Understanding causality and treatment of bacterial dysbiosis in the human gut using systems biology approaches

Jack Jansma. S2172003

ABSTRACT

Bacteria are living in symbiosis in the human gut. Here bacteria can have positive effects on the host immune system, central nervous system and digestive tract. A dysbiosis in the human gut may result in serious diseases, such as inflammatory bowel disease and major depressive disorder. However the onset of the disease and the mechanism of action of possible treatments is not well understood. Systems biology can discover emergent properties of a network via mathematical modeling, thus systems biology can help unraveling both problems. Systems biology approaches can be divided into top-down and bottom-up approaches. The top-down approach uses –omics approaches to construct a network based on quantitative, experimental data. Keystone species can be identified with this approach, which leads to a better understanding of the microbiota structure. The bottom-up approach uses mathematical models to look for novel interactions in a network. The bottom-up approach mainly uses different types of flux balance analysis. Via the addition of a probiotic or prebiotic to a model the mechanism of action can be determined. This study gives an overview of different systems biology approaches in microbial research and shows the possibilities of these approaches.

The gut microbiota

There are approximately 10^{13} bacterial cells in our body, which is roughly the same number of cells in a human body, most of these bacteria are found in the colon[1]. In the gut, bacteria play a vital role in several physiological processes of the host. Members of the genera *Bacteroides*, *Roseburia*, *Bifidobacterium* and *Faecalibacterium* are able to produce short-chain fatty acids (SCFA), such as acetate, butyrate and propionate, from non-digestible carbohydrates[2]. These SCFA are taken up by the gut and are used as an energy source, regulators of gene expression and signaling molecules by the host[3]. In addition, members of the microbiota can affect both the innate and adaptive immune response. For example, IgA expressing cells found in Peyer's patches are greatly reduced in germ-free (GF) mice. In turn microbial colonization of GF-mice activates the production of IgA. Furthermore, T_{reg} cells are also induced when the microbiota is present[4]. A healthy microbiota also prevents the onset of infectious diseases via direct and indirect colonization resistance. Direct colonization resistance is achieved by commensal bacteria that can utilize a carbon source faster than a pathogenic bacteria or by secreting an inhibitory factor that targets a pathogenic bacteria. Indirect colonization resistance is achieved by inducing the immune system of the host in such a way that it targets the

pathogenic bacteria[5]. For example, colonization of GF-mice with the symbiotic bacteria *Bacteroides thetaiotaomicron* enhances the production of antimicrobial lectins REGIIIγ and REGIIIβ[6]. Finally, the microbiota plays a role in brain development and brain function. SCFA produced by bacteria account for a less permeable blood-brain barrier, both during fetal development and in adults. Additionally, the microbiota is able to alter neurotransmitter levels in the host, such as serotonin, dopamine and GABA[7,8].

In a healthy gut situation bacterial species are balanced and in symbiosis with the host. However, when the balance is disrupted a dysbiosis can occur. A dysbiosis in the gut may cause several diseases, since the gut microbiota plays a key role in the previous mentioned processes in the host. Fecal samples from patients with inflammatory bowel disease (IBD) show a reduction in bacteria from the phylum Firmicutes[9] and an increase in bacteria from the phylum Proteobacteria[10]. Fecal samples from patients with type 2 diabetes mellitus show a decrease in butyrate-producing bacteria such as, *Faecalibacterium prausnitzii* and *Roseburia intestinalis*[11, 12]. Fecal samples from patients with Parkinson's disease show a decrease in bacteria from the genera Blautia, Coprococcus, and Roseburia. Additionally, bacteria from the genus Ralstonia were more abundant in patients with Parkinson's disease[13]. Fecal samples from patients with major depressive disorder (MDD) showed a reduction in Bacteroidetes and an increase in Actinobacteria[14].

Therapeutic strategies currently used to treat microbiota related diseases include pro- and prebiotics. Probiotics are living microorganisms that when consumed have a positive effect on human health. Prebiotics are dietary supplements that when consumed stimulate the growth and/or activity of beneficial bacteria within the human gut, thus having beneficial consequences for the host. Both pro- and prebiotics help push the gut microbiota out of the dysbiosis and into a healthy, balanced situation[15]. Studies show that administering probiotics can have a positive effect on health[16-18]. Additionally, administering a prebiotic such as fructo-oligosaccharides to patients with Crohn's disease decreased the symptoms[19, 20]. However, the effects of both pro- and prebiotics is still debated and both can have negative effects[21-23]. Therefore a better understanding of the working mechanisms and interactions with the gut microbiota is needed to develop new and better pro- and prebiotics.

The previously mentioned examples of a dysbiosis within the gut gives an indication that the microbiota might play a key role in diseases. However, these studies do not show a causal correlation between the disease state and dysbiosis. Is the dysbiosis a result of the onset of the disease or is the dysbiosis the cause of the disease? The investigation of this causal relationship is important for the development of novel therapeutic strategies. Currently, strategies to unravel the correlation between host health and the microbiota rely mostly on animal models. Studies using this approach colonized GF-mice with the microbiota from patients with MDD and monitored the physical and behavioral changes over time[14, 24]. However, GF models pose several limitations when

investigating causality within humans. The immune responses differ between mice and humans, lab-mice do not have the genetic and environmental diversity of humans and mouse experiments do not take the variability of the human diet and factors such as smoking into account[25]. Furthermore, studies in humans show not always the same pattern. For example, a study done in patients with IBD showed no significant differences in diversity between patients and healthy controls and another study showed an increase in *F. prausnitzii*[26, 27]. These studies are in contrast with the studies mentioned in the previous paragraph[9, 10, 28]. This might be due to the different experimental methods or due to the variability in humans.

To provide novel insights into causality in the relationship between the microbiota and disease and understand the mechanism of actions of pro- and prebiotics, a new approach can be useful.

Systems biology

Wolkenhauer defines system biology as follows: "Systems biology is the science that studies how biological function emerges from the interactions between the components of living systems and how these emergent properties enable and constrain the behavior of those components."[29] In systems biology mathematical tools are utilized to construct a model of biological components. With this constructed model, emergent properties can be studied[30]. An example of such an emergent property based on a mathematical model is that microbial diversity may arise due to cross-feeding, but only at a certain pace[31], which means that with a higher gut motility, the microbial diversity is decreasing. This was confirmed in patients with diarrhea associated irritable bowel syndrome (D-IBS)[32]. This example shows the potential for the use of systems biology in gut microbiota related research.

