A Graph-based Molecular Communications Model Analysis of the Human Gut Bacteriome

Samitha Somathilaka, *Student Member, IEEE,* Daniel P. Martins, *Member, IEEE,* Wiley Barton, Orla O'Sullivan, Paul D. Cotter, Sasitharan Balasubramaniam, *Senior Member, IEEE*

***Abstract*—Alterations in the human Gut Bacteriome (GB) can be associated with human health issues, such as type-2 diabetes and obesity. Both external and internal factors can drive changes in the composition and in interactions of the human GB, impacting negatively on the host cells. This paper focuses on the human GB metabolism and proposes a two-layer network system to investigate its dynamics. Furthermore, we develop an *in-silico* simulation model (virtual GB), allowing us to study the impact of the metabolite exchange through molecular communications in the human GB network system. Our results show that regulation of molecular inputs strongly affects bacterial population growth and creates an unbalanced network, as shown by shifts in the node weights based on the produced molecular signals. Additionally, we show that the metabolite molecular communication production is greatly affected when directly manipulating the composition of the human GB network in the virtual GB. These results indicate that our human GB interaction model can help to identify hidden behaviours of the human GB depending on molecular signal interactions. Moreover, the virtual GB can support the research and development of novel medical treatments based on the accurate control of bacterial population growth and exchange of metabolites.**

***Index Terms*—Biological network systems, graph analysis, molecular communications, human gut bacteriome, metabolic interactions.**

I. Introduction

THE Gut Bacteriome (GB) is an ecosystem of a massive number of bacterial cells which play a vital role in maintaining the stability of the host's metabolism [1]. The bacterial populations of the GB build complex interaction networks by exchanging metabolites with the host and/or other bacterial populations [2], resulting in the production of new metabolites, such as Short Chain Fatty Acids (SCFAs), proteins, and other molecules [3].

External factors such as the availability of nutrients, antibiotics, and pathogens can affect this interaction network [4]. These factors mainly alter the compositional balance of the human GB, subsequently disrupting the metabolite production [5]. In humans, these GB changes have a significant impact on the host's health and may lead to many diseases, including inflammatory bowel disease, type-2 diabetes, obesity and cancers [6], [7]. Although studying complex causal metabolic networks is challenging [8], several studies have been undertaken to precisely identify the causes for microbial behavioural alterations and their consequent health effects in humans and animals [9], [10]. For example, Yang et al. [11] performed a cross-sectional whole-genome shotgun metagenomics analysis of the microbiome and proposed a combinatorial marker panel to demarcate microbiome-related major depressive disorders from a healthy microbiome. From a different perspective, Kim et al. introduced a split graph model to analyse the microbial compositions of healthy or Crohn's disease microbiome compositions [12]. Inspired by these works, we propose a novel tool to further characterise the interactions among the bacterial populations often found in the human GB.

In this paper, we propose a two-layer interaction model supported by the exchange of molecular signals, i.e. metabolites, to model the human GB. Here, we identify the interactions between bacterial cells as **Molecular Communications** (MC) systems and their collective behaviour as a MC network. MC aims to model the communication between biological components [13] using molecules as information [14], [15] and it is fundamental to characterise the exchange of metabolites in our two-layer interaction model.

In the graph network, bacterial populations act as nodes while the edges represent the interactions between them. This interpretation allows quantifying the behaviours of the human GB using graph theoretical incorporating MC analysis to understand impacts from distances between different graph states and variations of node/edge weights. Moreover, conducting *in-vivo* or *in-vitro* experiments on the human GB to extract data related to each interaction of the network often requires a significant number of resources and time. On the other hand, calculating them theoretically using Flux Balance Analysis (FBA) is extensively complex due to the large number of variables that prompt the same number of equations to be solved (see in Section IV for further details). On top of that, FBA is known as a static approach that fails to capture the stochastic nature of biological networks. Hence, we designed an agent-based simulator (henceforth named virtual GB) to simulate the human GB, which produces the same set of data that we expect by conducting *in-vivo* or *in-vitro* experiments or FBA calculations. The virtual GB performs the behaviours of the human GB considering natural characteristics. Hence, the generated data represents bacterial behaviours that are influenced by the aforementioned stochastic parameters.

S. Somathilaka, Daniel P. Martins are with VistaMilk Research Centre and the Walton Institute for Information and Communication Systems Science, Waterford Institute of Technology, Waterford, X91 P20H, Ireland. E-mail: samitha.somathilaka,daniel.martins@waltoninstitute.ie.

S. Balasubramaniam is with the Department of Computer , University of Nebraska-Lincoln, 104 Schorr Center, 1100 T Street, Lincoln, NE, 68588-0150, USA. E-mail:sasitharanb@gmail.com

Wiley Barton, Orla O'Sullivan and Paul D. Cotter are with VistaMilk Research Centre and the Teagasc, Food Research Centre, Moorepark, Ireland, P61 C996. E-mail: paul.cotter@teagasc.ie

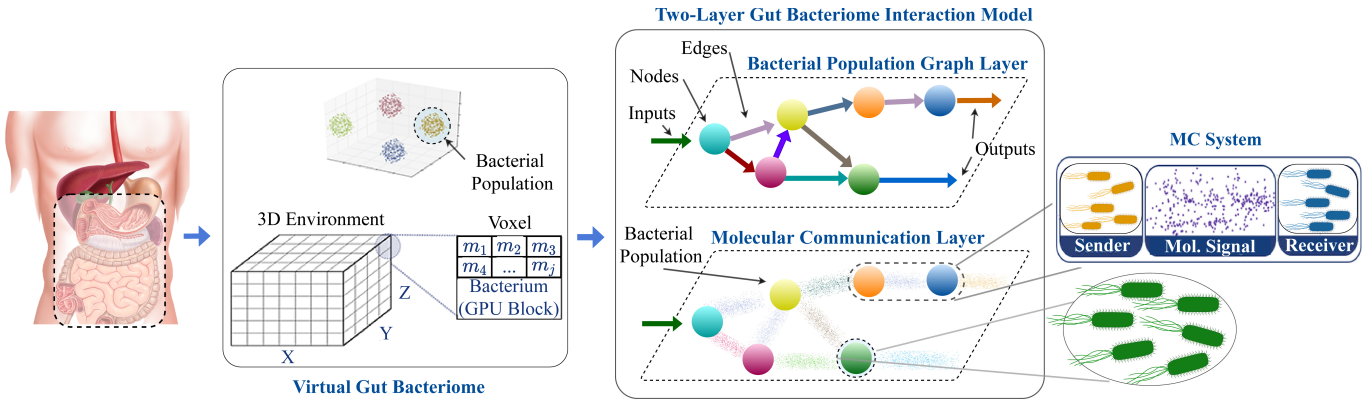


Fig. 1: Illustration of the system model. (a) We recreated the human GB functionalities on virtual GB using voxel architecture and parallel processing dedicating one GPU block for each bacterial cell to produce quantitative data on the MC layer, and (b) we propose a two-layer system model to investigate the molecular interactions simulated in the virtual GB.

Our main contributions are as follows:

- **Design of a two-layer interaction model of the human GB**: The collective gut bacteria metabolism forms a complex interaction network among the different bacterial populations. Hence, in this study, we design a layered interaction model to investigate the dynamics of the human GB based on the exchange of metabolites.
- **Analysing molecular communication impact on the human GB graph structure:** Deviations of bacterial populations' metabolism cause alterations in molecular interaction within the human GB, which may impact the graph layer structure. We analyse this relationship between the MC measures and the graph structure of the human GB in terms of graph nodes and edges behaviours.
- **Development of a human GB simulator to perform *in-silico* experiments:** We design and utilise an *in-silico simulation model* of the human GB to investigate the direct and hidden interactions among bacterial populations based on the exchange of metabolites.

In the next sections, we detail our approach to model the human GB and assess its network performance. In the section II, we describe the basics of the human GB and highlight the existent gaps that this research aims to address. Our proposed model is detailed in Section III. Then, the metrics considered in this paper are introduced in Section IV, and our analysis results are presented in Section V. Further, in Section V-A, we introduce the simulation environment built to utilise metagenomics data and perform *in-silico* experiments with the human GB. Finally, our conclusions are shown in Section VI.

