Genome scale metabolic modelling of human gut microbes to inform rational community design

Juan Pablo Molina Ortiz [a b] * jp.molinaortiz@csiro.au

Dale David McClure [c] dale.mcclure@brunel.ac.uk

Andrew Holmes [d e] andrew.holmes@sydney.edu.au

Scott Alan Rice [b f] scott.rice@csiro.au

Mark Norman Read [d g] mark.norman.read@gmail.com

Erin Rose Shanahan [d e] * erin.shanahan@sydney.edu.au

a. CSIRO Environment, Canberra, ACT 2601, Australia

b. CSIRO MOSH- Future Science Platform, Australia

c. Department of Chemical Engineering, College of Engineering, Design and Physical Sciences, Brunel University, London, UK

d. Charles Perkins Centre, the University of Sydney, Sydney, Australia

e. School of Life and Environmental Sciences, Faculty of Science, the University of Sydney, Sydney, Australia

f. CSIRO Agriculture and Food, Westmead, NSW 2145, Australia

g. School of Computer Science, Faculty of Engineering, the University of Sydney, Sydney, Australia

* Corresponding authors: jp.molinaortiz@csiro.au, erin.shanahan@sydney.edu.au

Abstract

Background: The human gut microbiome plays a pivotal role in health and disease, influenced significantly by diet, particularly through the intake of digestion-resistant carbohydrates (DRCs). Emerging evidence underscores the potential of DRC supplementation in modulating the gut microbiome towards health-promoting metabolic outputs, notably through the fermentation of DRCs into short-chain fatty acids (SCFAs) including butyrate. However, the effectiveness of such interventions is hindered by the inherent complexity of microbial communities and the variable functional capacity of microbiomes across individuals. Improved understanding of gut ecology is necessary to move past interventions with transient benefits.

Results: This study leverages genome-scale metabolic models (GEMs) to characterise the metabolic capabilities of 598 stable human gut colonising strains from the AGORA database. We infer the strains' abilities to utilize dietary carbohydrates of varying complexities and produce metabolites that mediate interactions with other microbes and the host.

Our analysis reveals a stratified functional landscape where prominent bacterial families show predispositions towards primary or secondary degrader roles based on their carbohydrate utilisation capabilities. Further, we identify metabolite production profiles that exceed phylogenetic variation in our sample. These results offer a comprehensive functional mapping of carbohydrate metabolism across a wide array of gut microbes, shedding light on the complex trophic networks underpinning the gut ecosystem.

Supporting DRC-based interventions with rationally designed microbial communities can better guarantee the delivery of the intended health-promoting metabolic outputs. Applying our functional assessment, and principles of reverse ecology and network analysis, we propose a novel framework for the rational design gut microbial communities, where trophic networks are optimised to produce target metabolites from selected DRCs. Our results further suggest that this framework can predict resilient minimal communities, an important trait in the constantly changing human gut nutritional environment.

Conclusion: Our work provides novel insights into gut microbial ecology as well as intervention and consortia design. The identified metabolic capabilities of individual strains inform the rational design of a purpose-based microbial community to optimise butyrate production from inulin degradation. The framework we propose herein sets a foundation for future efforts aimed at the rational design of interventions that target the human gut microbiome to improve health.

Keywords: gut microbiome, human, reverse ecology, intervention design, community design, consortia design, consortia engineering, metabolic network modelling, genome scale metabolic modelling, GEMs.

Background

Mounting evidence consistently links human gut microbiome composition and function with health outcomes (1-5). Scientific evidence further identifies diet composition, particularly digestion-resistant carbohydrates (DRC) content, as a strong gut microbiome modulator (6-11). Broad principles have been established in this regard. A diet high in DRCs has been associated with health promoting

metabolic outputs from the microbial community (6, 12-16), including fermentation into short chain fatty acids (SCFAs), particularly butyrate, which improve colonocyte fitness. Meanwhile, microbiomes with limited access to DRC may overproduce compounds associated with gut inflammation (e.g., hydrogen sulphide) (17-21). Therefore, an opportunity exists to intervene at the gut microbiome level, through DRC supplementation, aiming to improve health (22-26).