To discover emergent properties, first a network needs to be constructed. A network typically consist of nodes connected by edges. Nodes represent components of a system and edges are interactions between nodes. In the case of the gut-microbiota, a node can be a bacteria, metabolite or intestinal cell, an edge can be a conversion, uptake or secretion reaction[33, 34]. Figure 1 depicts a simplified example of such a network[35]. To get a system level understanding of a biological system the identity of all the components, the dynamic behavior and the interactions of the components need to be known[36]. In the example depicted in figure 1 only the uptake, conversion and secretion are presented. However, when modeling this system to look for emergent properties, more information needs to be added. In the model the glucose distribution i.e. how much of the glucose available is taken up by *Bifidobacterium adolescentis* and how much by *F. prausnitzii*, uptake rates, conversion rates, secretion rates, cell growth, bacterial ratio etc. need to be included in the system. Additionally, this system only looks at 3 metabolites and 2 bacteria. In the gut there are many more metabolites and bacteria present. Thus to construct a sufficient model a lot more information is needed. This information can be

gathered from experimental data, but can also be gathered from genomics, transcriptomics, proteomics and metabolomics databases. The type of data included in the model depends on the research question and the approach.

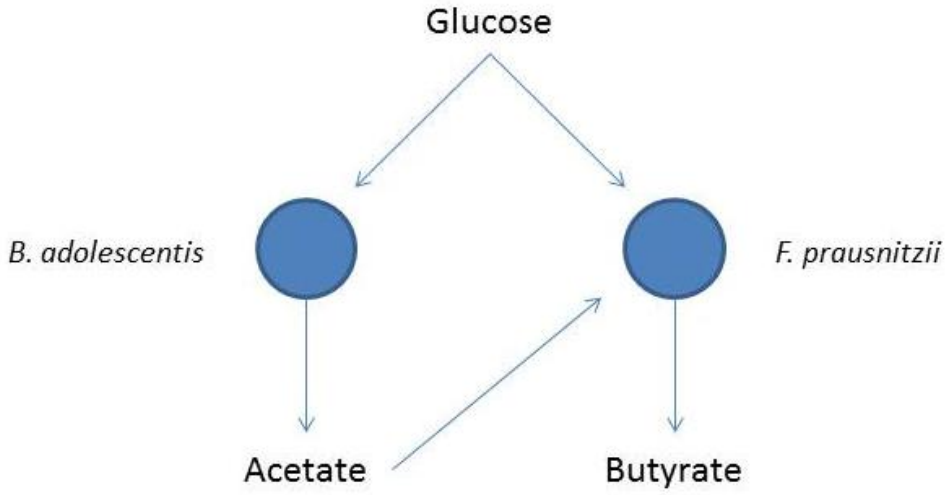

Figure 1: Simplified example of a microbial network. The blue circles are bacteria. The nodes in this example are both bacteria and the metabolites. The edges in this example represent uptake and secretion of metabolites[35].

Systems biology uses mainly two approaches: the top-down approach and the bottom-up approach. Each approach has its own applications and limitations. In the next paragraphs both approaches are discussed and examples of their use in microbiota research are given.

Top-down approach

The top-down approach is a data driven approach. This approach starts with gathering of high-throughput data from, for example RNA-seq data from stool samples. Next, the genes or bacterial species are annotated using a database and the abundances of the genes or bacteria in the sample are determined. Thereafter, a mathematical correlation tool is used to link the abundance of a gene or bacteria to an interaction between bacteria[37].

This approach is widely used to correlate co-occurrence of bacteria within the gut. With co-occurrence, interaction networks can be made. These interaction networks give information on how bacteria affect other bacteria within an environment. These interactions can be mutualistic, commensal, parasitic or ammensal[38]. Simple interaction networks can be made via similarity- and regression based network inference. Making such a network starts with the abundance information. A score is given for each bacteria pair with a similarity or distance measure such as, Pearson correlation, Euclidean distance measure or Spearman correlation. Next, the scoring procedure is repeated a number of times to generate a random score distribution. Bacterial pairs with a *P*-value below a certain

threshold are visualized in a network. However, only positive and negative relationships can be determined. This means that it is not clear if a bacterial pair with a positive correlation has a mutualistic or commensal relationship[39]. Nonetheless, these interaction networks are useful, because network properties can give a lot of information about the bacterial community. For example, keystone species in bacterial communities can be identified. The keystone species have a lot of interactions with other species, whereas non keystone species have a low number of interactions[40]. Additionally niches and separate communities can be visualized. Figure 2 shows an example of an interaction network where keystone species, niches and separate communities are highlighted.

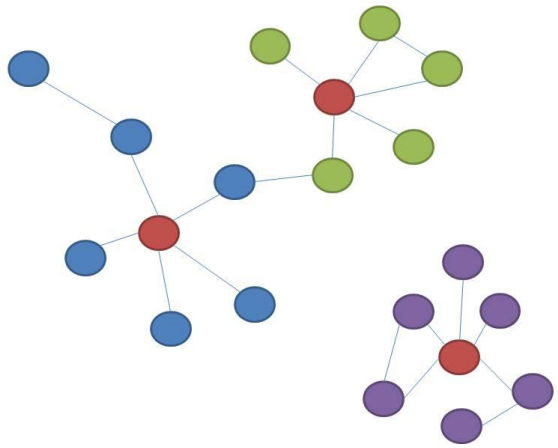

Figure 2: Example of an interaction network. Bacterial species are represented by a colored circles, an interaction is represented by an edge. Keystone species are depicted in red. Within a bacterial community, separate niches can be present. Both niches in this example are depicted with blue and green nodes. The purple nodes represent a separate community.

Fisher et al. used a regression based network to identify *Bacteroides fragilis* and *Bacteroides stercosis* as possible keystone species in the human gut[41]. *B. fragilis* is a bacteria that is known to cause colon cancer[42]. Thus *B. fragilis* might shape the microbial community in a way that it promotes colon cancer in the host. In other words: *B. fragilis* might be a keystone pathogen[43]. By identifying the keystone species in both health and disease, dietary or therapeutically interventions can be designed targeting the keystone species. In this way growth of keystone species associated with health can be stimulated and keystone species associated with disease may be eradicated[44]. However, a study showed that *B. fragilis* improved intestinal health in mice and possibly in humans[45]. Thus showing that both overgrowth and complete eradication of a single bacteria have consequences. Highlighting the importance of a balanced microbiota and showing that designing interventions is not straightforward.