II. Background on the Human GB Model

The human GB is the bacterial ecosystem residing inside the human digestive system, comprising approximately 1000 species interacting with each other and carrying out crucial functions such as nutrient metabolism and immunomodulation of the host [16]. These bacteria utilise products of host metabolism, metabolites produced by other bacteria or dietary components from the gastrointestinal tract to convert into various products essential for the host through different metabolic pathways [17]. Bacteria in the human GB manifest their cellular functions by exhibiting various social behaviours such as commensalism [18], and competition by interacting with other populations mainly using molecules (e.g., proteins, metabolites and *quorum* sensing) rather than individual entities [19]. We identify these interactions as MC system and assert that the communication process in the GB is quite similar to routing and relaying information in a conventional network system which has inspired different network models (including ours) of the human GB interactions. For example, Naqvi et al. used a network-based approach to characterise the human gut microbiome composition and analysed healthy vs diseased states using network statistics [20]. Another study focuses on the use of Boolean dynamic models that combines genome-scale metabolic networks to determine the metabolic deviations between community members, which was used to characterise their metabolic roles of interactions [21].

The composition of the human GB is a crucial driver for the processing of metabolites (i.e., small molecules produced and used in metabolic reactions) in the lower intestine, which significantly impacts the health of the host [22]. Human GB composition differs among individuals, and it depends on various factors, including dietary patterns, gut diseases, exercise regimes, antibiotic usage, age, and genetic profiles [23].

III. Two-Layer Human GB Interaction Model

In this paper, we represent the metabolic interactions of select representative bacterial genera of the human GB as a two-layer interaction model, as shown in Figure 1. First, the compositional and behavioural data on the human GB is extracted from the databases and literature and implemented the virtual GB (Figure 1a), see Section V-A for further details. The virtual GB then simulates the human GB functionalities according to various experimental setups (later explained in section V), producing data on bacterial, molecular and gut environmental behaviours. The produced data is analysed

according to the introduced two-layer interaction model as shown in Figure 1b.

The upper layer of this model, which is the bacterial population graph layer, defines the interconnections and overall structure of the human GB, where we model the bacterial populations and host as nodes and interactions between them as edges. To minimize the complexity of the model, this graph layer considers bacterial genera as nodes, as species in the same genus share a common ancestral origin and the data availability. Further, the edges of the network represent the direct connections between the nodes that produce a particular metabolite and the nodes that consume the corresponding metabolite. In this layer, we can investigate the network topology of the human GB.

The bottom layer consists of the cascading molecular communications systems created by the bacterial populations to establish their exchange of metabolites and support their network structure. Here, each node is viewed as a molecular transceiver, and the edges are the communications channels interconnecting the nodes. Furthermore, this model extends to the molecular signals that reach the human GB from the environment, as well as, the ones that are output from the human GB and return to the environment. The interactions represented in this layer are dynamic and will depend on several environmental conditions, such as media characteristics and human GB composition. Please note that this is the layer where we initially observe the impacts of any alterations on the human GB composition (we further model and analyse this effect in Section IV-A). The upper layer and the bottom layer are further described in the following sections.

A. Upper Layer - Bacterial Population Graph Layer

Bacteria display a wide variety of social behaviours, and this can lead to processes such as the metabolism of molecules or coordinated biofilm formation [24]. The bacteria's ability to consume and produce multiple metabolites results in dense interaction patterns that can lead to challenges in the analysis.

Our human GB interaction model aims to provide a better global view of the functionality of the human GB, leading to the understanding of the causes and effects of its imbalance and to propose precise alterations to fix such issues. Therefore, we model the human GB as follows. We first consider that all bacterial cells b^{B_k} of a bacterial population B_k (where k is the bacterial population identifier) perform the same series of metabolic functions to process the metabolite inputs in the human gut. Each node of the proposed graph layer is a bacterial population, and each edge is an interaction between two bacterial populations through metabolite exchange. The nodes of this layer comprise the collective metabolic functions of all cells within the corresponding population. Let Ω be the set of all agents in this study, i.e. host cells and bacterial populations, $\Omega = \{host, B_k\}$. In this case, the molecular intake of particular bacterial population $B_{k'}$ from Ω, $C_{(\Omega,B_{k'})}$ is considered $C_{(\Omega,B_{k'})} \simeq \sum c_{(\Omega,b^{B_{k'}})}$ where C represents population interactions, c represents the intercellular interactions and $c_{(\Omega,b^{B_{k'}})}$ is the molecular reception of bacterial cell $b^{B_{k'}}$ (a cell from the bacterial population $B_{k'}$) from a Ω source. In the same way, molecular emission of the population is considered the combined molecular emission of all bacterial cells of the particular population, $C_{(B_{k'},\Omega)} \simeq \sum c_{(b^{B_{k'}},\Omega)}$, where $C_{(B_{k'},\Omega)}$ is the molecular emission from population $B_{k'}$ to any receiver (host or other bacterial populations), and $c_{(b^{B_{k'}},\Omega)}$ is the molecular emission of a single bacterial cell of the population $B_{k'}$ to any receiver. Additionally, the metabolite consumed by the bacterial cell $b^{B_{k'}}$, $M_{Con}(b^{B_{k'}})$ is obtained as $M_{Con}(b^{B_{k'}}) = c_{(\Omega,b^{B'_k})} - c_{(b^{B'_k},\Omega)}$. Hence the metabolite consumption of a bacterial population is defined as $M_{Con}(B_{k'}) \simeq \sum M_{Con}(b^{B_{k'}})$.

Next, we map the interactions between bacterial populations to a directed multi-graph network, $\Gamma = (B, C, B^s, B^d, M)$, where B is the set of all bacterial populations, C is the set of all interaction in the human GB, $B^s \in B$ is the bacterial population interaction sources, $B^d \in B$ is the bacterial population interaction destinations, and M is the set of metabolites. In this work, we consider SCFAs production as the use case for our model on the bacterial population interactions.

B. Bottom Layer - MC System

As detailed in the previous section, the metabolism of nutrients by the human GB involves the reception, processing, production of metabolites. These activities are fundamental for the maintenance of the human GB, and this is modelled as the MC layer shown in Figure 1b. Our aim of having the two-layer model is to determine how the changes due to molecular signals of the metabolites will affect the relationship of the bacterial population graph layer. Therefore, any changes in the bottom layer directly affect the upper layer and vice-versa.

Here, we define the metabolites as the molecular signals that are exchanged by the nodes, which can assume different functions depending on the MC network structure. For example, when the node receives molecular signals, we model it as a receiver, and when processing and secreting molecular signals, we define them as transmitters based on the MC paradigm. The edges of the proposed MC network are represented as the MC channels to model the physical transport of molecular signals between the nodes by diffusion. Figure 1b shows a visual representation of the proposed bottom layer and its relationship with the upper layer.

The diffused molecular signal is received by the nodes which have the membrane receptors that will allow the metabolites to bind. The performance of this network node function (i.e., molecular reception) relies on many factors such as molecule size [25], ligand-receptor maximum attraction length and bond equilibrium [26], binding noise due to the Brownian motion of molecules near the receptors [27], and the minimum required concentration to be detected [28]. After receiving the molecular signals, the node will process them internally, which may result in the production of a new molecular signal to be transmitted to the next node (focus of this paper).

