* also produce pyruvate from lactate, but use different methods to further metabolize the pyruvate, and produce more ATP from lactate than *Bifidobacterium* at any concentration (Fig. 3D). We hypothesize that *Bifidobacterium* cannot compete with *E. coli* and *B. hansenii* species on lactate. This could explain the existence of *B. hansenii* even without lactose uptake.

To investigate the role of lactate in feeding secondary populations, we blocked the lactate uptake reaction for all species and ran 30 model 30 (Fig. 3E). Final *B. hansenii* populations were much smaller than in simulations with functional lactate uptake, not even significantly larger than the initial populations ($p>0.05$, two sample t-test).*E. coli* populations were also significantly smaller than those in simulations with lactate uptake, but also significantly larger than their initial population at the start of the run ($p<0.05$, two sample t-test). *Bifidobacterium* populations were similar to these in the first experiment (Fig. 2A). The average *B. vulgatus* population was also higher than in the previous experiment due to a single run with high *B. vulgatus*. Overall, these results show that in our model *B. hansenii* requires lactate to sustain a population, but neither *Bifidobacterium* spp. nor *E. coli* do.

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

In the anaerobic experiments that we have done so far, *Bifidobacterium* quickly dominated the population, but *in vivo* studies find that during the first few days species of the Enterobacteriaceae cluster are dominant [4, 18]. We examined whether initially present oxygen in the gut could suffice to explain early dominance of Enterobacteriacea followed by dominance of Bifidobacterium spp, a common hypothesis [46, 57]. This oxygen could have diffused through the gut wall [26] and accumulated in the newborn infant gut in the absence of microbes [37]. We ran the model with an initial 0.1 µmol of oxygen per square. This oxygen was allowed to diffuse, but was not moved distally every timestep, thus mimicking an early presence of oxygen. Interestingly, this relatively small pulse of initial oxygen sufficed to reproduce the pattern observed *in vivo* : *E. coli* is initially dominant and *Bifidobacterium* becomes dominant after a few weeks of simulated time (Fig. 4 A, S3 Video and S4 Video). The inital presence of oxygen also strongly affected the other species in the model - *B. hansenii* remained practically absent, while *B. vulgatus* had a higher abundance in the first days of the simulation compared to our experiments without initial oxygen.

It is known that intraluminal oxygen diffusion depends on host tissue oxygenation and colonic blood flow [1, 69]. In absence of precise data on the concentration and variation of oxygen levels in the infant gut, we tested oxygen values from 0.1 to 10 µmol initial oxygen per lattice site. *E. coli* to remained dominant for longer at higher intial concentrations of oxygen (Fig. 4B), while succession took much longer, or did not occur at all within the simulated timeframe (Fig. 4C). Over 30 simulations initialized with oxygen at 10 µmol per lattice site *Bifidobacterium* either dominated the population or remained much smaller with few intermediate cases leading to a bimodal distribution (Fig. 4D). This prediction matches *in vivo* observations: in some infants Enterobacteriaceae dominate the microbiota during the first 21 days [18]. The proportion of infants dominated by Enterobacteriaceae varies depending on the study populations, but can be e.g. 51% after 21 days, similar to the 50% that we observe [18]. However, it is also known that these inter-individual differences are affected by delivery mode and gestational duration [18].

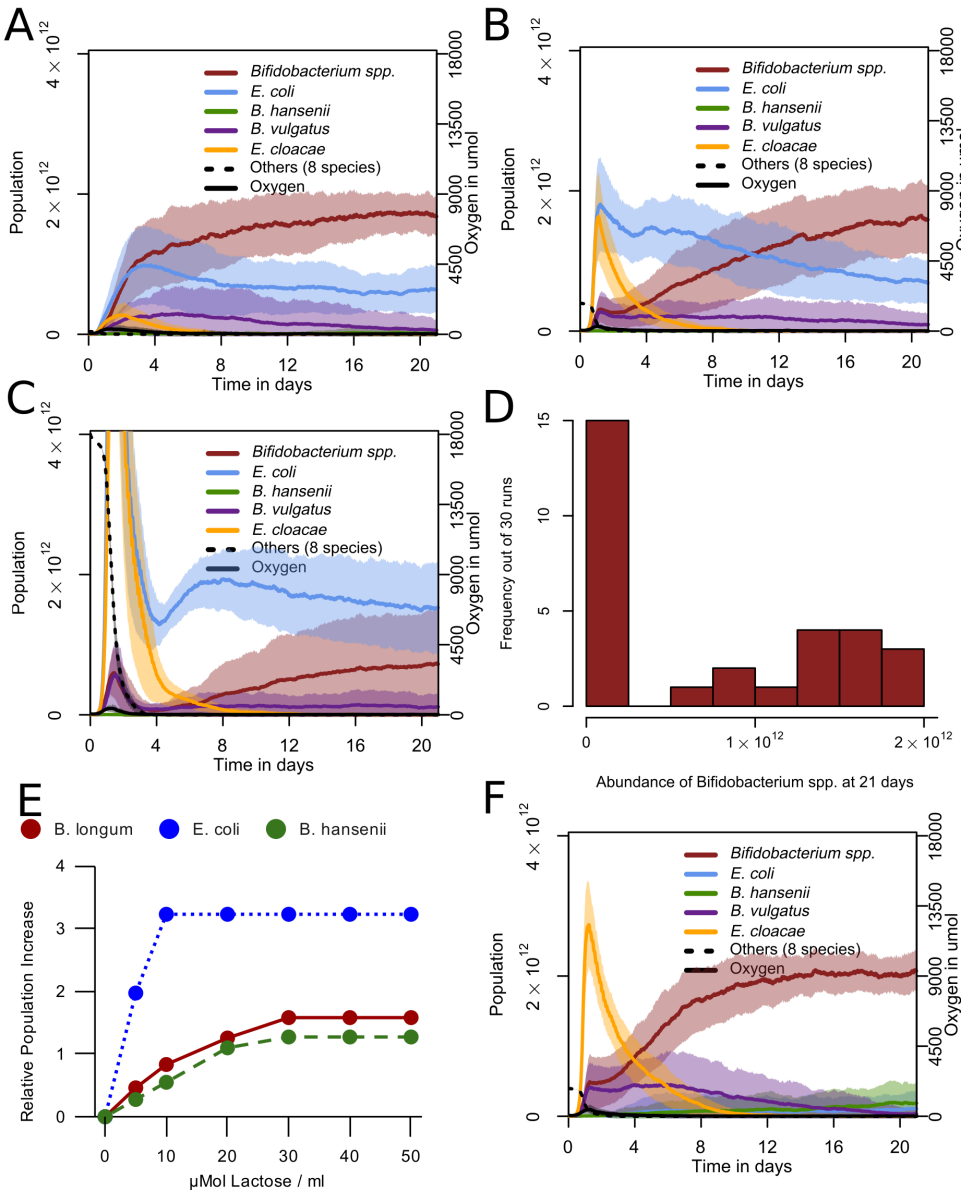

Fig 4. Initial oxygen leads to initial *E. coli* dominance and succession by *Bifidobacterium*