When interpreting an interaction network based on abundances caution is needed. In the ideal situation each sample is taken and prepared in the exact same manner. However, this is never completely true and differences in abundances can arise due to different sample preparations, which will influence the network. Next to this a different network will be constructed when using different

correlation tools[46]. Furthermore, constructing a network in this way gives no information about dynamics, but only gives the interactions between the bacteria at one point in time.

Bottom-up approach

The bottom up approach is a hypothesis driven approach. This approach uses biochemical and microbial knowledge to construct a network of mathematical equations that can be simulated under different physiological conditions. Information about conversion, uptake and secretion reactions is needed to construct such a network. This information can be extracted from databases such as, Kyoto encyclopedia of genes and genomes (KEGG) and Metacyc. These reactions need to be converted into a mathematical language to perform simulations. This can be done using software such as, constrained-based reconstruction and analysis (COBRA), which uses flux balance analysis (FBA) to investigate metabolic fluxes in an organism or between a set of organisms[37, 47].

A flux is the rate of turnover of a metabolite through a metabolic pathway. FBA calculates the distribution of metabolic fluxes in a metabolic network which optimizes a given objective function (e.g. growth rate). Additionally, FBA uses constraints to calculate this optimal distribution of fluxes. To perform FBA all the metabolic reactions in the network need to be represented by a set of linear equations. FBA works at steady-state i.e. the amount of metabolite produced is equal to the amount of metabolite consumed. The set of linear equations can be formulated as:

$$S * v = 0 \qquad \text{(Eq. 1)}$$

S represents a stoichiometric matrix as seen in figure 3 and v represents the flux distribution.

Figure 3: Stoichiometric matrix of the metabolic reactions depicted in the metabolic network of figure 1.

Upper- and lower bounds can be imposed to limit the maximum and minimum values that each flux can take. Additionally, an objective function needs to be formulated. This, the biomass function, contains all the metabolites in the system that are needed to make a new cell. Since Eq. 1 depicts a set of linear equations and there are typically more reactions than compounds, there is more than a single flux distribution possible. By optimizing the objective function in order to achieve a

certain goal, such as maximal growth or maximal production of a certain metabolite, an optimal flux distribution can thus be calculated[48].

FBA is an easy tool to use for many purposes. By adjusting the upper- and lower bounds of metabolites growth on different media can be simulated. By setting the flux of a certain metabolite to zero, a gene knock-out can be simulated. With these examples an indication of the viability of an organism under different conditions can be estimated. Furthermore, by deleting a gene or altering the media composition in silico, the production of a desirable product can be maximized[47, 48]. To predict the behavior of a cell using a FBA model, the model should have a high quality i.e. the model that has the highest number of reactions and constraints based on experimental data, has the highest quality. The quality of a FBA model is mainly dependent on the biomass composition, the protein interaction rules, the objective function, the exchange rates and the available nutrients[49]. Figure 4 shows an example of a model used for FBA.

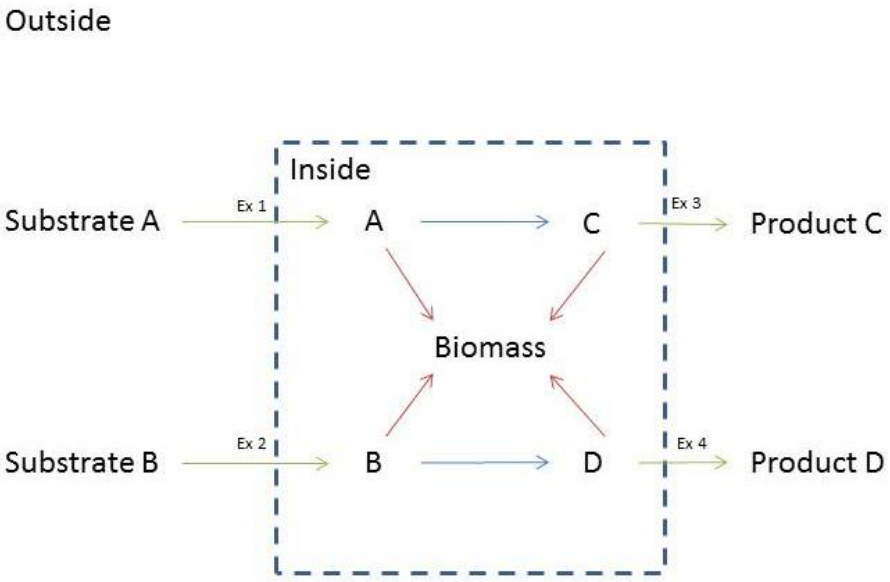

Figure 4: A simple model used for FBA. The cell is depicted with the dashed line. The environment is everything outside the dashed line. Substrates A and B are taken up by the cell and converted into product C and D respectively, represented by the blue arrows. Product C and D are secreted by the cell. The exchange reactions Ex1, Ex2, Ex3 and Ex4 are depicted as green arrows. All the metabolites present in this figure are part of the biomass reaction depicted by the red arrows.

At first glance it seems straightforward to apply classical FBA to a bacterial community. Simply think of the enzymes performing the conversion reactions shown in figure 3 as bacteria and nothing will change. However, since classical FBA depicts metabolic fluxes and takes place within a single cell, a biomass function is needed to account for growth. When modeling a community of different organisms a single biomass function is not sufficient anymore, because all the members of the community have different growth rates and biomass compositions. Furthermore, defining a suitable objective function can be difficult[49, 50]. FBA relies heavily on experimental data to construct a working model. Correct exchange rates determine the fluxes largely. When modeling a metabolic

network of a single cell, exchange rates are usually determined using ^{13}C-labeling experiments[51]. However, determining the contribution of each member of the community to the sum of the fluxes of the extracellular metabolites is not possible with ^{13}C-labeling experiments. To overcome the problems when applying classic FBA to bacterial communities, such as the human gut, new models are being developed all with their own advantages and disadvantages.