Received molecular signals are processed through signalling pathways and produce different metabolites that will be transmitted to the next node [29]. Even though we only focus on the genus level, the signal processing occurs in each

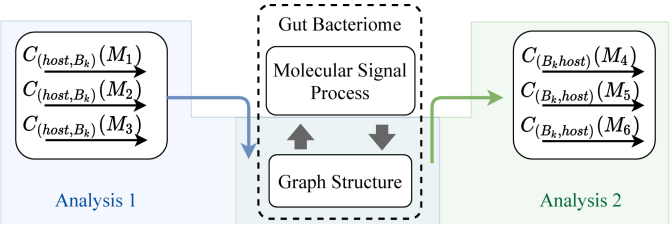


Fig. 2: Illustration of the analysis structure of this study. Analysis 1: Influence of inputs on the graph structure and Analysis 2: Behaviours of graph output against structural deviations.

bacterial cell. Accordingly, we present the signalling process performance of a bacterial cell b^{B_k} related to any metabolite M_j as $SPP_{b^{B_k}}(M_j)$. Let's assume that the cell b^{B_k} produces M_j by consuming another metabolite $M_{j'}$. Then the signal process performance $SPP_{b^{B_k}}(M_j)$ can be modelled by considering the metabolite $M_{j'}$ reception process (defined as $R_{b^{B_k}}(M_{j'})$), the encoding/decoding process from metabolite $M_{j'}$ to M_j (defined as $E_{b^{B_k}}(M_{j'}, M_j)$), and M_j metabolite secretion process by the cells in the bacterial population b^{B_k} (defined as $L_{b^{B_k}}(M_j)$). Hence, we represent the signal process performance as follows,

$$SPP_{b^{B_k}}(M_j) = f(R_{b^{B_k}}(M_{j'}), E_{b^{B_k}}(M_{j'}, M_j), L_{b^{B_k}}(M_j)). \tag{1}$$

Therefore, the SPP of the populations B_k can be modeled as follows,

$$SPP_{B_k}(M_j) = \sum SPP_{b^{B_k}}(M_j). \tag{2}$$

Since the output of the molecular signal processing is the emission of a particular molecular signal, it is fair to say,

$$SPP_{B_k}(M_j) = C^r_{(B_k,\Omega)}(M_j) \tag{3}$$

where $C^r_{(B_k,\Omega)}(M_j)$ is the rate of molecule M_j production by the bacterial population B_k to any node (either other bacterial populations or host cells).

IV. System Dynamics

We investigate the system dynamics of the human GB through a series of simulations using the virtual GB and propose a two-layer human GB model. First, we recreate the digital form of the human GB on the simulator, which is explained in depth later in Section V-A. Then we perform two main sets of experiments, as depicted in Figure 2. In the first set, we analyse the impact of the system's inputs on the connectivity structure of the virtual GB, and in the second set, we manipulate the composition of our virtual GB to investigate the impact on the metabolite production of our MC network. Through this second set of experiments, we aim to identify the nodes that can play a pivoting role in the GB imbalances.

In our analyses, first we define a standard graph state S_0, which represents the functionality of an average healthy human GB with the intention of quantifying structural changes and behavioural deviations relative to the standard structures. The average composition, interactions, and metabolite production dynamics were mainly considered in defining the S_0. The average composition and the interactions of S_0 for the case study of this paper is presented in Section V

A. Molecular input impact on the human GB structure

Due to the variety of bacterial behaviours induced by the exchange of molecules, some of the molecular input signals have a significant impact on the structure of the human GB (our focus), while others are directly converted into output metabolites. In this section, we detail how the molecular input signals impact the structure of our MC network. As the structural deviations of the graph is a crucial measurement in understanding the deviation of the human GB behaviour, the structural deviation is evaluated in terms of edges and nodes weight using the rates of the interaction of the nodes. Hence, we explain how the interaction rates can be calculated theoretically using FBA and are represented as follows,

$$F_{[k\times q]} \cdot \vec{C} = \vec{M}_{Con}(B_k) \tag{4}$$

where $F_{[k\times q]}$ is the stoichiometric matrix of k number of bacterial populations and q number of interactions based on the flux of metabolites between the nodes in the MC network. Here $\vec{C} = [C^r_1, C^r_2, ..., C^r_q]_{1\times q}$ and C^r_q is the rate of interactions for C_q. We can solve (4) as follows,

$$\begin{matrix} B_1 \\ B_2 \\ \vdots \\ B_k \end{matrix} \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,q} \\ a_{2,1} & a_{2,2} & \dots & a_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & \dots & a_{k,q} \end{pmatrix} \cdot \begin{pmatrix} C^r_1 \\ C^r_2 \\ \vdots \\ C^r_q \end{pmatrix} = \begin{pmatrix} \frac{\mathrm{d}M_{Con}(B_1)}{\mathrm{d}t} \\ \frac{\mathrm{d}M_{Con}(B_2)}{\mathrm{d}t} \\ \vdots \\ \frac{\mathrm{d}M_{Con}(B_k)}{\mathrm{d}t} \end{pmatrix} \tag{5}$$

where, $a_{k,q}$ is the stoichiometry of the interaction C^r_q for bacterial population B_k.

Based on (5), we can extract the relationship between rates of interactions starting from the node B_k using Mass Balance Equation (MBE), which is based on the following relationship

$$\frac{\mathrm{d}M_{Con}(B_k)}{\mathrm{d}t} = \sum_q a_{(k,q)} C^r_q. \tag{6}$$

On the other hand, the rate of molecular consumption can be modeled as follows [30],

$$\frac{\mathrm{d}M_{Con}(B_k)}{\mathrm{d}t} = -U_1 \left(\mu_k \frac{M_{Con}(B_k)}{M_{Con}(B_k) + K_{S1}} \right) N_{B_k} \tag{7}$$

where N_{B_k} is the bacterial concentration, μ_k is maximum growth rate, K_{S1} is the half-saturation constant of the bacteria, and U_1 is an utility parameter. Hence,

$$-U_1 \left(\mu_k \frac{M_{Con}(B_k)}{M_{Con}(B_k) + K_{S1}} \right) N_{B_k} = \sum_q a_{(k,q)} C^r_q. \tag{8}$$

By solving the series of MBEs, all the interaction rates can be calculated. This is a highly complex calculation due to the massive number of nodes, edges of the network, and a large number of parameters associated with the structural connections. The introduced virtual GB produces data on the rates of interactions avoiding complex FBA calculations.

The extracted rates of interactions are then used to quantify the graph structural changes in two ways. First, we investigate

the graph structural changes considering the behaviours of the node weights. Here, the statistical distances between the weights of the same node in different graph states are measured. The node weight, $B_k^w(S_g)$ of the bacterial population B_k in the graph state S_g is considered as the collective SPP and can be evaluated as follows,

$$B_k^w(S_g) = \sum_j SPP_{B_k}(M_j). \tag{9}$$

Alternatively, using (3) we compute the node weight as follows,

$$B_k^w(S_g) = \sum_j C^r_{(B_k,\Omega)}(M_j). \tag{10}$$

Based on this, $d(B_k^w : S_g, S_0)$ represents the distance of node B_k between the two graph states S_0 and S_g is evaluated as follows,

$$d(B_k^w : S_g, S_0) = B_k^w(S_g) - B_k^w(S_0). \tag{11}$$

Next, we quantify the structural deviation of the graph using the interaction changes. In this study, we consider static snapshots of different graph states that can enable the use of the *Hamming Distance* to evaluate graphical distances for two states [31] among other techniques. The Hamming distance $d_h(S_0, S_g)$ between the graph states S_g and the standard state S_0 is defined as the difference of two adjacent matrices corresponding to the two graph states. First, we define the adjacency matrix of the graph state S_g as follows,

$$\begin{array}{c} \\ B_1 \\ B_2 \\ \vdots \\ B_k \end{array} \begin{array}{c} \begin{array}{cccc} B_1 & B_2 & \cdots & B_k \end{array} \\ \begin{pmatrix} C^w_{(B_1,B_1)} & C^w_{(B_1,B_2)} & \cdots & C^w_{(B_1,B_k)} \\ C^w_{(B_2,B_1)} & C^w_{(B_2,B_2)} & \cdots & C^w_{(B_2,B_k)} \\ \vdots & \vdots & \ddots & \vdots \\ C^w_{(B_k,B_1)} & C^w_{(B_k,B_2)} & \cdots & C^w_{(B_k,B_k)} \end{pmatrix} \end{array} \tag{12}$$

where $C^w_{(*,*)}$ is the weight of the interaction $C_{(*,*)}$. Note that the weights of interactions in the main diagonal of the above matrix represents the interactions that take place within the same bacterial population, which is a type of interaction that cannot be observed in the metabolic network we considered in this study. Further, we define the weight of the interaction $C_{(B_k,B_{k'})}(M_j)$ between any bacterial population B_k and B'_k through metabolite M_j as follows,

$$C^w_{(B_k,B_{k'})}(M_j) = \frac{C^r_{(B_k,\Omega)}(M_j)}{\sum_k C^r_{(B_k,\Omega)}(M_j)} \cdot \frac{C^r_{(\Omega,B_{k'})}(M_j)}{\sum_{k'} C^r_{(\Omega,B_{k'})}(M_j)}. \tag{13}$$