The hierarchical structure and dynamic metabolic roles within the human gut microbiome present significant challenges for designing lasting interventions. Another broadly accepted principle governing the human gut microbiome is the existence of a hierarchical order, where primary degraders (e.g., DRC degraders) enable access to the energy of the more complex compounds by producing structurally simpler carbohydrates and organic acids (e.g., SCFAs) to microbes that otherwise lack the required metabolic capabilities (secondary degraders) (27-29). Importantly, short term changes in substrate availability, due to dietary content and circadian cycle, may lead to niche adaptation (18), and thus a microbe's role as a primary or secondary degrader may change according to substrate availability at a given point in time. Cross-feeding at various levels of carbohydrate complexity leads to the emergence of complex trophic networks with non-linear dynamics. Characteristics common to the gut environment, including both high rates of substrate availability and intense competition for resources, favour partial degradation of nutrients and accumulation of intermediate metabolites, such as formate, which can then be cross-fed to alternative populations in the community (30-32). Additionally, the capabilities to interact with specific compounds may vary between individuals' gut microbiomes (33). These factors collectively contribute to different degrees of intervention response (10, 34-37), complicating the design of dietary- or nutrient-based interventions targeting the gut microbial community. Supplementation of microbial strains (probiotics), with known metabolic capabilities, has been proposed to address such variability (38-40). Yet, most of these efforts result in comparatively transient improvements to host health, if any (41-44). For example, it has been suggested that short-term colonisation of probiotics in the gut is linked to a limited integration to the existing microbiome (45-47), and eventual extinction due to intense competition (47) and potentially incompatible auxotrophies (48).

The delivery of interventions that rationally combine DRC and probiotics/microbial consortia have the potential to improve outcomes. The design of a purpose-based microbial community, where members: a) have been hierarchically and functionally selected to produce beneficial metabolites from a selected DRC; while b) guaranteeing community persistence in the gut through cross-feeding interactions, would constitute a more robust approach (33, 49). However, rational community design through conventional techniques usually requires iterative design and testing, which demands substantial resource investment in terms of time, personnel and laboratory equipment for example. Further, limited understanding of the human gut microbiome, particularly of the complex trophic networks that emerge from DRC-degradation (carbohydrate dimension) (50-52), complicates exploration of this combinatorial challenge.

Reverse ecology has recently emerged as a bottom-up approach which, in contrast to traditional ecological studies, aims to predict ecological roles from multi-omics data (53-56). Genome scale metabolic models (GEMs) are computational tools built from multi-omics and experimental data, for the reconciliation of genetic traits at the cellular level (57-59). Through the inspection of fluxes and the use of metabolic constraints, these computational models allow inference of functional traits, often at the strain level, as they are integrated with a cellular metabolic network (60-64). This offers a more nuanced assessment of metabolic capabilities encoded in strains compared to common omics approaches. Strain-level GEMs can further be modelled in consortia to predict nutritional interactions among members, therefore serving as a reverse ecology framework and enabling the rational design of purpose-based microbial communities (21, 64-68).

Here, we first use GEMs to functionally characterise a group of 816 bacteria and archaea strains from AGORA (assembly of gut organisms through reconstruction and analysis) (69) in terms of their capabilities to a) consume carbohydrates from various levels of structural complexity; and b) produce metabolites that can be cross-fed. The resulting functional data enables us to promote understanding of human gastrointestinal tract (GIT) microbiome traits through the application of reverse ecology. We additionally propose a framework to develop rationally designed interventions through purpose-based communities capable of optimising the production of health promoting metabolites from select DRC

supplementation. Such frameworks have the potential to transform the way we approach gut microbiome intervention design and hence, improve the delivery of health outcomes.

Methods

Computational resources

The GEMs employed for this project were obtained from www.vmh.life (70), and are part of the *without mucins* version of AGORA (Assembly of Gut Organisms through Reconstruction and Analysis) (69). This was selected based on our aim of focusing on dietary-derived molecules (with the 1423 O-glycans represented in the standard (*with mucins*) version being out of the scope of this body of work). GEMs of strains *Clostridium sporogenes* ATCC 15579 and *Lactobacillus helveticus* DPC 4571 were excluded from this analysis due to inconsistencies, as outlined previously (63).