Modeling a bacterial community using FBA

Constructing a mathematical community model starts with gathering the metabolic properties of each bacteria in the model. Genomic-scale metabolic reconstructions (GENREs) also called genomic-scale metabolic models (GEMs) represent the complete set of metabolic properties of an organism based on genomic and experimental data[52]. GENREs can be automatically generated from the annotated genome sequence. The set of metabolic enzymes gives all the metabolic properties of the organism. These GENREs can be used to make a community model and perform simulations[53]. However, not all metabolic properties of an organism will be active in each environment. Therefore the GENRE need to be further refined. This has to be done manually and is therefore labor intensive. Currently, a number of algorithms are developed to speed up this process. However, the evaluation of the GENRE has still to be done manually[52]. Thiele et al. published a step-by-step overview on how to generate a high quality GENRE[54].

Compartmentalized community models put two GENREs of bacteria together in the same environment, which can represent growth media or the lumen of the gut, as shown in figure 5. In this model, two organisms are separated from each other by being in different compartments. Both compartments are combined in a larger compartment, which represents the environment. In this model different organisms can compete for a substrate as shown with metabolite B in figure 5 or may benefit from each other as shown with metabolite C in figure 5. Thus competitive and commensal relationships can be integrated in this model[47, 49, 55]. With a compartmentalized community model consisting of *B. adolescentis* and *F. prausnitzii* El-semman and coworkers showed that the growth of *F. prausnitzii* and butyrate production is limited with a low acetate supply[35]. Figure 1 shows the network used in this model. Compartmentalized models can also be used when modeling more complex communities, by adding more bacterial compartments in the model[56]. Additionally, the interaction between the gut of the host and the bacterial community can be investigated, by adding a compartment depicting the gut within the environment[57].

As mentioned before, the objective function of classic FBA is usually to optimize for biomass formation. However, with multiple organisms in the system that all have their own biomass function with different growth rates and composition, the objective function cannot be the optimization of a single biomass function. The OptCom framework gives a solution for this problem. With the OptCom

method a multilevel objective function is introduced. This objective function consists of a community level objective function, which maximizes the total community biomass. Furthermore this objective function consists of biomass optimization function for the growth of each individual member of the community. Thus incorporating both the egoistic growth of each individual, while still keeping a healthy community[35, 58].

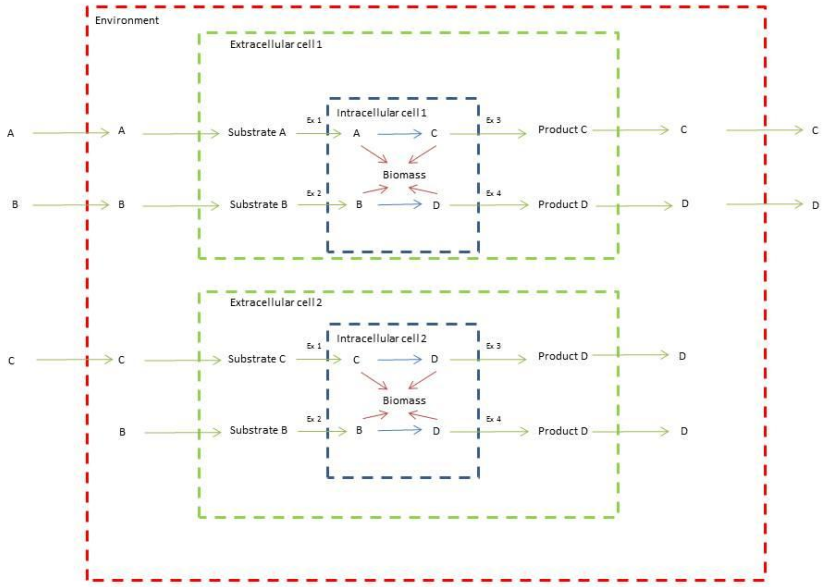

Figure 5: A simple model representing a compartmentalized community model. The bacterial cells have both an intracellular (blue dashed lines) and extracellular space (dashed green lines). The shared environment is everything between the red and green dashed lines. In the shared environment substrate A, B, C and D are present. Substrate A and B are taken up by cell 1. Substrate C and D are taken up by cell 2. The blue arrows depict conversion of a substrate into a product. The green arrows depict exchange of a metabolite. With this model cell 1 is producing C and releasing it in the shared environment. Cell 2 can take up this metabolite and converting it to product D.

During the day the metabolite availability within the gut changes due to different nutrient intake. Thus the metabolite availability is dynamic. Compartmentalized community models are working under steady-state conditions. Thus dynamic changes in the environment and dynamic interactions between microbial species cannot be modelled[49, 55]. With another type of model these dynamic interactions can be investigated. These models are called population-based models. A population-based model combines compartmentalized community models with the static optimization approach (SOA) from dynamic flux balance analysis (dFBA)[47]. The SOA-dFBA essentially makes a series of snapshots using FBA. The starting conditions of each individual snapshot are based on the optimal solution of the previous snapshot[59]. By making a series of snapshots from a compartmentalized community model, such as shown in figure 5, metabolite concentrations and the dynamic relationship between organisms can be studied. For this dynamic approach a multi-level OptCom framework was also developed[60]. With a population based model a recent study showed that metabolic reactions are dynamic over time in a multi-species interaction network[61]. Furthermore, a

study conducted with a community of nine gut species was able to predict the abundances under different nutrient compositions[62].

Both compartmentalized community models and population-based models can only be applied when the individual cells are all exposed to the same environment. Thus each cell can only use the metabolites that are present in the shared environment. However, this does not represent the human gut. Figure 6 shows a schematic overview of the human gut. Different parts of the gut have different numbers of bacteria and a different species distribution, both due to different conditions in each part[63]. For example, the stomach has a high acidity, there is an oxygen gradient along the human gut and metabolites produced by bacteria can only be utilized by bacteria further down the gut.