Moreover, from released molecules by a bacterial population, only a fraction is consumed directly by the other populations and the rest will get accumulated in the environment. This means the most significant portion of molecular consumption by the bacterial populations is from the environment. We define this process with the help of a memory component concept as depicted in Figure 3. Since the metabolites are accumulated in the environment, we consider it a memory, then model the metabolite accumulation as an interaction starting from a bacterial population that releases the metabolites and ending with the memory, $C_{(B_k,Mem)}$. In the same way, the metabolite consumption from the environment is modelled as an interaction starting from the memory and ending with a bacterial population that consumes the particular metabolite, $C_{(Mem,B_k)}$. Hence, we modify (13) by applying the memory, which is represented as follows,

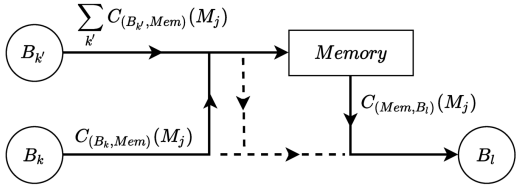


Fig. 3: Illustration of the environment working as a memory of molecules.

$$C^w_{(B_k,B_{k'})}(M_j) = \frac{C^r_{(B_k,Mem)}(M_j) C^r_{(Mem,B'_k)}(M_j)}{\sum_k C^r_{(B_k,Mem)}(M_j) \sum_{k'} C^r_{(Mem,B'_k)}(M_j)}. \tag{14}$$

Then, the Hamming distance, $d_h(S_0, S_g)$ can be represented as,

$$d_h(S_0, S_g) = \sum_{k,k'} |C^w_{(B_k,B_{k'})}(S_g) - C^w_{(B_k,B_{k'})}(S_0)| \tag{15}$$

where, $C^w_{(B_k,B_{k'})}(S_g)$ and $C^w_{(B_k,B_{k'})}(S_0)$ are the weights of interaction $C_{(B_k,B_{k'})}$ in graph states S_g and S_0 respectively.

B. Human GB structure impact on the molecular output

This analysis explores the impact of interaction variations of the human GB on the output. Here, we keep the inputs at an optimal level and manually alter the graph structure by changing the population sizes, which leads to variations in the SPP of the nodes. Then the output of the system is measured in different graph states and the weights of the edges are calculated using (13) to determine the molecular output of the MC layer using graph theory.

The ratio between the three SCFAs can be identified as a critical measurement to evaluate the metabolite production accuracy of the bacteriome. We adopt the signal to noise ratio (SNR) to evaluate the consistency of the output signal ratios. In this analysis, we calculate SNR of any signal $SNR(M_j)$, considering the other output signals, $M_{j'}$ as noise. This SNR value directly indicates the ratio between the molecular signal M_j and other metabolite signals $M_{j'}$. Then $SNR(M_j)$ is calculated as follows,

$$SNR(M_j) = \sum_k \frac{C_{(B_k,host)}(M_j)}{\sum_{j'} C_{(B_k,host)}(M_{j'})}. \tag{16}$$

Moreover, some bacterial populations do not produce specific SCFAs, but have an indirect influence on them. For example, *Bacteroides* cells do not produce butyrate, but the acetate produced by the *Bacteroides* cells is a substrate for the butyrate production by *Faecalibacterium* and *Roseburia* cells. Hence, the *Bacteroides* population indirectly influences the butyrate production. Considering the above mentioned effect,

a correlation matrix is generated for variation of node weights vs the collective SCFA output of the human GB to analyse the impact of various bacterial populations in SCFA production. Here, we denote the rate of SCFA M_j output by all the bacterial populations as $O^r(M_j)$ where

$$O^r(M_j) = \sum_k C^r_{(B_k,host)}(M_j). \tag{17}$$

Then, the correlation coefficient $r(B_k)$ of node weight B^w_k versus the collective output of M_j is calculated as follows.

$$r(B_k) = \sum_g \frac{\overline{B^w_k} - B^w_k(S_g)(\overline{O^r}(M_j) - O^r(M_j))}{\sqrt{(\overline{B^w_k} - B^w_k(S_g))^2}\sqrt{(\overline{O^r}(M_j) - O^r(M_j))^2}} \tag{18}$$

where, $\overline{O^r}(M_j)$ is the standard collective output rate for M_j by all the bacterial populations and $\overline{B^w_k}$ is the weight of the node B_k in the standard state S_0.

V. Analytical Results

In this section, we describe the development of the virtual GB and the results from our analysis that is based on the models presented in Section IV.

A. Virtual GB Design

We developed the virtual GB using metagenomic data to characterise the bacterial populations signalling interactions and their impact on the network relationships. The virtual GB is inspired by the BSim agent-based cell simulator [32]. The virtual GB is written in C++ with CUDA platform for parallel processing to increase the simulation performance and most importantly, mimic the parallel processing typically executed by the bacterial populations. We dedicate one GPU block for each bacterial cell, and the threads of that block to intracellular functions of the corresponding cell. To simulate the bacterial interactions, we model the exchange of molecules using metabolic flux in a diffusive media. The simulator has a 3D environment with voxel architecture (Figure 1a), which provides the ability of extracting data on each metabolite and bacterial cell separately. Moreover, we can introduce any new cell type by creating their internal metabolic pathways and other physiological characteristics such as motility, shape, size, etc. Therefore, the simulator can be used for a range of setups including other metabolic functions, microbial ecosystems in different habitats or targeting specific bacterial behaviour like quorum sensing. Further, the simulator can log data on the metabolite consumption/production/accumulation and bacterial proliferation. In this study, we setup the virtual GB to simulate the SCFA production using metagenomic and metabolomic data obtained in [29], KEGG [33]–[35], NJS16 [36], and MetaCyc databases [37].

Here, we present a series of analyses conducted on SCFA production within the human GB using the two-layer model. First, we defined the average composition of the human GB using the average relative abundance (RA) (see Table I) calculated based on data extracted from 352 samples of the MicrobiomeDB [38].

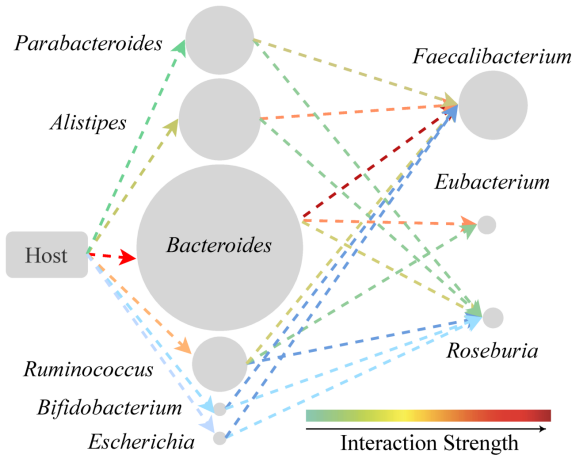


Fig. 4: Representation of the subgraph, Γ_{SCFA} considered in the case study which contains the nodes and edges related to SCFA production.

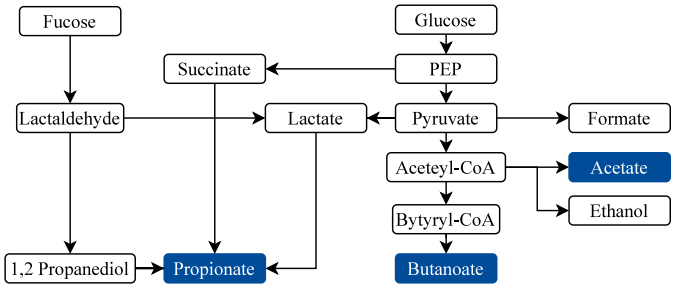


Fig. 5: Combined and simplified SCFA production pathway of converting fucose and glucose, into SCFAs.

TABLE I: Average RAs of bacterial populations

Genus	Average RA
Bacteroides	0.4899173
Alistipes	0.05960802
Faecalibacterium	0.04329791
Parabacteroides	0.04096428
Ruminococcus	0.03320183
Roseburia	0.01039938
Eubacterium	0.0093219
Bifidobacterium	0.00179366
Escherichia	0.00185639

Using these RA data along with the extracted interaction data from the databases mentioned earlier, we created a graph network for SCFA production, Γ_{SCFA} following the definitions presented in Section III-A. We only considered nine bacterial populations based on their RA, their metabolic activities, and data availability. We include *Bacteroides*, *Alistipes*, *Faecalibacterium*, *Parabacteroides*, and *Ruminococcus* in the model as they are the most abundant bacterial genera. To add further metabolic diversity to the network, we include other bacterial genera used in this study as they perform different metabolic functions compared to the most abundant bacterial genera. Figure 4 illustrates the Γ_{SCFA} where node sizes indicate the RAs of the respective bacterial genera shown in Table I. Furthermore, the edges are colour-coded to highlight the strengths of the interactions which are quantified using

TABLE II: Parameters utilised in Section V-B and V-C.