Abundances for grouped species. One day is 480 timesteps. Shaded area indicates one standard deviation. n=30. (A) 0.1 µmol initial oxygen per lattice site (B) 1 µmol initial oxygen per lattice site (C) 10 µmol initial oxygen per lattice site (D) Distribution of total *Bifidobacterium* abundance at 21 days (10080 timesteps) performed with 10 µmol initial oxygen per lattice site n=30. (E) Growth (through ATP production) per timestep by lactose concentration for some major bacterial species using our FBA with flux limit method. Water and oxygen uptake were unbounded. (F) Abundances for grouped species with oxygen uptake disabled for *E. coli*. One day is 480 timesteps. Shaded area indicates one standard deviation. n=30.

We further examined the causes of *E. coli* dominance over *Bifidobacterium* in our model. Fig. 4E shows that under aerobic conditions *E. coli* has a much higher energy production per timestep than *Bifidobacterium* for all concentrations of lactose, even though *Bifidobacterium* also produces more ATP than it does under anaerobic conditions. *B. hansenii*, in contrast, does not use oxygen in its metabolism at all in the model.

To test if the direct uptake of oxygen was indeed responsible for E. coli's dominance of the microbiota, we disabled the oxygen uptake of *E. coli* for 30 simulations with 1 µmol initial oxygen per lattice site initial oxygen. Under these conditions, *E. coli* failed to dominate the population (Fig. 4F), and other non-*Bifidobacterium* species such as *Enterobacter cloacae* became dominant in the early stages (Fig. 4F). They did not maintain their position for as long as *E. coli* did, and are eventually replaced by a population composition similar to that from the anerobic experiments. Altogether, the model predicts that *E. coli* relies on direct consumption of oxygen to dominate the microbiota under aerobic conditions in the model. It also indicates that some other bacterial species in our model can benefit from aerobic conditions, but do not have an anaerobic metabolism competitive enough to sometimes become dominant over *Bifidobacterium* as *E. coli* can.

Sensitivity analysis

Although many of the parameters values in our model were based on data, a number of key parameters remained undetermined, in particular the flux constraint, the growth per unit of ATP, and the input of new bacterial populations (Table 2). To test the effect of these parameter values on the system, we performed a sensitivity analysis under anaerobic conditions.

We first systematically varied the flux limit (Fig. 5 A), which represents the maximum allowed, summed metabolite flux sustained by all reaction in a local population in a single timestep. The flux limit is crucial for a diverse microbiota in which *Bifidobacterium* can become dominant. Increasing the limit to values higher than our default (Table 2) led to a much weaker dominance of *Bifidobacterium*, and further increase to *Bifidobacterium* no longer being dominant. It was overtaken by other species whose metabolisms were more efficient under these conditions. Making the limit infinite (i.e. removing it) led to the practical extinction of *Bifidobacterium* in the model. Without the flux limit, *Bifidobacterium* also no longer uses the bifid shunt. Lowering the flux limit below the default led to a decrease in diversity, with only *Bifidobacterium* remaining, with a much smaller population size. This is also an unrealistic condition - other species are known to coexist with *Bifidobacterium* even when it dominates the microbiota.

We also systematically changed the growth ratio (Fig. 5B). The growth ratio represents the ratio between ATP production and population growth. At a higher ratio, less new population is formed with the same amount of ATP production. An increase of the growth ratio compared to our default parameters (Table 2) led to a much smaller population for all species, but especially the non-*Bifidobacterium* species. This indicates a much lower level of secondary consumption, as various metabolisms are no longer productive enough to exceed the death rate. A decrease in the growth ratio led to much larger populations for all species, particularly The non-*Bifidobacterium* species. This might indicate that there is a more complete fermentation of secondary products with this larger population.

Removing colonization, the continued influx of populations after initialization of the simulation, led to all non-*Bifidobacterium* species going extinct in 7 out of the 30 runs, and all *Bifidobacterium* species going extinct in 6 out of 30 runs. This extinction did not occur in any of the runs with colonization. Increasing the colonization to be above the default value led to a slightly larger but still generally small population for these species. We have maintained it at a moderate level, to show that the system is robust against invasion, and prevent unrealistic complete extinctions from taking place.

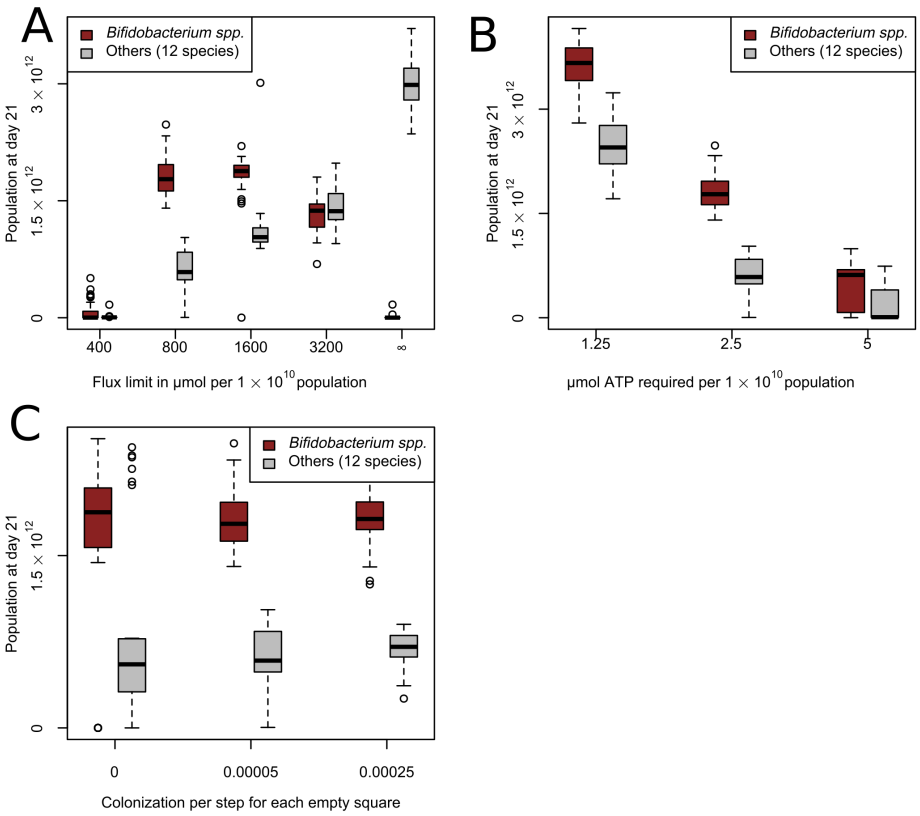

Fig 5. Sensitivity analysis for estimated parameters
Abundances for grouped species over time with different parameters. One day is 480 timesteps. Shaded area indicates one standard deviation. n=30. (A) Effect of flux limit on abundance of *Bifidobacterium* spp. and other species (B) Effect of amount of ATP required for growth on abundance of *Bifidobacterium* spp. and other species (C)Effect of variation in colonization per step per square after the initial population has been placed on abundance of *Bifidobacterium* spp. and other species