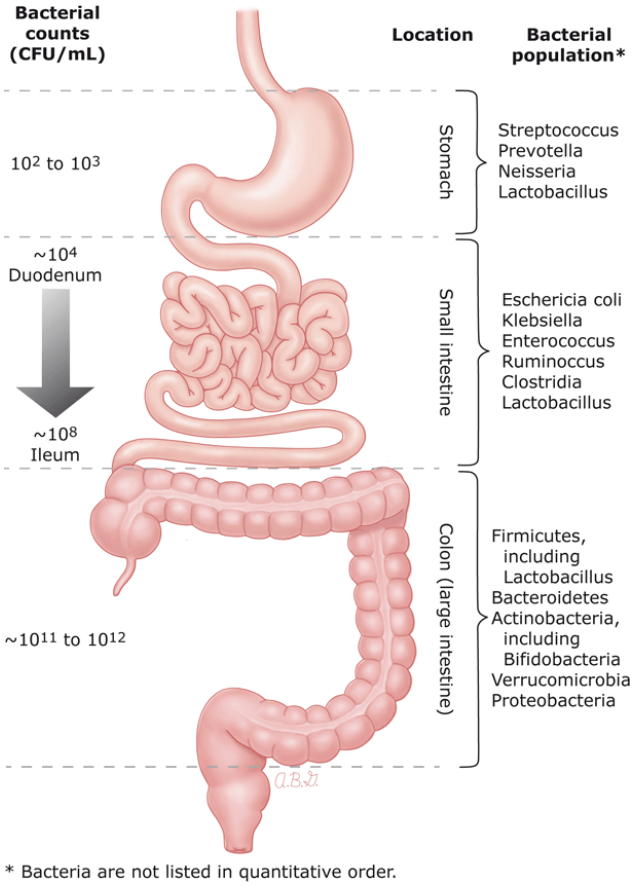

Figure 6: Overview of the human gut with the amount and distribution of bacteria colonizing the different parts[63].

Population based models do not take the spatial differences of the human gut into account. To model these spatial differences, individual-based community models can be made. This approach combines spatial-FBA with dFBA. Essentially this method places several compartmentalized models next to each other. Differential equations are being used to create a concentration gradient for each metabolite. In this way each compartment has its own concentration of metabolites based on the metabolite concentration in the compartment next to it. Furthermore, dFBA as previously described can be used to mimic the dynamic interactions of the different compartments[31, 47, 64]. Therefore

complex dynamic behavior arises as emergent properties. Figure 7 shows a simplified version of an individual-based community model.

With an individual-based community model van Hoek et al. showed that a bacterial community rearranges itself along metabolite gradients, whereby each bacteria makes optimal use of its metabolic capabilities[31]. Bauer et al. constructed an indiviual-based community model that consists of seven gut microbes and investigated the community structure in the presence and absence of a mucus glycan gradient. The site of the model with highest concentration of mucus glycan represents the gut epithelia since that is the place with the mucus layer. With a mucus glycan gradient, the bacterial community shaped itself whereby a mucus degrading organism (*B. thetaiomicron)* was closest to the mucus layer. In the absence of the gradient no community shape was present, but there was an overgrowth of *E. coli*[64]. This is in accordance with experimental studies showing the spatial organisation of mucus degrading organisms in the gut[65]. A healthy mucus layer is important since a defect in the mucus layer may cause diseases[66]. Therefore a mucus gradient may shape the spatial organisation of the gut bacteria and thus protect against disease[64].

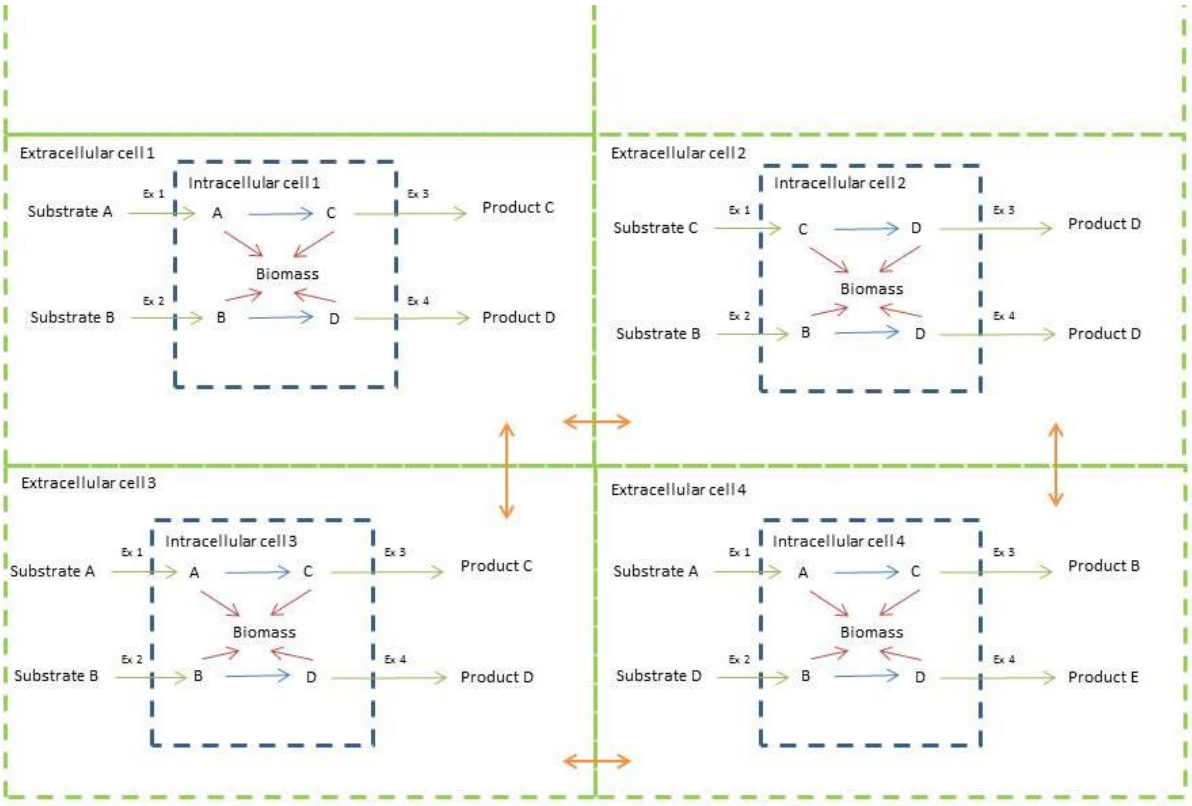

Figure 7: A simplified individual-based community model. Each cell depicts a different bacterial species (blue dashed lines) and sits within its own environment (green dashed lines). The orange arrows depicts the exchange reactions between different compartments. With this model each cell can have different concentrations of metabolite in its own compartment, which will affect the fluxes of each bacteria. Empty cells can be used to simulate moving cells or dividing cells.