Metabolite flux for each GEM was modelled using Flux Balance Analysis (FBA) and the constraint-based reconstruction and analysis (COBRA) toolbox for the Python coding language (COBRApy) (71). FBA calculates growth in terms of biomass generation in grams of dry weight per hour (gDW/h) based on nutrient fluxes. COBRApy further allows a) inspection of individual metabolic reactions and metabolite consumption and synthesis and b) seamless addition or removal of specific molecules to growth media. Modelling of GSMs presented in this manuscript were performed on a personal computer.

Community modelling was performed using MICOM (65), a python package for modelling individual GEMs in a community context. Fluxes were derived from utilising the cooperative_tradeoff method, which optimises the community metabolic network for maximum community and individual member growth. Parsimonious FBA (pFBA) was used to model community fluxes, a variation of traditional FBA, which aims to optimise the metabolic network, while minimising the total sum of fluxes that contribute to a solution, reducing the solution space, which is highly desirable during community modelling. COBRApy and MICOM simulations were performed with an academic license for the Gurobi solver (72). Custom Python scripts were developed, in addition to those that form part of GEMNAST, for data analysis and figure plotting.

Expanding GEMNAST for nutrient utilisation inference

Previously, we introduced GEMNAST (GEMs Metabolic and Nutritional Assessment), a Python-based algorithm that samples metabolic capabilities of a predetermined set of GEMs qualitatively, based on combinatorial media modifications (63). GEMNAST was devised as a high-throughput pipeline to identify auxotrophies and prototrophies in GEMs. For this study, GEMNAST was extended to assess the following on individual GEMs:

Nutrient utilisation capabilities

A COBRApy-based python script was developed to identify molecules individual GEMs can utilise as nutrients. The implementation requires three main inputs: a) a directory to load GEMs from; b) base media ingredients; c) a list of molecules which *should not* be part of the original base media. Both inputs b and c should be listed in the form of compatible exchange reactions, as codified in the GEMs from input a. Using the provided base media, the pipeline generates culture media containing the first molecule specified in input c. A strain is cultured (modelled) in the resulting media and is constrained to utilise the molecule, if possible. If FBA results report that the strain utilises the tested molecule *and* grows above a predetermined growth rate threshold, a positive outcome (1) is assigned to the experiment; both conditions must be true (Figure S1). These conditions work as safeguards, guaranteeing that positive outcomes are not due to potential confounders in base media. If a strain cannot be constrained to utilise the molecule, meaning it does not encode the required reaction(s), a negative outcome (0) is assigned to the experiment. This cycle is repeated with every molecule of interest, starting from base media, and when completed, the analysis resumes with the next strain provided in input a. Individual outcomes are arranged in a Boolean table with strain names as rows and molecule names as columns.

Metabolite export capabilities following nutrient utilisation

A pipeline to determine metabolites that can be exported from the utilisation of molecules was developed. It would be impractical to assess the output of every metabolite generated following FBA; therefore, we designed a script that enables a predetermined set of compounds to be assessed. This script requires four inputs. Inputs a, b and c follow the same requirements as those described in the Nutrient utilisation capabilities section. In addition to these, a list of metabolites (input d) that can potentially be produced from the digestion of the nutrients from input c should be provided and *should not* be part of base media (input b). Similar to the above-described pipeline, this script adds one of the molecules from input c to base media and the strain is constrained to utilise it, if possible. In this instance, however, FBA outcomes should fulfil three conditions for the experiment to be assigned a positive outcome (1): the strain should grow above a predefined growth threshold, it should have utilised the tested molecule *and* the first metabolite from input d should have been exported as a waste product. Otherwise, a negative outcome (0) is assigned to the current iteration. This process is repeated with every metabolite from input d, then with every molecule from input c and every GEM in the directory provided as part of input a (Figure S2); to determine if strains have the capability of exporting a given metabolite, they are constrained to do so. Outcomes are organised in a Boolean table with strain names as rows and cultures, labelled using the following format: name of molecule to be degraded, name of metabolite to be exported, as columns. It is important to highlight that this analysis cannot guarantee that the assessed metabolites are directly generated from the utilisation of specific molecules as nutrients, as the script does not trace catabolic pathways within the cell. Instead, this script assesses a strain's capability to generate a metabolite while utilising the assessed molecule and achieving a specified growth rate.