Discussion

To develop new hypotheses for the development of the infant gut microbiota, we have presented a dynamic multiscale model. Our simulations suggest that the initial dominance of Enterobacteriaceae, particularly *E. coli*, over *Bifidobacterium* is mediated through the initial presence of oxygen. Our model predictions on oxygen also suggest that variations in oxygen may explain some of the observed variation in infant Bifidobacterium abundance. An absence or strong delay of *Bifidobacterium* development is also observed in some infants [19]. While it is known that there is variation in colonic oxygenation [1, 69], the range of this variation in newborns is not known. Our model suggests that oxygen might be an important influence on the microbiota development. Supporting this, in an *in vivo* study of premature infants, those given supplemental oxygen for longer were more likely to have a strong Enterobacteriaceae signature in their gut microbiota [63]. However, this study was purely observational, and only very few subjects developed a *Bifidobacterium* microbiota, even among those not given supplemental oxygen. Nonetheless, it might be desirable to lower oxygen concentration faster to achieve the positive health effects of *Bifidobacterium* more consistently. This may be manipulable through the infant's nutrition. The direct oxidation of lipids and other organic substrates may contribute to creating anaerobic conditions in the colon [26]. This may allow for a decrease in oxygen in the colon without stimulating facultative anaerobic bacteria. There are also indications that short-chain fatty acids produced by *Bifidobacterium* may play a role in decreasing oxygen levels by stimulating the oxygen use of colonocytes [39]. Although the decrease of oxygen through short-chain fatty acids may play a role in decreasing the oxygen concentration in the gut lumen, our model predicts *Bifidobacterium* is not competitive against facultative anaerobic species such as *E. coli* under aerobic conditions, and cannot reach a large population size under aerobic conditions.

Our simulations also suggest that the dominance of *Bifidobacterium* in the anaerobic infant system can be explained from its bifid shunt metabolism that allows for lactose to be digested faster, despite the ATP yield per molecule of lactose being lower than that of other species. This is a result of our imposition of metabolic limitations. The bifid shunt in our model is an example of overflow metabolism, a common mechanism in bacteria, first characterized in *E. coli*, whereby growth rate is maximized by only generating a part of the potential ATP yield [29]. This process only occurs under resource rich conditions - in our model this process also only occurs at high local levels of lactose. At lower levels of lactose, *Bifidobacterium* in our model switch to a metabolism with a higher yield, which has also been described *in vitro* [15]. The primary difference from *E. coli* here is that for *Bifidobacterium* both metabolisms are anaerobic. Overflow metabolism is found across a wide range of organisms, being known as the Crabtree effect in yeasts and as the Warburg effect in cancer cells [17]. The FBA with flux limit method we use to implement metabolic limits is similar to Flux Balance Analysis with Molecular Crowding [7], which we have used in our previous model [32]. However, in the absence of good reaction-specific thermodynamic information for our current model, we do not utilize the reaction-specific technique of crowding coefficients. We instead base a flux limit only on the size of the local population, which is sufficient for modeling this overflow metabolism [44].

In terms of metabolites produced by *Bifidobacterium* metabolism, the model predicts an acetate:lactate ratio of about 3:1, comparable to that of formula-fed infants [45, 50]. A possible cause is the considerable consumption of lactate by *Bifidobacterium* populations in our model (Fig. 2G). While the proteins used in this pathway exist, it is not known if this process takes place *in vivo*. Strong lactate uptake has only been found *in vitro* when *Bifidobacterium* was grown in co-culture with lactate consuming bacteria [22]. Disabling lactate uptake by *Bifidobacterium* in future versions of the model might lead to higher lactate concentrations in the system. As we have found that lactate is an important crossfeeding metabolite, this might enable different or larger non-*Bifidobacterium* crossfeeding populations.

We do observe a notable production of ethanol (Fig. 2 B). Ethanol is only rarely reported in feces *in vivo* [25]. However *Bifidobacterium* is known to be capable of producing ethanol *in vitro* [15]. This discrepancy might be explained by the lack of gut wall absorption in our model, through which ethanol can be taken up particularly easily [52]. While newborns have reduced ethanol dehydrogenase activity compared to adults, they are capable of metabolising ethanol [35]. We do not expect a large

effect on the metabolism or composition of the microbiota if this were to be implemented - ethanol is little used by the microbiota in our model.

In addition to the dynamics of *Bifidobacterium*, our models also suggests niches for the existence of other species, particularly *B. hansenii* and *E. coli*. A notable discrepancy between the model and the known *in vivo* data is the large presence of *B. hansenii* in the model without initial oxygen. While *Bifidobacterium* and *E. coli* are frequently abundant *in vivo*, *B. hansenii* is present only marginally [4]. Other species that fill the same lactate-consuming niche as *B. hansenii* in our experiment are quite common *in vivo* [51].

Our modeling approach is similar in a broad sense to several previously published frameworks, such as Steadycom, BacArena, and COMETS [6, 12, 31]. For example, Steadycom has previously been used to study longitudinal variation in microbiota occurence, though in an adult context [11]. Our model is distinguished primarily by the focus on carbon dissimilation and by the focus on the infant gut microbiota. While useful as frameworks for many applications, we chose to continue developing our own line of gut models based on our previous work [32]. The previous model on a dynamic, evolving microbiota was a good fit to the infant gut microbiota, which is characterised by a succession of species. It also allowed for the easy integration of advection and a reasonably low computational overhead. Nonetheless, further integration of methods would allow for better comparison between models, and our model could potentially be implemented in the new COMETS 2 framework [21].

We have generated testable hypotheses on the causes and mechanics of succession in the infant gut microbiota, potentially laying a base for nutritional interventions that could improve the health of infants. It should be emphasized that the current and future work on this model represents a tool for generating hypotheses, and to test potential mechanical explanations. We cannot fully represent the complexities of the infant gut microbiota, and any generated hypotheses must necessarily be validated in *in vitro* and *in vivo* conditions. Detailed analysis of the infant microbiota has revealed a species composition that is much more diverse in both species and metabolites than our model results show [4, 64]. There is future work to be done to further study the relevant factors in the infant gut, and to bring these results into a more realistic model context. We aim to do so by integrating prebiotics and protein content of nutrition into the model. This may lead to the existence of more niches, and a diversity closer to that of the *in vivo* system. We will also continue to improve the metabolic models, such as by disabling reverse lactate dehydrogenase in *Bifidobacterium*. This may lead to novel future insights on the interactions between different infant nutrition and succession in the infant microbiota.