Challenges and limitations of community modeling

Currently, a lot of research is done to improve the modeling of the gut microbiota. Different software is being developed that utilize the previously mentioned FBA methods. For example, OptCom[58] and classical FBA[35] use compartmentalized community models. D-OptCom[60], SteadyCom[62] and COMETS[67] use population based models. BacArena[64] and MatNet[68] use individual-based community models. Choosing the best modeling technique and software tool depends entirely on the research question. When interested in the interaction between two bacteria, compartmentalized community models are sufficient. However, when interested in the relationship between the gut bacteria and a disease state, population-based or individual-based community models are the best option.

A limitation in choosing the sufficient model is the model complexity. By adding dynamic and spatial properties, the model complexity increases. Furthermore, by increasing the number of bacteria in a community the complexity increases too. For population based models the increase in complexity scales linear with the number of species[47] and individual-based models scale linear with the number of individual bacteria[64]. The increasing complexity accounts for an increase in simulation time. Furthermore, increasing complexity leads to an increase in data. In turn, data analysis and data visualization will be more difficult. Therefore choosing the right modeling approach is always a compromise.

A further limitation of modeling the microbiota is the limited amount of information that describes the metabolic properties of the individual species present in the human gut. Recently Magnusdottir et al. published a list of 773 reconstructed metabolic models of bacterial species from the human gut[33]. However this is far from the total number of bacterial species found in the human gut[69]. Furthermore, the abundances of each bacterial species need to be determined, together with the presence of other organisms such as viruses and fungi. Especially since viruses and fungi might play a role in diseases such as IBD[25]. Furthermore, a single bacterial species can play a role the onset of disease, but can also play a role in keeping a healthy gut as shown with *B. fragilis*[41, 45]. These different roles may arise due to a different intestinal environment. Therefore models for each intestinal environment need to be made.

To prove that modeling the gut microbiota gives emergent properties and is able to come up with viable hypotheses, models need to be experimentally validated. However, designing experiments that take spatial distribution and dynamic interaction of several species at the same time into account is almost impossible. Furthermore, bacterial isolation is sometimes not possible or extremely difficult[70]. Thus making experimental validation of a model difficult too. Fortunately, new techniques come available in time. For example, with the rise of culturomics a lot of new bacteria can be cultured in the lab[71].

Perspectives of systems biology in microbiota research

Modeling and simulating metabolic interactions of the gut microbiota is a relatively novel way of studying emergent properties of the microbiota. However, with a FBA approach the effects of the addition or removal of a single or combination of multiple bacterial species can be studied. Therefore, the introduction of a probiotic to a complex bacterial community can be studied. In this manner the mechanism of action of the probiotic can be made clear, thus giving scientific evidence for the beneficial effects of the probiotic. Furthermore, adding novel bacteria to an in silico community might reveal possible targets for the development of novel probiotics. The same accounts for prebiotics. By adding the prebiotic to a human gut model, the abundances of bacterial community members can shift towards a higher abundance of beneficial bacteria. Therefore giving a mechanistic explanation of the workings of the probiotic. The effects of antibiotics or medicine can be studied in a similar way. Antibiotics such as vancomycin specifically target gram-positive bacteria, thus giving gram-negative bacteria a chance to dominate the human gut. This dysbiosis may lead to less diversity, thereby possibly causing diseases[72].

The top-down approach gives insight into the structure of the interactions between bacteria in the human gut. It can identify important bacteria as keystone species. By comparing healthy individuals with disease individuals, species that play a pivotal role in disease can be identified. Thus giving evidence to the cause of the disease. However, the identification of the microbiota as the cause of disease can only be identified if a keystone species can be identified[42]. By developing bottom-up networks that include intestinal cells, the cause of disease onset might be studied, by manually altering the abundances of community members and look at the effects on the intestinal cells. Additionally, the output of the intestinal cells can be altered and the effects on the microbiota can be studied.

To accurately determine the effect of pro- and prebiotics and determining the cause of microbial related disease further developments in both the bottom-up and top-down approach are needed. Models should include intestinal cells of the host. Currently, only a small number of models include intestinal cells of the host[57, 73]. Although recently 773 metabolic models were constructed[33], more metabolic models of gut bacteria need to be constructed and validated. Furthermore, differences in microbial composition between individuals with a disease and healthy individuals need to be elucidated for the accurate identification of keystone species.

References

1. Sender Ron R, Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biology* 14 (2016).

2. Louis Petra P & Flint, H. J. Formation of propionate and butyrate by the human colonic microbiota. *Environ. Microbiol.* 19, 29-41 (2017).

3. Koh Ara A, De Vadder, F., Kovatcheva-Datchary, P. & Backhed, F. From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* 165, 1332-1345 (2016-6-02).

4. Honda Kenya K & Littman, D. R. The microbiota in adaptive immune homeostasis and disease. *Nature* 535, 75-84 (2016-07).

5. Buffie Charlie G CG & Pamer, E. G. Microbiota-mediated colonization resistance against intestinal pathogens. *Nature Reviews Immunology* 13, 790-801 (2013-11).

6. Sonnenburg Justin L JL, Chen, C. T. L. & Gordon, J. I. Genomic and metabolic studies of the impact of probiotics on a model gut symbiont and host. *PLoS Biology* 4 (2006-11).

7. Sampson Timothy R TR & Mazmanian, S. K. Control of brain development, function, and behavior by the microbiome. *Cell Host and Microbe* 17, 565-76 (2015-5-13).

8. Yano Jessica M JM, Yu, K., Donaldson, G. P., Shastri, G. G. & Ann, P. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* 161, 264-76 (2015-4-09).

9. Manichanh C C, Rigottier-Gois, L., Bonnaud, E., Gloux, K. & Pelletier, E. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 55, 205-11 (2006-2).

10. Mukhopadhya Indrani I, Thomson, J. M., Hansen, R., Berry, S. H. & El-Omar, E. M. Detection of Campylobacter concisus and other Campylobacter species in colonic biopsies from adults with ulcerative colitis. *PLoS ONE* 6 (2011).

11. Qin Junjie J, Li, Y., Cai, Z., Li, S. & Zhu, J. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55-60 (2012-10-04).

12. Larsen Nadja N, Vogensen, F. K., van den Berg, F. W. J., Nielsen, D. S. & Andreasen, A. S. Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS ONE* 5 (2010-2-05).

13. Keshavarzian Ali A, Green, S. J., Engen, P. A., Voigt, R. M. & Naqib, A. Colonic bacterial composition in Parkinson's disease. *Movement Disorders* 30, 1351-60 (2015-9).