Modelling growth media

The previously reported Universally Defined Media (UDM, Table S1) (63), which was engineered to guarantee growth of every GEM in AGORA, was modified for the purposes of this study. Molecules identified as potential confounders (i.e., simple sugars) were removed from the formulation

(carbohydrate-free UDM, cfUDM, Table S1) to allow determination of carbohydrate degradation capabilities in individual GEMs.

Meaningful growth

To determine a strain's capacity to grow in media an extremely conservative threshold of 0.09 h^{-1} growth rate (eight hours doubling time) was established on the basis that any microbe growing at a lower rate would present relatively low probabilities of surviving in the human gut, where transit times average just over 24 hours, but can be shorter than 14 hours in some individuals (73). This threshold represents a theoretical minimum growth rate for gut survival as less than two replication cycles would not guarantee permanence in the human colon (74).

Carbohydrate utilisation and metabolite export experimental design

Carbohydrates and metabolites for our analysis were selected based on a compilation of relevant carbon sources reported in the Virtual Metabolic Human (VMH) database (70), where AGORA GEMs are housed (69) and supported in the literature with respect to DRC, related dietary molecules and potential metabolic products. A list of the 66 selected compounds, divided into four categories based on structure complexity, are shown in Table S2.

Following our customised GEMNAST script, we sampled the carbohydrate utilisation capabilities of 816 strains represented in AGORA (input a) based on all the molecules in Table S2 (input c) and cfUDM as base media (input b). Each molecule was individually added to cfUDM, and molecule utilisation and growth were assessed (Figure S3). Results from this analysis are shown in Supplementary file 2.1, 2.2 and 2.3 for monosaccharides, oligosaccharides and polysaccharides, respectively.

The GEMNAST script we developed to assess metabolite export was applied to identify which of the selected metabolites (input d) AGORA strains (input a) can export as waste while utilising each

monosaccharide from Table S2 (input c), with cfUDM (input b) as base media (Figure S4). Results of this part of our analysis can be found in Supplementary file 2.4.

Classification of AGORA strains into topo-physiological categories

AGORA strains were divided into groupings (Supplementary file 2.5) based on their relationship with the host (pathogens, pathobionts) and their more prevalent anatomical site (upper GIT, large bowel, skin). This process was done manually with reference to the literature and relevant databases (75-78).

Phylogeny assignment

Taxonomy IDs of 590 large bowel AGORA strains were retrieved from the Virtual Metabolic Human database (70) and were used to assemble a .phy file using NCBI's online taxonomy browser (79) (Supplementary file 4) to retrieve updated phylogenetic details for such strains and families relevant for our analyses.

t-Distributed stochastic neighbour embedding (t-SNE) and Kernel density estimation (KDE)

t-SNE and KDE were applied to the Boolean tables that resulted from the analysis of each category of molecules to visualise how strains from prevalent large bowel families are distributed in the space, based on their individual carbohydrate degradation/metabolite production profiles. The TSNE scikit-learn Python module was used for this purpose. We utilised the default Euclidean distance and exact t-SNE, which optimises the Kullback-Leibler divergence of distributions between original and embedded spaces, to measure the dissimilarity between datapoints (80). KDE, from the SciPy, was used to estimate the number of overlapping strains in on a t-SNE plot (81).

Heatmaps and Boxplots for results visualisation

Seaborn (82), a Python library for data visualization, based on matplotlib (83), was used to generate heatmaps and boxplots derived from the carbohydrate utilisation and metabolite production datasets. Matplotlib was used to generate stacked bar charts (83).

Percent change in butyrate production rates

To report the difference in butyrate production rates between two modelled communities or different nutritional contexts we rely on Equation 1 to calculate percent changes.

Equation 1 Percentage Change

$$PC = \frac{CV - OV}{OV} \times 100$$

Where PC is the percent change, OV is the original value; in this case, the predicted rate for butyrate production in our purpose-based 6-strain community in cfUDM supplemented with inulin. CV is the compared value, which corresponds to alternative butyrate production rates. Absolute predicted values can be found in the corresponding sections of Supplementary file 2 as part of community modelling fluxes.