Methods

We use a spatially explicit model to represent the newborn infant colon (Fig. 1B). Our model is based on an earlier model of the adult gut microbiota [32]. The model consists of a regular square lattice of 225 by 8 lattice sites, where each lattice site represents 2mm by 2mm of space, resulting in a infant colon of 450mm by 16mm. Each lattice site can contain any number of metabolites of the 723 types represented in the model in any concentration and a single simulated bacterial population. The metabolism of these bacterial populations is based on genome-scale metabolic models (GEMs) from the AGORA set of species-specific GEMs [42]. From this set, we have chosen 15 GEMs based on a consortium of known infant microbiota species [4, 16]. Their metabolic inputs and outputs are calculated using Dynamic Flux Balance Analysis [68] with a limit to the total flux through the network [44]. The effects of the FBA solution are applied to the spatial environment and then recalculated each timestep, creating a spatial dynamic FBA (sdFBA).

We identify the two narrow ends of the rectangle with the proximal and distal ends of the colon. In each timestep metabolites both diffuse and flow from the proximal to the distal end. To mimic adhesion to the gut wall and mixing through contractions [3], bacterial populations diffuse, but do not flow distally as metabolites do.

At the start of each run, we initialize the system with a large number of small populations. We let these perform their metabolism each timestep. They take metabolites from the environment, and deposit the resulting products. We let the populations grow and divide according to their energy

output. Both the initial placement locations and movement of the populations is random, introducing stochasticity in the model. The system develops differently depending on initial conditions, into a diverse and complex ecosystem.

Species composition

Each population is represented by a genome-scale metabolic model (GEM) of a species. 15 different metabolic models are used in our spatially explicit model (Table 1), selected based on previous research [16]. From this list of genera, the most prevalent species within that genus in the vaginally born newborn data set from Bäckhed *et al.* was used [4]. One group from the list could only be determined at the family level - Lachnospiraceae. *Ruminococcus gnavus* was chosen to represent this group, based on its high prevalence among this group and the prior inclusion of species from the *Blautia* and *Dorea* genera [55]. Because the genus *Bifidobacterium* is known to be particularly diverse [65], we represented it with models of three different strains. All models are based on the GEMs created by Magnúsdóttir et al. in the AGORA project [42].

Bacterial metabolism

FBA with flux limit

Each timestep of the model, a modified version of Flux Balance Analysis with a total flux limit is used by each population to determine what reactions it should use to achieve the most biomass production from the metabolites available to it [44, 48]. First, each GEM is converted to a stoichiometric matrix S. All reversible reactions are converted to two irreversible reactions, so that flux is always greater than or equal to 0. All reactions identified as exchange, sink, or demand in the metabolic reconstruction are marked as exchange in the matrix. Each timestep, all reactions are then assumed to be in steady state:

$$\frac{d\vec{x}}{dt} = S \cdot \vec{f} = 0 \tag{1}$$

Where $\vec{x}$ is a vector of all metabolites and $\vec{f}$ is a vector of all metabolic fluxes through each reaction in the network. Reactions that take up metabolites from the environment are constrained by an upper bound set dynamically by the environment [68]:

$$\vec{f} \leq \vec{f}_{ub} \tag{2}$$

These represent the limited availability of most nutrients from the environment. Only water uptake is unbound. There is an additional constraint on the total flux [44]:

$$\sum f_i \leq B \cdot a \tag{3}$$

where B is the size of the local bacterial population, and a a scaling factor. This functions as a limit on the total flux through the system.

FBA then identifies the solution space that adheres to these constraints and from this spaces identifies the solution that optimizes the objective function. Our objective function takes ATP and returns ADP, P_i, and a proton. Our FBA is performed using the GNU Linear Programming Kit (GLPK).

The solution consists of a set of exchange fluxes and a growth rate. These exchange fluxes are taken as the derivatives of a set of partial-differential equations to model the transport of intermediary metabolites (see below). The size of the population increases in proportion to the growth rate produced by the solution.

The addition of a flux limit constraint is similar, to previous models [44]. Our method is also similar to FBA with molecular crowding in that we use an additional constraint to model metabolic capacity [7]. However, we do not utilize the necessarily reaction-specific crowding coefficients.

Checking the validity of the GEMs

All GEMs used in the model were tested individually to ensure that they could grow on a substrate of lactose. Only *Veillonella* did not pass this test, which is consistent with experimental observations [53]. *Veillonella* did pass when lactate, a common infant gut metabolite, was used instead. We also tested all GEMs individually for spurious growth in absence of substrates. To this end, all uptake bounds except water were set to 0. None of the GEMs grew under these conditions. All models were tested for having a net neutral exchange of hydrogen, carbon, oxygen and nitrogen at each timestep of the model during the experiments. This process revealed a discrepancy in several reactions, where protons were destroyed. We corrected these reactions (Table S1). With the corrections we applied to the model, these tests always passed - allowing for rounding errors of less than $1 \cdot 10^{-8}$ µmol per FBA solution.

The thermodynamic correctness of all reactions was checked by calculating the net difference in Gibbs free energy between input and output metabolites, using a pre-existing dataset of thermodynamic source data [24]. The conditions assumed for the calculation of the Gibbs free energy were a pH of 7 and an ionic strength of 0.1M [24]. This was recorded for each population at each timestep over the course of a full run of 10080 timesteps. 99.9999% of all FBA solutions in the experiment of Fig. 2 passed the test by containing less free energy than the associated inputs (Fig. 1 E). The remaining 0.0001% all had an amount of energy in the outputs equal to that of the inputs. All of these solutions had very low growth rates, less than 0.001% of the average growth rate.

Changes from AGORA

We use updated versions of the AGORA GEMs [41], to which we have applied checks and modifications. Firstly, the objective function was changed from the biomass reaction included in the models to a reaction only requiring ATP production. As this reaction yields only ADP and P_i, it is is mass-neutral. This allows us to focus on carbon dissimilation within the GEMs, and the differences in underlying metabolism. ATP yield has been a good proxy for biomass production in previous studies [58]. Focusing on carbon dissimilation means we can leave all unknown uptake bounds at 0, instead of using an arbitrary or randomized level, as in some other studies [6, 12].