14. Zheng P P, Zeng, B., Zhou, C., Liu, M. & Fang, Z. Gut microbiome remodeling induces depressive-like behaviors through a pathway mediated by the host's metabolism. *Mol. Psychiatry* 21, 786-96 (2016).

15. Iannitti T T & Palmieri, B. Therapeutical use of probiotic formulations in clinical practice. *Clinical Nutrition* 29, 701-25 (2010-12).

16. Aizawa Emiko E, Tsuji, H., Asahara, T., Takahashi, T. & Teraishi, T. Possible association of Bifidobacterium and Lactobacillus in the gut microbiota of patients with major depressive disorder. ***J. Affect. Disord.*** **202, 254-7 (2016-15).**

17. Messaoudi Michaël M, Lalonde, R., Violle, N., Javelot, H. & Desor, D. Assessment of psychotropic-like properties of a probiotic formulation (Lactobacillus helveticus R0052 and Bifidobacterium longum R0175) in rats and human subjects. ***Br. J. Nutr.*** **105, 755-64 (2011-3).**

18. Slykerman R F RF, Hood, F., Wickens, K., Thompson, J. M. D. & Barthow, C. Effect of Lactobacillus rhamnosus HN001 in Pregnancy on Postpartum Symptoms of Depression and Anxiety: A Randomised Double-blind Placebo-controlled Trial. ***EBioMedicine*** **24, 159-165 (2017-10).**

19. Lindsay J O JO, Whelan, K., Stagg, A., Gobin, P. & Al-Hassi, H. Clinical, microbiological, and immunological effects of fructo-oligosaccharide in patients with Crohn's disease. ***Gut*** **55, 348-55 (2006-3).**

20. Benjamin Jane L JL, Hedin, C. R. H., Koutsoumpas, A., Ng, S. C. & McCarthy, N. E. Randomised, double-blind, placebo-controlled trial of fructo-oligosaccharides in active Crohn's disease. ***Gut*** **60, 923-9 (2011-7).**

21. Suez Jotham J, Zmora, N., Zilberman-Schapira, G., Mor, U. & Dori-Bachash, M. Post-Antibiotic Gut Mucosal Microbiome Reconstitution Is Impaired by Probiotics and Improved by Autologous FMT. ***Cell*** **174, 1406-1423 (2018-9-06).**

22. Rondanelli Mariangela M, Faliva, M. A., Perna, S., Giacosa, A. & Peroni, G. Using probiotics in clinical practice: Where are we now? A review of existing meta-analyses. ***Gut Microbes*** **8, 521-543 (2017-02).**

23. Zmora Niv N, Zilberman-Schapira, G., Suez, J., Mor, U. & Dori-Bachash, M. Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. ***Cell*** **174, 1388-1405 (2018-9-06).**

24. Ridaura Vanessa K VK, Faith, J. J., Rey, F. E., Cheng, J. & Duncan, A. E. Gut microbiota from twins discordant for obesity modulate metabolism in mice. ***Science*** **341 (2013-9-06).**

25. Ni Josephine J, Wu, G. D., Albenberg, L. & Tomov, V. T. Gut microbiota and IBD: causation or correlation? ***Nature Reviews Gastroenterology and Hepatology*** **14, 573-584 (2017-10).**

26. Assa Amit A, Butcher, J., Li, J., Elkadri, A. & Sherman, P. M. Mucosa-Associated Ileal Microbiota in New-Onset Pediatric Crohn's Disease. ***Inflamm. Bowel Dis.*** **22, 1533-9 (2016).**

27. Hansen Richard R, Russell, R. K., Reiff, C., Louis, P. & McIntosh, F. Microbiota of de-novo pediatric IBD: increased Faecalibacterium prausnitzii and reduced bacterial diversity in Crohn's but not in ulcerative colitis. ***Am. J. Gastroenterol.*** **107, 1913-22 (2012-12).**

28. Van Meervelt Veerle V, Soskine, M. & Maglia, G. Detection of two isomeric binding configurations in a protein-aptamer complex with a biological nanopore. ***ACS Nano*** **8, 12826-35 (2014-12-23).**

29. Wolkenhauer Olaf O. Why model? ***Frontiers in Physiology*** **5 (2014).**

30. Stelling Jörg J. Mathematical models in microbial systems biology. *Curr. Opin. Microbiol.* 7, 513-8 (2004-10).

31. Hoek Milan J A van MJAV & Merks, R. M. H. Emergence of microbial diversity due to cross-feeding interactions in a spatial model of gut microbial metabolism. *BMC Systems Biology* 11 (2017-5-16).

32. Carroll I M IM, Ringel-Kulka, T., Siddle, J. P. & Ringel, Y. Alterations in composition and diversity of the intestinal microbiota in patients with diarrhea-predominant irritable bowel syndrome. *Neurogastroenterology and Motility* 24, 521-30 (2012-6).

33. Magnúsdóttir Stefanía S, Heinken, A., Kutt, L., Ravcheev, D. A. & Bauer, E. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* 35, 81-89 (2017-1).

34. Faust Karoline K, Sathirapongsasuti, J. F., Izard, J., Segata, N. & Gevers, D. Microbial co-occurrence relationships in the human microbiome. *PLoS Computational Biology* 8 (2012).

35. El-Semman Ibrahim E IE, Karlsson, F. H., Shoaie, S., Nookaew, I. & Soliman, T. H. Genome-scale metabolic reconstructions of Bifidobacterium adolescentis L2-32 and Faecalibacterium prausnitzii A2-165 and their interaction. *BMC Systems Biology* 8 (2014-4-03).

36. Kitano Hiroaki H. Systems biology: a brief overview. *Science* 295, 1662-4 (2002-3-01).

37. Shahzad Khuram K & Loor, J. J. Application of Top-Down and Bottom-up Systems Approaches in Ruminant Physiology and Metabolism. *Curr. Genomics* 13, 379-94 (2012-8).

38. Grosskopf Tobias T & Soyer, O. S. Synthetic microbial communities. *Curr. Opin. Microbiol.* 18, 72-7 (2014-4).

39. Faust Karoline K & Raes, J. Microbial interactions: from networks to models. *Nature Reviews Microbiology* 10, 538-50 (2012-7-16).

40. Berry David D & Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in Microbiology* 5 (2014).

41. Fisher Charles K CK & Mehta, P. Identifying keystone species in the human gut microbiome from metagenomic timeseries using sparse linear regression. *PLoS ONE* 9 (2014).