Parameter	Value	Description
Standard setup		
Bacteroides cell count	7.3×10^5	
Alistipes cell count	8.9×10^4	
Faecalibacterium cell count	6.4×10^4	
Parabacteroides cell count	6.1×10^4	Calculated based on the RA (see Table I) and to keep the total cell count
Ruminococcus cell count	4.9×10^4	lesser than 1.5×10^6 (maximum number of cells was limited by the memory
Roseburia cell count	1.5×10^4	availability of our GPU server).
Eubacterium cell count	1.3×10^4	
Bifidobacterium cell count	1.5×10^3	
Escherichia cell count	1.5×10^3	
Maximum GPU blocks utilised	1.7×10^6	Calculated based on number of voxels.
Maximum threads per utilised per block	24	Calculated based on number of metabolites used in the simulations.
Analysis 1: Molecular input impact on the human GB structure		
Glucose input rate (min-max)	0.000-16.605 μmol/m^{3}s	Calculated the using the number of cells and the stoichiometry of metabolic pathways [37], [39].
All bacterial population cell count	Fixed at standard values	Values are same as the standard setup.
Analysis 2: Human GB structure effect on the graph outputs		
Glucose input rate	6.642 μmol/m^{3}s (Fixed)	Obtained from simulations to match SCFA production of average human GB.
***Bacteroides* setup**		
Bacteroides cell count (min-max)	0-1.6×10^6	Calculated using the stoichiometry of metabolic pathways of *Bacteroides* and the number of cells to obtain results range with significant changes [37].
Other bacterial population counts	Fixed at standard values	Values are same as the standard setup.
***Faecalibacterium* setup**		
Faecalibacterium cell count (min-max)	0-1.4×10^5	Calculated using the stoichiometry of metabolic pathways of *Faecalibacterium* and the number of cells to obtain results range with significant changes [37].
Other bacterial population counts	Fixed at standard values	Values are same as the standard setup.

(13).

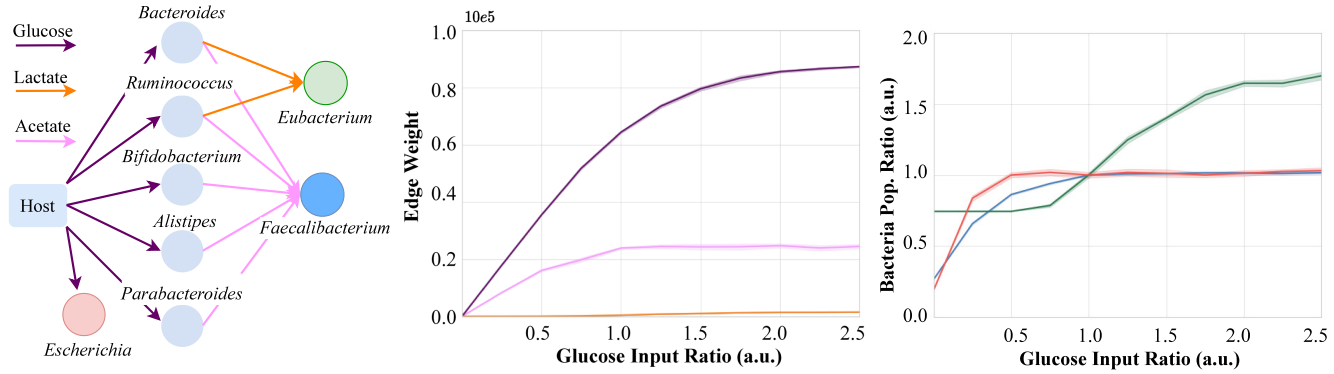
For illustration purposes, we combine the metabolic processes executed on different bacterial cells and simplify the SCFA pathway to focus on the most important steps that leads to the production of the three most abundant SCFAs in the human GB, namely acetate, butyrate and propionate (see Figure 5) [40]. The parameters utilised in V-B and V-C are presented in Table II. As we explained earlier, the bacterial cell counts for the standard setup are calculated based on the calculated RA and to keep the total cell count less than 1.5×10^6. The number of GPU blocks equals to the number of voxels in the system and the maximum number of threads per block calculated based on the number of metabolites in the environment. Further, the glucose input rate is extracted by an array of iterative experiments to match the ratio of SCFA abundance of an average human GB. Please note that in a typical human GB, SCFA abundance ratios range from 3:1:1 to 10:2:1 [41]. The maximum glucose input rate of the Analysis 1, and the maximum *Bacteroides* and *Faecalibacterium* cell counts of the Analysis 2 are fixed at certain values to obtain results with significant behaviours. Beyond those maximum values, the results only continue the trends without significant changes.

B. Analysis 1: Molecular input effects on the graph structure

Here, we present the results for the analyses mentioned in Section IV-A. The analyses are conducted by regulating the input glucose rate $C^r_{(host,Mem)}(M_{glu})$ and fucose rate $C^r_{(host,Mem)}(M_{fse})$ from the host cells to the system that contains the memory of existing metabolites and evaluating the human GB compositional changes. The simulation for these experiments only contains growth dynamics of *Faecalibacterium*, *Eubacterium* and *Escherichia* bacteria as their growths are supported by the metabolites involved in the same SCFA production. Further, with the data availability, the model can be extended to analyse the growth dynamics of other bacterial genera as well.

Figure 6 illustrates the impact of glucose on the three bacterial populations based on Γ_{glu} ($\Gamma_{glu}\subseteq\Gamma_{SCFA}$), shown in Figure 6a. The colours used in Figures 6b and 6c follow the same colour scheme as in Figure 6a. Figures 6b and 6c shows the behaviours of edge weight and variation of population sizes as a fraction of that in S_0 due to the changes in $C^r_{(host,Mem)}(M_{glu})$ respectively. The variations of the input rate $C^r_{(host,Mem)}(M_{glu})$ alters the intermediate interaction from any bacterial population B_k to other population $B_{k'}$ through acetate, $C_{(B_k,B_{k'})}(M_{ace})$ and lactate $C_{(B_k,B_{k'})}(M_{lact})$, which are required for the growth of *Faecalibacterium* and *Eubacterium*, respectively. Figure 6b explains the graph theoretical behaviour of indirect influence on the growth dynamics of the respective bacterial populations. The growth of *Eubacterium* keeps increasing steadily until the $C^r_{(host,Mem)}(M_{glu})$ is twice the standard level, while the growths of the other two bacterial populations converge to the standard static level. This is due to the stoichiometry of the metabolite conversion, where an acetate molecule is produced by one glucose molecule while a lactate molecule requires two. The growth of *Escherichia* and *Faecalibacterium* are directly altered by the variations of glucose inputs and the behaviours. We calculated the maximum Mean Standard Error (MSE) as 0.03374 for any metabolite by iterating the experiment 20 times.

Deviations of a bacterial population concentration refer to deviations in node weights according to the (2) and (9). Figure 7 represents the node weight deviation compared to standard graph state S_0 due to the variability in inputs. This



(a) (b) (c)

Fig. 6: Deviation of population sizes of *Faecalibacterium*, *Eubacterium* and Escherichia from the standard levels due to different input concentrations of glucose: (a) subgraph for the glucose consumption, (b) edge weight behaviours of the intermediate interactions, and (c) population growth behaviours.

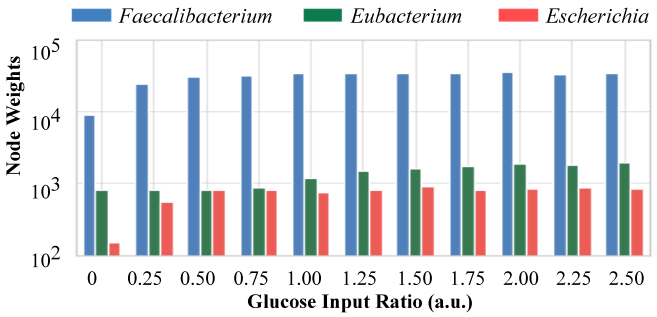


Fig. 7: Changes of node weights due to the variations in molecular signal inputs.