We also checked the metabolic networks for reactions that allow for the occurrence of unrealistic FBA solutions, and we have added additional reactions. In the *Bifidobacterium* models the ATP synthase reaction was made reversible. This allows it to function as a proton pump, which *Bifidobacterium* uses to maintain its internal pH [56]. Lactose permease reactions were added to all *Bifidobacterium* species in the model [38] based on those available in other models in the AGORA dataset. Combined with the existing reactions in the model and the metabolic constraints this leads to a set of reactions simulating a realistic bifid shunt yield of 5 mol ATP, 2 mol lactate and 3 mol acetate per mol of lactose [61]. Lactose permease was also added to *Streptococcus salivarius* and *Ruminococcus gnavus* to bring them in line with existing literature on their *in vitro* behavior [33, 40]. *Veillonella dispar* is the only species in the model that does not have any lactose uptake [53]. A complete list of changes is presented in supplemental file S1 Table.

Environmental metabolites

We modeled a set of 723 extracellular metabolites - the union of all metabolites that can be exchanged with the environment by at least one GEM in the model. In practice, only a fraction occurs in any reaction that has flux over it, and only 17 metabolites are ever present in the medium in more than micromolar amounts in our experiments outside of the sensitivity analysis.

To mimic advection, the complete metabolic contents of the lattice are moved towards the distal end by one lattice site per timestep, i.e. once every 3 simulated minutes. This leads to an average transit time of approximately 11 hours. Due to diffusion the first metabolites of a food input will emerge from the distal end after around 10 hours, in agreement with the observed cecum to rectum transit time in newborns [34, 54]. Metabolites moving out of the distal end are removed entirely and analysed separately (see Section ‘ Analysis’ for detail).

Every timestep all metabolites diffuse between adjacent lattice sites and are added and removed by bacterial populations, yielding

$$\frac{d\vec{c}(\vec{x},t)}{dt} = \vec{F}(\vec{x},t)B(\vec{x},t) + \frac{D}{L^2}\sum_{\vec{i}\in NB(\vec{x})}\left(\vec{c}(\vec{i},t) - \vec{c}(\vec{x},t)\right). \tag{4}$$

All lattice sites initially contain water as their only metabolite, except for the aerobic conditions where oxygen is added. The upper uptake flux bound f_{ub} of each metabolite exchange reaction is given by the local concentration

$$\vec{f}_{ub} = \vec{c}(\vec{x},t) \tag{5}$$

Metabolites representing the food intake are inserted into the first six columns of lattice sites every three hours (60 timesteps) to approximate a realistic interval for neonates [30]. It consists solely of lactose, in a concentration in line with human milk [5], assuming 98% host uptake of carbohydrates before reaching the colon, a commonly used assumption [12]. Water is provided as a metabolite in unlimited quantities. No other metabolites are available, other than those produced during as a result of bacterial metabolism within the model.

Initial conditions

The simulation is initialized by placing a number of very small populations of the various species randomly across different lattice sites of the environment (Table 2). This is considered to be equivalent to a plausible initial total load of approximately $3 \cdot 10^{10}$ [49], assuming a total colon volume of approximately 88 ml. As there is little information on the relative abundance of species in the very early infant gut, we place all species at equal abundance. In oxygenated conditions, oxygen is also placed as a metabolite now. Water is always considered to be present everywhere. No other metabolites are initially present except for the first feeding.

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

Population dynamics

Each local population solves the FBA problem based on its own possible reactions, a set flux limit, and the local concentrations of metabolites at each timestep, and applies the outcome to the environment and its own size. If an adjacent lattice site is empty, large populations (200 times the initial size, Table 2) split into two. Each new population is then assigned half of the old population, in such a way that the total size is preserved. To mimic colonization events new populations are introduced at random into empty lattice sites during the simulation with a small probability (Table 2). The new population is equal in size to an initial population. Each population dies out at a probability of 0.75%per timestep.

To mimic the spread of the bacterial populations, the lattice sites swap the population contents with adjacent lattice sites each timestep. Random lattice sites are taken and swapped with a random adjacent lattice site using the Moore neighbourhood. This only affects any present population or populations, not the metabolites. No lattice site is swapped more than once per timestep. The lattice site swaps are repeated until no more swaps are possible, i.e. all lattice sites that have not been swapped only have adjacent lattice sites that have been swapped. This procedure is similar to one commonly used for modelling liquid solutions [62]. This method of movement does not allow populations to move out of the gut.

Analysis

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

To detect any irregularities, we also record the net carbon, hydrogen, and oxygen flux of every population and the system as a whole. The difference in Gibbs free energy per timestep is also recorded per population and over the whole system. Estimated Gibbs free energy is derived from the Equillibrator database [47]. Energy loss g per step per population is recorded as follows, where j are metabolites, F is the fluxes and E contains the Gibbs free energy value for each metabolite,

$$g = \sum_j F(j) \cdot E(j) \tag{6}$$

For specific experiments, reactions are removed from the models. This is performed by deleting some reactions from the GEM before the conversion to the stoichiometric matrix. For other experiments, the uptake of certain metabolites is disabled. This is done by placing the upper flux bound $\vec{f}_{ub}$ of the exchange reaction at 0 for the relevant populations.

Parameters

Relevant parameters are listed in Table 2. Based on measurements of the typical length and diameter of the infant colon [13,66] we estimated a volume of 88 ml. Combined with the average abundance per ml of around 10^{10} after the first days [49], this leads to a very rough estimate of 10^{12} bacteria in the young infant colon. To remain computationally feasible, while still modelling at a high resolution, we divide this population into units of at most $1 \cdot 10^{10}$. Local populations at or above this limit will divide into two equally sized populations when space is available. This prevents the local population from becoming unrealistically high. Local populations of more than $2 \cdot 10^{10}$, which can only form if no space is available to divide for a longer time period, cease metabolism.

Table 2. Default parameters of the model

Parameter	Value	Unit
Timestep	180	seconds
Lattice site length	2	millimetre
Width of the lattice	225	lattice sites
Height of the lattice	8	lattice sites
Colon transit time	10	hours
Number of initial populations	540	-
Initial size per population	$5 \cdot 10^7$	number of bacteria
Population size required for division	$1 \cdot 10^{10}$	number of bacteria
Death probability	0.0075	per timestep per population
Flux limit	0.02	µmol flux per step per $1 \cdot 10^{10}$ population
New species placement probability	0.00005	per timestep per empty lattice site
Diffusion	$4.4 \cdot 10^{-5}$	cm^2/s
Growth per µmol ATP	$2.5 \cdot 10^7$	number of bacteria
Lactose input	210	µmol per three hours

The lactose input is estimated from the known intake of milk, its lactose concentration, and an estimate of pre-colonic lactose absorption of 98% [2,12]. Little data is available on the growth rate of bacteria within the human colon. Growth rates are expected to be much lower than those found in *in vitro* cultures of individual species [28]. In absence of precise data for infants, here we use a value based on the estimated doubling time of the whole gut microbiota in mice, 5.8 h [27]. The colonic transit

time is based on data for total transit time gathered with carmine red dye [54], adjusted for the mouth to cecum transit time [34]. The timestep interval was set at three simulated minutes, to be able to capture individual feedings at a high resolution. Other parameters were determined experimentally to reach plausible outcomes on a metabolic and species level while maintaining computational feasibility.