42. Wu Shaoguang S, Rhee, K., Albesiano, E., Rabizadeh, S. & Wu, X. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* 15, 1016-22 (2009-9).

43. Hajishengallis George G, Darveau, R. P. & Curtis, M. A. The keystone-pathogen hypothesis. *Nature Reviews Microbiology* 10, 717-25 (2012-10).

44. Banerjee Samiran S, Schlaeppi, K. & van der Heijden, M. G. A. Keystone taxa as drivers of microbiome structure and functioning. *Nature Reviews Microbiology* 16, 567-576 (2018-9).

45. Hsiao Elaine Y EY, McBride, S. W., Hsien, S., Sharon, G. & Hyde, E. R. Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* 155, 1451-63 (2013-12-19).

46. Kuczynski Justin J, Liu, Z., Lozupone, C., McDonald, D. & Fierer, N. Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. ***NATURE METHODS*** **7, 813-9 (2010-10).**

47. Bauer Eugen E & Thiele, I. From Network Analysis to Functional Metabolic Modeling of the Human Gut Microbiota. ***mSystems*** **3 (2018 May-Jun).**

48. Orth Jeffrey D JD, Thiele, I. & Palsson, B. O. What is flux balance analysis? ***Nat. Biotechnol.*** **28, 245-8 (2010-3).**

49. Gottstein Willi W, Olivier, B. G., Bruggeman, F. J. & Teusink, B. Constraint-based stoichiometric modelling from single organisms to microbial communities. ***Journal of The Royal Society Interface*** **13 (2016).**

50. Khandelwal Ruchir A RA, Olivier, B. G., Roling, W. F. M., Teusink, B. & Bruggeman, F. J. Community flux balance analysis for microbial consortia at balanced growth. ***PLoS ONE*** **8 (2013).**

51. Sauer Uwe U. Metabolic networks in motion: 13C-based flux analysis. ***Molecular Systems Biology*** **2 (2006).**

52. Magnúsdóttir Stefanía S & Thiele, I. Modeling metabolism of the human gut microbiome. ***Curr. Opin. Biotechnol.*** **51, 90-96 (2018-6).**

53. Zelezniak Aleksej A, Andrejev, S., Ponomarova, O., Mende, D. R. & Bork, P. Metabolic dependencies drive species co-occurrence in diverse microbial communities. ***Proceedings of the National Academy of Sciences of the United States of America PNAS*** **112, 6449-54 (2015-5-19).**

54. Thiele Ines I & Palsson, B. O. A protocol for generating a high-quality genome-scale metabolic reconstruction. ***NATURE PROTOCOLS*** **5, 93-121 (2010-1).**

55. Klitgord Niels N & Segre, D. Environments that induce synthetic microbial ecosystems. ***PLoS Computational Biology*** **6 (2010-11-18).**

56. Heinken Almut A & Thiele, I. Anoxic Conditions Promote Species-Specific Mutualism between Gut Microbes In Silico. ***Appl. Environ. Microbiol.*** **81, 4049-61 (2015-6-15).**

57. Heinken, A., Sahoo, S., Fleming, R. M. T. & Thiele, I. Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. ***Gut Microbes*** **4, 28-40 (2013).**

58. Zomorrodi Ali R AR & Maranas, C. D. OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. ***PLoS Computational Biology*** **8 (2012-2).**

59. Mahadevan Radhakrishnan R, Edwards, J. & Doyle, F. Dynamic flux balance analysis of diauxic growth in Escherichia coli. ***Biophys. J.*** **83, 1331-40 (2002-9).**

60. Zomorrodi Ali R AR, Islam, M. M. & Maranas, C. D. d-OptCom: Dynamic multi-level and multi-objective metabolic modeling of microbial communities. ***ACS Synthetic Biology*** **3, 247-57 (2014-4-18).**

61. Granger Brian R BR, Chang, Y., Wang, Y., DeLisi, C. & Segre, D. Visualization of Metabolic Interaction Networks in Microbial Communities Using VisANT 5.0. ***PLoS Computational Biology*** **12 (2016-4).**

62. Chan Siu Hung Joshua SHJ, Simons, M. N. & Maranas, C. D. SteadyCom: Predicting microbial abundances while ensuring community stability. ***PLoS Computational Biology*** **13 (2017-5).**

63. , - 487.

64. Bauer Eugen E, Zimmermann, J., Baldini, F., Thiele, I. & Kaleta, C. BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. ***PLoS Computational Biology*** **13 (2017-5).**

65. Earle Kristen A KA, Billings, G., Sigal, M., Lichtman, J. S. & Hansson, G. C. Quantitative Imaging of Gut Microbiota Spatial Organization. ***Cell Host and Microbe*** **18, 478-88 (2015-10-14).**

66. Johansson Malin E V ME. Mucus layers in inflammatory bowel disease. ***Inflamm. Bowel Dis.*** **20, 2124-31 (2014-11).**

67. Harcombe William R WR, Riehl, W. J., Dukovski, I., Granger, B. R. & Betts, A. Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. ***Cell Reports*** **7, 1104-15 (2014-5-22).**

68. Biggs Matthew B MB & Papin, J. A. Novel multiscale modeling tool applied to Pseudomonas aeruginosa biofilm formation. ***PLoS ONE*** **8 (2013).**

69. Thursby Elizabeth E & Juge, N. Introduction to the human gut microbiota. ***Biochem. J.*** **474, 1823-1836 (2017-16).**

70. Eckburg Paul B PB, Bik, E., Bernstein, C., Purdom, E. & Dethlefsen, L. Diversity of the human intestinal microbial flora. ***Science*** **308, 1635-8 (2005-6-10).**

71. Lagier Jean-Christophe JC, Dubourg, G., Million, M., Cadoret, F. & Bilen, M. Culturing the human microbiota and culturomics. ***Nature Reviews Microbiology*** **16, 540-550 (2018-01).**

72. Meijnikman Abraham S AS, Gerdes, V. E., Nieuwdorp, M. & Herrema, H. Evaluating Causality of Gut Microbiota in Obesity and Diabetes in Humans. ***Endocr. Rev.*** **39, 133-153 (2018-01).**

73. Heinken, A. & Thiele, I. Systematic prediction of health-relevant human-microbial co-metabolism through a computational framework. ***Gut Microbes*** **6, 120-130 (2015).**