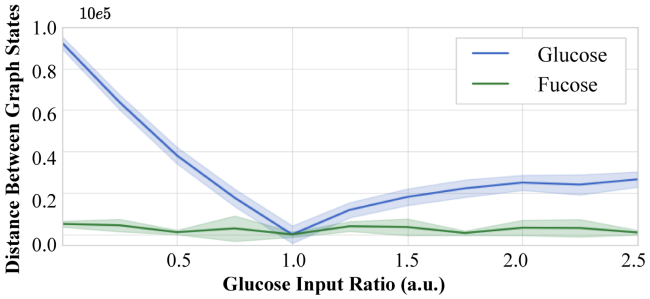


Fig. 8: Behaviours of overall graph weights against the changes in inputs and their concentrations.

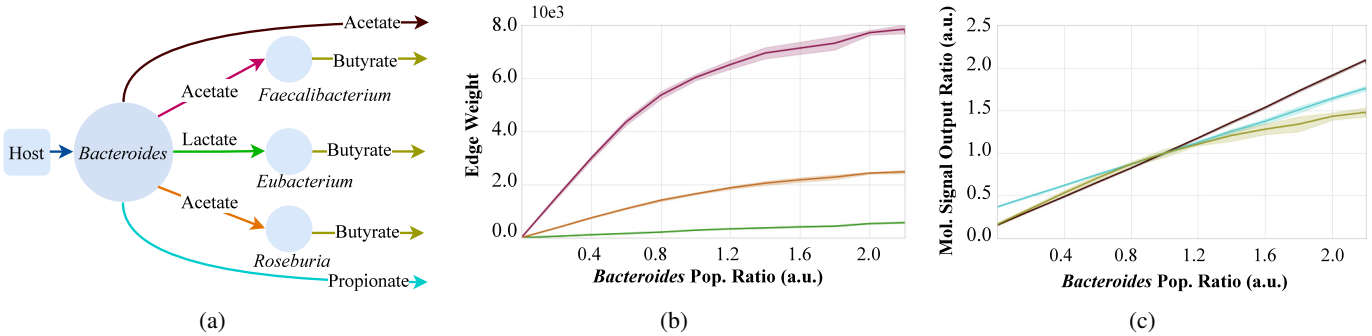
analysis reveals the impact of different input conditions on the molecular signal performance SPP of bacterial populations.

While Figure 7 explains the node weight variations, Figure 8 focuses on the overall interaction weight behaviours compared to S_0. This graph provides an insight into how the structure is being modified by the input variability. When the $C^r_{(host,Mem)}(M_{glu})$ is low compared to the standard level, the graph deviates significantly from the standard level, and when the $C^r_{(host,Mem)}(M_{glu})$ exceeds the standard level, the graph starts to deviate again from the standard structure, but with a lesser magnitude compared to a weaker signal (the standard level is 1.0). This reveals that the human GB is more sensitive to low glucose concentrations. The experiment is repeated for the fucose input rates $C^r_{(host,Mem)}(M_{fse})$ as well, but the impact is minimal compared to $C^r_{(host,Mem)}(M_{glu})$.

C. Analysis 2: Human GB structure effect on the graph outputs

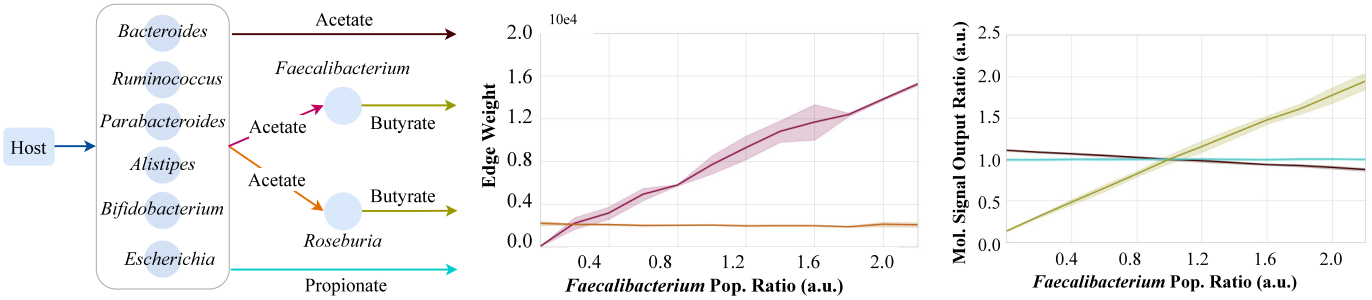
In this section, we analyse the direct and indirect impacts of the human GB compositional changes on the network behaviours. The analyses are conducted by altering the bacterial population sizes manually on the virtual GB and extracting the metabolite production data with respect to each alteration. The resulting behaviours of the MC layer are explained using the graph analyses. Although we conduct similar experiments for all the nine populations, we only show results on *Bacteroides* (Figure 9) and *Faecalibacterium* (Figure 10) populations as they provide a better understanding of the metabolite production dynamics of the human GB.

Figure 9 shows the impact of *Bacteroides* population size variation on the human GB SCFA production. In this experiment, we focus on the graph Γ_{Bct} ($\Gamma_{Bct} \subseteq \Gamma_{SCFA}$) considering only the interactions that are related to the *Bacteroides* population, as shown in Figure 9a. The colour scheme used in Figures 9b and 9c follow the same colour scheme as in Figure 9a. The metabolite inputs to the graph and the population sizes are maintained fixed at the standard level except for the *Bacteroides* population size. We modify the population size of *Bacteroides* ($|B_{Bct}|$) from zero cells to 2.2 times the standard population size. Figure 9b explains the behaviours of the intermediate links from *Bacteroides* to *Faecalibacterium* node through acetate, *Bacteroides* to *Eubacterium* populations through lactate, and *Bacteroides* to *Roseburia* populations through acetate, while Figure 9c shows SCFA production behaviours in the MC layer due to changes in the population size. From Figure 9c, it is evident that all the SCFAs have strong positive relationships with the population size of *Bacteroides*. Acetate and propionate are direct products of *Bacteroides* cells. As a result of that, acetate and propionate outputs show steady trends against the increase of *Bacteroides* population sizes. Moreover, the edge weight



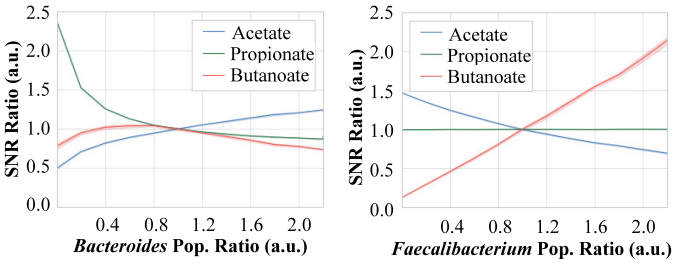
(a) (b) (c)

Fig. 9: Behaviours of SCFA production for various in *Bacteroides* population sizes: (a) subgraph of Bacteroides population interactions, (b) edge weight behaviours, and (c) SCFA output.



(a) (b) (c)

Fig. 10: Behaviours of SCFA production for various in *Faecalibacterium* population sizes: (a) subgraph related to the interactions of *Faecalibacterium* population, (b) edge weight behaviours, and (c) SCFA output.



(a) (b)

Fig. 11: Simulation results for SNR of three output signals with the changes in population sizes of *Bacteroides* and Feacalibacterium. (a) SNR results for *Bacteroides* population and (b) SNR results for Feacalibacterium population.

variations are shown in Figure 9b justify the butyrate signal behaviour in the MC layer shown in Figure 9c. To be more precise, the butyrate output curve starts to become flat when the *Bacteroides* population size $|B_{Bct}|$ is greater than 0.8 times the standard value. The graph theoretical quantification of links also shows the same trend in Figure 9b, emphasizing that the graph theoretical measures can be used to explain the metabolite production behaviours.