Implementation

The model was implemented in C++11. This code was based on one used earlier to model the gut microbiota [32]. The GEMs are loaded using libSBML 5.18.0 for C++. We use the GNU Linear Programming Kit 4.65 (GLPK) as a linear programming tool to solve each FBA with flux limit problem. We use the May 2019 update of AGORA, the latest at time of writing, from the Virtual Metabolic Human Project website (vhm.life). Thermodynamic data is extracted from the eQuilibrator API (December 2018 update) [47] using Python 3.6. Model screenshots are made using the libpng16 and pngwriter libraries. Other visualizations and statistical analysis were performed with R 3.4.4.

Supporting information

S1 Table. Table of changed or deleted reactions.csv
A table of changes made to the AGORA models as a .csv file.

S2 Video. Video of the first 5000 steps of a run without initially oxygenated conditions

S3 Video. Video of the first 1300 steps of a run under initially oxygenated conditions

S4 Video. Video of step 9000 to step 10000 of a run under initially oxygenated conditions

Data availability

The code and data that support the findings of this study are available from the corresponding author on reasonable request

References

[1] L. Albenberg, T. V. Esipova, C. P. Judge, K. Bittinger, J. Chen, A. Laughlin, S. Grunberg, R. N. Baldassano, J. D. Lewis, H. Li, S. R. Thom, F. D. Bushman, S. A. Vinogradov, and G. D. Wu. Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology*, 147(5):1055–1063.e8, 2014.

[2] P. O. Anderson and V. Valdés. Variation of milk intake over time: Clinical and pharmacokinetic implications. *Breastfeeding Medicine*, 10(3):142–144, 2015.

[3] M. Arnoldini, J. Cremer, and T. Hwa. Bacterial growth, flow, and mixing shape human gut microbiota density and composition. *Gut microbes*, 9(6):559–566, 2018.

[4] F. Bäckhed, J. Roswall, Y. Peng, Q. Feng, H. Jia, P. Kovatcheva-Datchary, Y. Li, Y. Xia, H. Xie, H. Zhong, M. T. Khan, J. Zhang, J. Li, L. Xiao, J. Al-Aama, D. Zhang, Y. S. Lee, D. Kotowska, C. Colding, V. Tremaroli, Y. Yin, S. Bergman, X. Xu, L. Madsen, K. Kristiansen, J. Dahlgren, and W. Jun. Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host and Microbe*, 17(5):690–703, 2015.

[5] O. Ballard and A. L. Morrow. Human Milk Composition. Nutrients and Bioactive Factors. *Pediatric Clinics of North America*, 60(1):49–74, 2013.

[6] E. Bauer, J. Zimmermann, F. Baldini, I. Thiele, and C. Kaleta. BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLoS Computational Biology*, 13(5):1–22, 2017.

[7] Q. K. Beg, A. Vazquez, J. Ernst, M. A. De Menezes, Z. Bar-Joseph, A. L. Barabási, and Z. N. Oltvai. Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proceedings of the National Academy of Sciences of the United States of America*, 104(31):12663–12668, 2007.

[8] B. Berger, N. Porta, F. Foata, D. Grathwohl, M. Delley, D. Moine, A. Charpagne, L. Siegwald, P. Descombes, P. Alliet, G. Puccio, P. Steenhout, A. Mercenier, and N. Sprenger. Linking human milk oligosaccharides, infant fecal community types, and later risk to require antibiotics. *mBio*, 11(2):1–18, 2020.

[9] S. L. Bridgman, M. B. Azad, C. J. Field, A. M. Haqq, A. B. Becker, P. J. Mandhane, P. Subbarao, S. E. Turvey, M. R. Sears, J. A. Scott, D. S. Wishart, and A. L. Kozyrskyj. Fecal Short-Chain Fatty Acid Variations by Breastfeeding Status in Infants at 4 Months: Differences in Relative versus Absolute Concentrations. *Frontiers in Nutrition*, 4(April), 2017.

[10] V. Bucci and J. B. Xavier. Towards predictive models of the human gut microbiome. *Journal of Molecular Biology*, 426(23):3907–3916, 2014.

[11] Chan, Friedman, Wu, and Maranas. Predicting the Longitudinally and Radially Varying Microbial Metabolic Modeling. *processes*, 2019.

[12] S. H. J. Chan, M. N. Simons, and C. D. Maranas. SteadyCom: Predicting microbial abundances while ensuring community stability. *PLoS Computational Biology*, 13(5):1–25, 2017.

[13] W. S. Davis, R. Parker Allen, and B. E. Favara. Neonatal small left colon syndrome. *American journal of Roentgenology*, 120(2):322–329, 1974.

[14] W. de Vries and A. H. Stouthamer. Pathway of glucose fermentation in relation to the taxonomy of bifidobacteria. *Journal of bacteriology*, 93(2):574–576, 1967.

[15] L. De Vuyst, F. Moens, M. Selak, A. Rivière, and F. Leroy. Summer Meeting 2013: Growth and physiology of bifidobacteria. *Journal of Applied Microbiology*, 116(3):477–491, 2014.

[16] D. Delsing, C. Lindner, H. van Dijck, E. Looijesteijn, E. S. Lima, M. Heerikhuizen, V. Agamennone, V. Triantis, and I. Bovee-Oudenhoven. Gut microbiota modulating effects of GOS and 2'-FL studied by in vitro infant and toddler faecal fermentations. *Journal of Pediatric Gastroenterology and Nutrition*, 68(Supplement 1):1037, 2019.

[17] R. Diaz-Ruiz, M. Rigoulet, and A. Devin. The Warburg and Crabtree effects: On the origin of cancer cell energy metabolism and of yeast glucose repression. *Biochimica et Biophysica Acta - Bioenergetics*, 1807(6):568–576, 2011.

[18] S. Dogra, O. Sakwinska, S. E. Soh, C. Ngom-Bru, W. M. Brück, B. Berger, H. Brüssow, N. Karnani, Y. S. Lee, F. Yap, Y. S. Chong, K. M. Godfrey, and J. D. Holbrook. Rate of establishing the gut microbiota in infancy has consequences for future health. *Gut Microbes*, 6(5):321–325, 2015.

[19] S. Dogra, O. Sakwinska, S.-E. E. Soh, C. Ngom-Bru, W. M. Brück, B. Berger, H. Brüssow, Y. S. Lee, F. Yap, Y.-S. S. Chong, K. M. Godfrey, J. D. Holbrook, GUSTO Study Group, O. Sakwinska, C. Ngom-Bru, W. M. Brück, B. Berger, H. Brüssow, S.-E. E. Soh, F. Yap, K. M. Godfrey, and J. D. Holbrook. Dynamics of infant gut microbiota are influenced by delivery mode and gestational duration and are associated with subsequent adiposity. *mBio*, 6(1):1–9, 2015.

[20] R. M. Duar, D. Kyle, and G. Casaburi. Colonization Resistance in the Infant Gut: The Role of B. infantis in Reducing pH and Preventing Pathogen Growth. *High-Throughput*, 9(2):7, 2020.

[21] I. Dukovski, D. Bajić, J. M. Chacón, M. Quintin, J. C. Vila, S. Sulheim, A. R. Pacheco, D. B. Bernstein, W. J. Rieh, K. S. Korolev, A. Sanchez, W. R. Harcombe, and D. Segrè. Computation Of Microbial Ecosystems in Time and Space (COMETS): An open source collaborative platform for modeling ecosystems metabolism. 2020.

[22] S. H. Duncan, P. Louis, and H. J. Flint. Lactate-Utilizing Bacteria, Isolated from Human Feces, That Produce Butyrate as a Major Fermentation Product. *Applied and Environmental Microbiology*, 70(10):5810–5817, 2004.

[23] K. Faust and J. Raes. Microbial interactions: From networks to models. *Nature Reviews Microbiology*, 10(8):538–550, 2012.

[24] A. Flamholz, E. Noor, A. Bar-Even, and R. Milo. EQuilibrator - The biochemical thermodynamics calculator. *Nucleic Acids Research*, 40(D1):770–775, 2012.

[25] R. Francavilla, M. Calasso, L. Calace, S. Siragusa, M. Ndagijimana, P. Vernocchi, L. Brunetti, G. Mancino, G. Tedeschi, E. Guerzoni, F. Indrio, L. Laghi, V. L. Miniello, M. Gobbetti, and M. De Angelis. Effect of lactose on gut microbiota and metabolome of infants with cow's milk allergy. *Pediatric Allergy and Immunology*, 23(5):420–427, 2012.

[26] E. S. Friedman, K. Bittinger, T. V. Esipova, L. Hou, L. Chau, J. Jiang, C. Mesaros, P. J. Lund, X. Liang, G. A. FitzGerald, M. Goulian, D. Lee, B. A. Garcia, I. A. Blair, S. A. Vinogradov, and G. D. Wu. Microbes vs. chemistry in the origin of the anaerobic gut lumen. *Proceedings of the National Academy of Sciences of the United States of America*, 115(16):4170–4175, 2018.

[27] R. J. Gibbons and B. Kapsimalis. Estimates of the overall rate of growth of the intestinal microflora of hamsters, guinea pigs, and mice. *Journal of Bacteriology*, 93(1):510–512, 1967.

[28] B. Gibson, D. J. Wilson, E. Feil, and A. Eyre-Walker. The distribution of bacterial doubling times in the wild. *Proceedings of the Royal Society B: Biological Sciences*, 285(1880), 2018.

[29] M. Gralka, R. Szabo, R. Stocker, and O. X. Cordero. Trophic Interactions and the Drivers of Microbial Community Assembly. *Current Biology*, 30(19):R1176–R1188, 2020.

[30] L. M. Grummer-Strawn, K. S. Scanlon, and S. B. Fein. Infant feeding and feeding transitions during the first year of life. *Pediatrics*, 122(SUPPL. 2), 2008.

[31] W. R. Harcombe, W. J. Riehl, I. Dukovski, B. R. Granger, A. Betts, A. H. Lang, G. Bonilla, A. Kar, N. Leiby, P. Mehta, C. J. Marx, and D. Segrè. Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Reports*, 7(4):1104–1115, 2014.

[32] M. J. A. Hoek and R. M. H. Merks. Emergence of microbial diversity due to cross-feeding interactions in a spatial model of gut microbial metabolism. *BMC Systems Biology*, 11(1):1–18, 2017.

[33] L. V. Holdeman. Emendation of Bacteroidaceae and B utyrivibrio and Descriptions of Desulfornonas gen. nov. and Ten New Species in the Genera Desulfomonas, Butyrivibrio, Eubacterium, Clostridium, and Ruminococcus. *International Journal of Systematic Bacteriology*, 26(2):238–252, 1976.

[34] J. S. Hyams, M. A. Geertsma, N. L. Etienne, and W. R. Treem. Colonic hydrogen production in infants with colic. *The Journal of Pediatrics*, 115(4):592–594, 1989.

[35] J. Idänpään-Heikkilä, P. Jouppila, H. K. Åkerblom, R. Isoaho, E. Kauppila, and M. Koivisto. Elimination and metabolic effects of ethanol in mother, fetus, and newborn infant. *American Journal of Obstetrics and Gynecology*, 112(3):387–393, 1972.

[36] M. Kalliomaki, M. Collado, S. S, and E. Isolauri. Early differences in fecal microbiota composition in children may. *Am J Clin Nutr*, 87(1):534–538, 2008.

[37] K. M. Kennedy, M. J. Gerlach, T. Adam, M. M. Heimesaat, L. Rossi, M. G. Surette, D. M. Sloboda, and T. Braun. Fetal meconium does not have a detectable microbiota before birth. *Nature Microbiology*, 6(7):865–873, 2021.

[38] F. Krzewinski, C. Brassart, F. Gavini, and S. Bouquelet. Characterization of the lactose transport system in the strain Bifidobacterium bifidum DSM 20082. *Current Microbiology*, 32(6):301–307, 1996.

[39] C. Lay, C. W. Chu, R. W. Purbojati, E. Acerbi, D. I. Drautz-Moses, P. F. de Sessions, S. Jie, E. Ho, Y. J. Kok, X. Bi, S. Chen, S. Y. Mak, M. C. Chua, A. E. Goh, W. C. Chiang, R. Rao, S. Chaithongwongwatthana, N. Khemapech, V. Chongsrisawat, R. Martin, Y. Koh, S. R. Lohar, I. C. H. Tan, W. A. Anng, C. Jie, N. Bartke, K. Ben-Amor, I. B. Renes, F. Wong, G. Roeselers, Y. S. Ho, M. L. Hibberd, S. C. Schuster, and J. Knol. A synbiotic intervention modulates meta-omics signatures of gut redox potential and acidity in elective caesarean born infants. *BMC Microbiology*, 21(1):1–17, 2021.

[40] C. Lessard, A. Cochu, J. D. Lemay, D. Roy, K. Vaillancourt, M. Frenette, S. Moineau, and C. Vadeboncoeur. Phosphorylation of Streptococcus salivarius Lactose Permease (LacS) by HPr(His∼P) and HPr(Ser-P)(His∼P) and Effects on Growth. *Journal of Bacteriology*, 185(23):6764–6772, 2003.