In the same way, Figure 10 illustrates the results for a similar experiment on *Faecalibacterium* population. Figures 10a, 10b and 10c represent the subgraph Γ_{Fae} ($\Gamma_{Fae} \subseteq \Gamma_{SCFA}$), edge and the MC layer behaviours, respectively. Similarly to the previous analysis, we modify the population size of *Faecalibacterium* $|B_{Fae}|$ ranging from zero cells to 2.2 times the standard population size. As the *Faecalibacterium* cells consume acetate and produce butyrate, the rate of acetate consumption from the environment increases when the $|B_{Fae}|$ is increased. Hence, the weight of interaction between environment and *Faecalibacterium* population increases, which can be observed in Figure 10b, and the resulting reduction in acetate output is visible in Figure 10c. Moreover, since *Faecalibacterium* population is one of the key butyrate producers, there is a clear positive relationship evident between $|B_{Fae}|$ and butyrate. Due to the smaller population size of the *Roseburia* population, the influence on the metabolite production is relatively low, which can be observed from Figure 10b. For all the graphs, the maximum MSEs are calculated below 0.03087.

The MC layer results presented for the two analyses on *Bacteroides* and *Faecalibacterium* populations (Figures 9c and 10c) are then interpreted in terms of SNR in Figure 11. In the plots of this figure, SNR values are shown as ratios of the SNR value at the standard state of the human GB, and the bacterial population sizes are increased similar to the previous analyses. Here, we show the three SNRs of acetate, propionate, and butyrate of two bacterial populations: *Bacteroides* and *Faecalibacterium*. Figure 11a shows the SNR behaviours of the three SCFAs against the $|B_{Bct}|$. It is clearly evident that the acetate production is higher compared to the other two SCFAs when the $|B_{Bct}|$ is increased. This means, when the composition of human GB is changed as the $|B_{Bct}|$

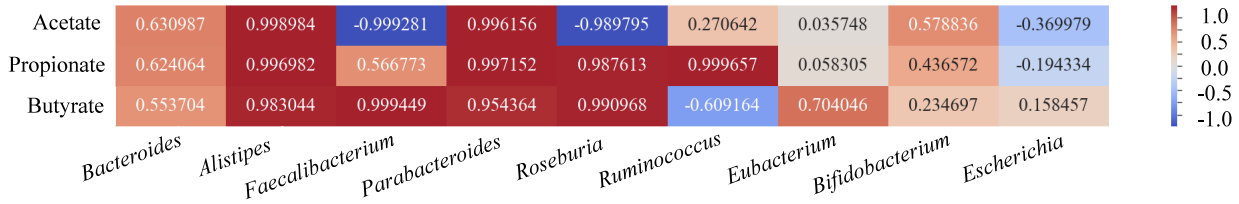


Fig. 12: Pearson correlation heat map of the impact on the three output signals by nine bacterial populations.

increases, the output of the GB also loses balance and tends to produce more acetate compared to the other two SCFAs. On the contrary, the propionate production rate reduces when the $|B_{Bct}|$ increases. When the population size of *Bacteroides* $|B_{Bct}|$ is smaller than the standard level, the system tends to produce molecular signal with higher deviated ratios, but when $|B_{Bct}|$ is greater than the standard level, the deviation is relatively low. Figure 11b shows the SNR behaviours of the three SCFAs against the $|B_{Fae}|$. Since *Faecalibacterium* is the main butyrate producer of this network, the butyrate SNR increases with the $|B_{Fae}|$ increment. Hence, compositional imbalance related to *Faecalibacterium* causes a significant imbalance in output molecular signal ratios. Furthermore, due to the acetate consumption of *Faecalibacterium*, the acetate signal becomes weaker, resulting in the acetate SNR deviating from the standard level.

Figure 12 explains the correlation between each bacterial population and the SCFA abundance in the gut environment. Although *Bacteroides* are the biggest producer of all the SCFAs, it has a weak correlation with SCFAs compared to other producers such as *Alistipes* and *Parabacteroides*. This reveals that the reduction of glucose consumption by *Bacteroides* increases the other bacterial population, resulting in boosted SCFA production. Note that, even the SCFA production of the other bacterial population is boosted in the absence of *Bacteroides*, the overall production is low. Since the *Faecalibacterium* and *Roseburia* consume acetate, the heat map shows a strong negative correlation with acetate. Interestingly, this heat map indicates metabolic switching by *Escherichia*, from a SCFA producer to a high acetate concentration consumer. This is the same for the *Ruminococcus* when the fucose concentration is not sufficient for the increased population, it switches from fucose to glucose consumption reducing the intermediate metabolite production, which causes a reduction in butyrate production.

VI. Conclusion

The gut bacteriome has been largely investigated due to its importance to the human health. We contribute to this research topic by introducing a two-layer GB interaction model to investigate the impacts of bacterial population compositional changes on the overall structure of the human GB utilising data collected from MicrobiomeDB and NJS16 databases. Our proposed human GB interaction model combines a bacterial population graph layer, which models the structure typically found in the human GB (i.e. bacterial populations genus and sizes), with a molecular communications layer, which models the exchange of metabolites by the bacterial populations in this structure. Supported by these models, we also developed a virtual GB to simulate the metabolic interactions that typically occurs in the human GB. These simulations allowed us to study the impacts caused by the metabolite exchanges on the human GB structure (i.e. nodes weight and hamming distance). Through our analyses, we found that the molecular inputs affect the bacterial populations in the human GB differently by modifying the nodes and edges weights of our GB interaction model. Our results also show that modifications in the human GB structure, specifically changing the sizes of *Bacteroides* and *Faecalibacterium* populations can lead to improvement/reduction in the production of SCFA, which may result in metabolic diseases in humans. Based on our results, we also infer that there is an intrinsic relationship between the investigated bacterial populations sizes, the increase/decrease of specific metabolites (SCFAs), and the overall balance of the human GB. These results can support the development of novel strategies to treat unbalanced human GB, and can provide insights on the role of other metabolites and molecules on the maintenance of a healthy gut bacteriome.

Acknowledgement

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) and the Department of Agriculture, Food and Marine on behalf of the Government of Ireland under Grant Number [16/RC/3835].

References

[1] L. K. Ursell, J. L. Metcalf, L. W. Parfrey, and R. Knight, "Defining the human microbiome," *Nutrition reviews*, vol. 70, no. suppl_1, pp. S38–S44, 2012.

[2] S. Balasubramaniam, N. Lyamin, D. Kleyko, M. Skurnik, A. Vinel, and Y. Koucheryavy, "Exploiting bacterial properties for multi-hop nanonetworks," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 184–191, 2014.

[3] S. Sanna, N. R. van Zuydam, A. Mahajan, A. Kurilshikov, A. V. Vila, U. Võsa, Z. Mujagic, A. A. Masclee, D. M. Jonkers, M. Oosting, *et al.*, "Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases," *Nature genetics*, vol. 51, no. 4, pp. 600–605, 2019.

[4] L. Wen and A. Duffy, "Factors influencing the gut microbiota, inflammation, and type 2 diabetes," *The Journal of nutrition*, vol. 147, no. 7, pp. 1468S–1475S, 2017.

[5] A. Iljazovic, U. Roy, E. J. Gálvez, T. R. Lesker, B. Zhao, A. Gronow, L. Amend, S. E. Will, J. D. Hofmann, M. C. Pils, *et al.*, "Perturbation of the gut microbiome by prevotella spp. enhances host susceptibility to mucosal inflammation," *Mucosal Immunology*, pp. 1–12, 2020.

[6] M. A. Henson and P. Phalak, "Microbiota dysbiosis in inflammatory bowel diseases: in silico investigation of the oxygen hypothesis," *BMC systems biology*, vol. 11, no. 1, p. 145, 2017.

[7] J. T. Bjerrum, Y. Wang, F. Hao, M. Coskun, C. Ludwig, U. Günther, and O. H. Nielsen, "Metabonomics of human fecal extracts characterize ulcerative colitis, crohn's disease and healthy individuals," *Metabolomics*, vol. 11, no. 1, pp. 122–133, 2015.

[8] S. M. Hill, L. M. Heiser, T. Cokelaer, M. Unger, N. K. Nesser, D. E. Carlin, Y. Zhang, A. Sokolov, E. O. Paull, C. K. Wong, *et al.*, "Inferring causal molecular networks: empirical assessment through a community-based effort," *Nature methods*, vol. 13, no. 4, pp. 310–318, 2016.

[9] V. K. Gupta, M. Kim, U. Bakshi, K. Y. Cunningham, J. M. Davis, K. N. Lazaridis, H. Nelson, N. Chia, and J. Sung, "A predictive index for health status using species-level gut microbiome profiling," *Nature communications*, vol. 11, no. 1, pp. 1–16, 2020.