[41] S. Magnúsdóttir, A. Heinken, R. M. T. Fleming, and I. Thiele. Reply to "Challenges in modeling the human gut microbiome". *Nature Biotechnology*, 36(8):686–691, 2018.

[42] S. Magnúsdóttir, A. Heinken, L. Kutt, D. A. Ravcheev, E. Bauer, A. Noronha, K. Greenhalgh, C. Jäger, J. Baginska, P. Wilmes, R. M. Fleming, and I. Thiele. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nature Biotechnology*, 35(1):81–89, 2017.

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

[44] R. A. Majewski and M. M. Domach. Simple constrained-optimization view of acetate overflow in E. coli. *Biotechnology and Bioengineering*, 35(7):732–738, 1990.

[45] F. P. J. Martin, S. Moco, I. Montoliu, S. Collino, L. Da Silva, S. Rezzi, R. Prieto, M. Kussmann, J. Inostroza, and P. Steenhout. Impact of breast-feeding and high-and low-protein formula on the metabolism and growth of infants from overweight and obese mothers. *Pediatric Research*, 75(4):535–543, 2014.

[46] S. Matamoros, C. Gras-Leguen, F. Le Vacon, G. Potel, and M. F. De La Cochetiere. Development of intestinal microbiota in infants and its impact on health. *Trends in Microbiology*, 21(4):167–173, 2013.

[47] E. Noor, H. S. Haraldsdóttir, R. Milo, and R. M. Fleming. Consistent Estimation of Gibbs Energy Using Component Contributions. *PLoS Computational Biology*, 9(7), 2013.

[48] J. D. Orth, I. Thiele, and B. O. Palsson. What is flux balance analysis? *Nature Biotechnology*, 28(3):245–248, 2010.

[49] C. Palmer, E. M. Bik, D. B. DiGiulio, D. A. Relman, and P. O. Brown. Development of the human infant intestinal microbiota. *PLoS Biology*, 5(7):1556–1573, 2007.

[50] V. T. Pham, C. Lacroix, C. P. Braegger, and C. Chassard. Early colonization of functional groups of microbes in the infant gut. *Environmental microbiology*, 18(7):2246–2258, 2016.

[51] V. T. Pham, C. Lacroix, C. P. Braegger, and C. Chassard. Lactate-utilizing community is associated with gut microbiota dysbiosis in colicky infants. *Scientific Reports*, 7(1):1–13, 2017.

[52] S. Quesada-Vázquez, G. Aragonès, J. M. Del Bas, and X. Escoté. Diet, Gut Microbiota and Non-Alcoholic Fatty Liver Disease: Three Parts of the Same Axis. *Cells*, 9(1):176, 2020.

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

[54] F. F. Rubaltelli and G. Largajolli. Effect of Light Exposure on Gut Transit Time in Jaundiced Newborns. *Acta Pædiatrica*, 62(2):146–148, 1973.

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

[56] B. Sánchez, C. G. De Los Reyes-Gavilán, and A. Margolles. The F1F0-ATPase of Bifidobacterium animalis is involved in bile tolerance. *Environmental Microbiology*, 8(10):1825–1833, 2006.

[57] P. A. Scholtens, R. Oozeer, R. Martin, K. B. Amor, and J. Knol. The Early Settlers: Intestinal Microbiology in Early Life. *Annual Review of Food Science and Technology*, 3(1):425–447, 2012.

[58] R. Schuetz, L. Kuepfer, and U. Sauer. Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Molecular Systems Biology*, 3(119), 2007.

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

[60] I. Sekirov, S. L. Russell, L. Caetano M Antunes, and B. B. Finlay. Gut microbiota in health and disease. *Physiological Reviews*, 90(3):859–904, 2010.

[61] D. A. Sela, A. Adeuya, J. H. Kim, T. R. Whitehead, J. Chapman, D. S. Rokhsar, P. M. Richardson, A. Lapidus, J. B. German, N. P. Price, F. Chen, C. B. Lebrilla, and D. A. Mills. The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome. *Proceedings of the National Academy of Sciences*, 105(48):18964–18969, 2008.

[62] C. S. Shida and V. B. Henriques. Smart and wrong versus correct Monte Carlo simulations of Kawasaki-Ising model. *International Journal of Modern Physics C*, 11(5):1033–1036, 2000.

[63] K. Sim, A. G. Shaw, P. Randell, M. J. Cox, Z. E. McClure, M. S. Li, M. Haddad, P. R. Langford, W. O. Cookson, M. F. Moffatt, and J. S. Kroll. Dysbiosis anticipating necrotizing enterocolitis in very premature infants. *Clinical Infectious Diseases*, 60(3):389–397, 2015.

[64] N. Tsukuda, K. Yahagi, T. Hara, Y. Watanabe, H. Matsumoto, H. Mori, K. Higashi, H. Tsuji, S. Matsumoto, K. Kurokawa, and T. Matsuki. Key bacterial taxa and metabolic pathways affecting gut short-chain fatty acid profiles in early life. *ISME Journal*, 2021.

[65] F. Turroni, C. Peano, D. A. Pass, E. Foroni, M. Severgnini, M. J. Claesson, C. Kerr, J. Hourihane, D. Murray, F. Fuligni, M. Gueimonde, A. Margolles, G. de Bellis, P. W. O'Toole, D. van Sinderen, J. R. Marchesi, and M. Ventura. Diversity of bifidobacteria within the infant gut microbiota. *PLoS ONE*, 7(5):20–24, 2012.

[66] J. Valentin and C. Streffer. Basic anatomical and physiological data for use in radiological protection: Reference values - ICRP Publication 89. *Annals of the ICRP*, 32(3-4):1–277, 2002.

[67] N. van Best, M. W. Hornef, P. H. Savelkoul, and J. Penders. On the origin of species: Factors shaping the establishment of infant's gut microbiota. *Birth Defects Research Part C - Embryo Today: Reviews*, 105(4):240–251, 2015.

[68] A. Varma and B. Ø. Palsson. Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use. *Biotechnology*, 12(October):994–998, 1994.

[69] L. Zheng, C. J. Kelly, and S. P. Colgan. Physiologic hypoxia and oxygen homeostasis in the healthy intestine. A review in the theme: Cellular responses to hypoxia. *American Journal of Physiology - Cell Physiology*, 309(6):C350–C360, 2015.

Acknowledgments

The work was carried out in part on the Dutch national e-infrastructure with the support of SURF Cooperative. This work was performed in part using the ALICE compute resources provided by Leiden University.

Contributions

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

Competing interests

This study was financially supported by FrieslandCampina. R.S, E.L., P.V.J., M.P. and J.M.W.G. are currently or were previously employed by FrieslandCampina.