[10] S. V. Lynch, S. C. Ng, F. Shanahan, and H. Tilg, "Translating the gut microbiome: ready for the clinic?," *Nature Reviews Gastroenterology & Hepatology*, vol. 16, no. 11, pp. 656–661, 2019.

[11] J. Yang, P. Zheng, Y. Li, J. Wu, X. Tan, J. Zhou, Z. Sun, X. Chen, G. Zhang, H. Zhang, *et al.*, "Landscapes of bacterial and metabolic signatures and their interaction in major depressive disorders," *Science advances*, vol. 6, no. 49, p. eaba8555, 2020.

[12] S. Kim, I. Thapa, L. Zhang, and H. Ali, "A novel graph theoretical approach for modeling microbiomes and inferring microbial ecological relationships," *BMC genomics*, vol. 20, no. 11, pp. 1–13, 2019.

[13] V. Petrov, S. Balasubramaniam, R. Lale, D. Moltchanov, Y. Koucheryavy, *et al.*, "Forward and reverse coding for chromosome transfer in bacterial nanonetworks," *Nano Communication Networks*, vol. 5, no. 1-2, pp. 15–24, 2014.

[14] I. F. Akyildiz, M. Pierobon, and S. Balasubramaniam, "An information theoretic framework to analyze molecular communication systems based on statistical mechanics," *Proceedings of the IEEE*, vol. 107, no. 7, pp. 1230–1255, 2019.

[15] I. F. Akyildiz, M. Pierobon, and S. Balasubramaniam, "Moving forward with molecular communication: From theory to human health applications [point of view]," *Proceedings of the IEEE*, vol. 107, no. 5, pp. 858–865, 2019.

[16] S. M. Jandhyala, R. Talukdar, C. Subramanyam, H. Vuyyuru, M. Sasikala, and D. N. Reddy, "Role of the normal gut microbiota," *World journal of gastroenterology: WJG*, vol. 21, no. 29, p. 8787, 2015.

[17] A. Visconti, C. I. Le Roy, F. Rosa, N. Rossi, T. C. Martin, R. P. Mohney, W. Li, E. de Rinaldis, J. T. Bell, J. C. Venter, *et al.*, "Interplay between the human gut microbiome and host metabolism," *Nature communications*, vol. 10, no. 1, pp. 1–10, 2019.

[18] L. V. Hooper and J. I. Gordon, "Commensal host-bacterial relationships in the gut," *Science*, vol. 292, no. 5519, pp. 1115–1118, 2001.

[19] M. Hasan, E. Hossain, S. Balasubramaniam, and Y. Koucheryavy, "Social behavior in bacterial nanonetworks: Challenges and opportunities," *IEEE Network*, vol. 29, no. 1, pp. 26–34, 2015.

[20] A. Naqvi, H. Rangwala, A. Keshavarzian, and P. Gillevet, "Network-based modeling of the human gut microbiome," *Chemistry & biodiversity*, vol. 7, no. 5, pp. 1040–1050, 2010.

[21] S. N. Steinway, M. B. Biggs, T. P. Loughran Jr, J. A. Papin, and R. Albert, "Inference of network dynamics and metabolic interactions in the gut microbiome," *PLoS computational biology*, vol. 11, no. 6, p. e1004338, 2015.

[22] R. K. Singh, H.-W. Chang, D. Yan, K. M. Lee, D. Ucmak, K. Wong, M. Abrouk, B. Farahnik, M. Nakamura, T. H. Zhu, *et al.*, "Influence of diet on the gut microbiome and implications for human health," *Journal of translational medicine*, vol. 15, no. 1, pp. 1–17, 2017.

[23] M. S. R. Rajoka, J. Shi, H. M. Mehwish, J. Zhu, Q. Li, D. Shao, Q. Huang, and H. Yang, "Interaction between diet composition and gut microbiota and its impact on gastrointestinal tract health," *Food Science and Human Wellness*, vol. 6, no. 3, pp. 121–130, 2017.

[24] B. D. Unluturk, S. Balasubramaniam, and I. F. Akyildiz, "The impact of social behavior on the attenuation and delay of bacterial nanonetworks," *IEEE transactions on nanobioscience*, vol. 15, no. 8, pp. 959–969, 2016.

[25] N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo, "A comprehensive survey of recent advancements in molecular communication," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1887–1919, 2016.

[26] Y. Chahibi and I. F. Akyildiz, "Molecular communication noise and capacity analysis for particulate drug delivery systems," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 3891–3903, 2014.

[27] M. Kuscu, E. Dinc, B. A. Bilgin, H. Ramezani, and O. B. Akan, "Transmitter and receiver architectures for molecular communications: A survey on physical design with modulation, coding, and detection techniques," *Proceedings of the IEEE*, vol. 107, no. 7, pp. 1302–1341, 2019.

[28] I. Llatser, A. Cabellos-Aparicio, M. Pierobon, and E. Alarcón, "Detection techniques for diffusion-based molecular communication," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 12, pp. 726–734, 2013.

[29] K. Oliphant and E. Allen-Vercoe, "Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health," *Microbiome*, vol. 7, no. 1, p. 91, 2019.

[30] D. P. Martins, M. T. Barros, and S. Balasubramaniam, "Using competing bacterial communication to disassemble biofilms," in *Proceedings of the 3rd ACM International Conference on Nanoscale Computing and Communication*, pp. 1–6, 2016.

[31] M. M. Deza and E. Deza, "Encyclopedia of distances," in *Encyclopedia of distances*, pp. 1–583, Springer, 2009.

[32] T. E. Gorochowski, A. Matyjaszkiewicz, T. Todd, N. Oak, K. Kowalska, S. Reid, K. T. Tsaneva-Atanasova, N. J. Savery, C. S. Grierson, and M. di Bernardo, "BSim: An agent-based tool for modeling bacterial populations in systems and synthetic biology," *PLoS ONE*, vol. 7, p. e42790, Aug. 2012.

[33] M. Kanehisa, S. Goto, and K.E.G.G., "Kyoto encyclopedia of genes and genomes," *Nucleic Acids Res*, vol. 28, p. 27–30, 2000. pubmed] [doi.

[34] M. Kanehisa, "Toward understanding the origin and evolution of cellular organisms," *Protein Sci*, vol. 28, p. 1947–1951, 2019. pubmed] [doi.

[35] M. Kanehisa, M. Furumichi, Y. Sato, M. Ishiguro-Watanabe, and M. Tanabe, "Kegg: integrating viruses and cellular organisms," *Nucleic Acids Res*, vol. 49, p. 545– 551, 2021. pubmed] [doi.

[36] J. Sung, S. Kim, J. J. T. Cabatbat, S. Jang, Y.-S. Jin, G. Y. Jung, N. Chia, and P.-J. Kim, "Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis," *Nature communications*, vol. 8, no. 1, pp. 1–12, 2017.

[37] R. Caspi, T. Altman, R. Billington, K. Dreher, H. Foerster, C. A. Fulcher, T. A. Holland, I. M. Keseler, A. Kothari, A. Kubo, *et al.*, "The metacyc database of metabolic pathways and enzymes and the biocyc collection of pathway/genome databases," *Nucleic acids research*, vol. 42, no. D1, pp. D459–D471, 2014.

[38] C. Huttenhower, D. Gevers, R. Knight, S. Abubucker, J. H. Badger, A. T. Chinwalla, H. H. Creasy, A. M. Earl, M. G. FitzGerald, R. S. Fulton, *et al.*, "Hmp phase i (v3-v5)," 2012.

[39] A. O'Callaghan and D. van Sinderen, "Bifidobacteria and their role as members of the human gut microbiota," *Frontiers in microbiology*, vol. 7, p. 925, 2016.

[40] I. Rowland, G. Gibson, A. Heinken, K. Scott, J. Swann, I. Thiele, and K. Tuohy, "Gut microbiota functions: metabolism of nutrients and other food components," *European journal of nutrition*, vol. 57, no. 1, pp. 1–24, 2018.

[41] G. Den Besten, K. Van Eunen, A. K. Groen, K. Venema, D.-J. Reijngoud, and B. M. Bakker, "The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism," *Journal of lipid research*, vol. 54, no. 9, pp. 2325–2340, 2013.