Results

In silico culturing with a stratified carbohydrate dimension

DRCs, commonly termed dietary fibre, are not digested in the small intestine and therefore represent a key carbon and energy source for microbes in the large bowel. Comprehensive *in silico* characterisation of the functional traits of AGORA strains in the carbohydrate nutritional dimension requires equally comprehensive representation of this group of molecules, as well as careful

consideration of the anabolic and catabolic activities from which trophic networks arise. To adequately represent the structural complexity within carbohydrates, relevant to the human gut microbiota, we designated three categories based on molecule structure:

- Polysaccharides (25 compounds): carbohydrates with over ten monosaccharides in their structure.
- Oligosaccharides (14 compounds): molecules composed of two to ten monosaccharides.
- Monosaccharides (16 compounds): single sugar units that cannot be further broken down by hydrolysis.

Each of these categories was populated with compounds from a list of nutrient requirements used for the curation of AGORA GEMs (70). This resulted in 55 carbohydrates (Table S2) and an equal number of *in silico* media formulations, based on a carbohydrate-free universally defined growth media (cfUDM, Methods and Table S1) supplemented with each carbohydrate.

To characterise the functional traits from which carbohydrate trophic networks arise, we extended our GEMNAST pipeline (63), which facilitates inference of metabolic traits encoded in GEMs (Methods). To determine if a strain is capable of consuming a given molecule, its GEM counterpart should predict that the strain can grow above a predetermined threshold while utilising the molecule. Using this pipeline, each AGORA strain was modelled once in each media to explore carbohydrate utilisation capabilities, totalling 53,040 *in silico* cultures (Supplementary files 2.1 to 2.3).

Genome-scale metabolic modelling also allows inference of metabolites, products of microbial metabolism that can accumulate in the environment from nutrient foraging. Leveraging this functionality, we aimed to identify intermediate and terminal metabolites that can be produced when large bowel strains utilise specific monosaccharides due to their importance as cross-feeding currency and means of interaction with the host. The literature suggests that both terminal metabolites (e.g. SCFAs) and even intermediate metabolites can drive metabolic interactions in the human gut (32) and

other microbial systems (84). To obtain a more comprehensive trophic network perspective, and given that different metabolic strategies may lead to the production of alternative metabolites (32, 35, 85), we relied on modelling constraints to 'induce' individual GEMs to export a given metabolite while utilising a monosaccharide, if metabolically feasible. A total of 11 metabolites (four intermediate and seven terminal, Table S2) were selected for this analysis based on AGORA curation data and literature reports (69), which resulted in 105,248 *in silico* cultures (Supplementary file 2.4).

Metabolic capabilities of large bowel microbiota

The focus of our analysis is the human large intestinal microbial community, as this is a key target for interventions. Although AGORA comprises GEMs of strains that have been identified in, and/or can inhabit the human large intestine, we focused our analysis on taxa that could be considered stable colonisers of the large bowel, excluding for example, oral taxa and known pathogens (Supplementary file 2.5). Comparison of the carbohydrate utilisation capabilities between large bowel strains and other groups of strains from AGORA can be found in Table S3 and Figure S5. We analysed the distribution of carbohydrate utilisation capabilities in strains from the large bowel to better understand their predisposition to different ecological roles. At a broad level we found that carbohydrate utilisation capabilities are increasingly rare at higher levels of nutritional complexity. For large bowel strains, 173 out of 598 could utilise one or more polysaccharides, a proportion that progressively increased for less complex groups of molecules (Figure 1A). This suggests that a larger portion of large bowel microbiota are predisposed to occupy secondary degrader roles, and that fewer strains possess the functional traits needed for a primary degrader role.

Figure 1 Carbohydrate utilisation capabilites and metabolite production profiles of stable large bowel colonisers in AGORA. A) Number of strains capable of utilising at least one carbohydrate from the given structural category. A lower number of strains are inferred to utilise molecules within a group as structural complexity increases. B-C) Linear regression and Pearson correlation coefficient (PCC) between carbohydrate utilisation capablilities of strains in terms of two structural categories. Larger datapoints show a higher number of overlapping strains. D) Percentage of large bowel strains capable of producing a

given metabolite based on monosaccharide availability. E) Metabolite production profiles shared by five strains or more. Number of members of the seven representative families of large bowel strains in AGORA is also shown. Number of produced metabolites for a given group of strains (stacked bar) are shown as x-axis labels.

Weak, yet significant positive correlations were identified between the number of monosaccharides and the number of more complex carbohydrates a strain can utilise (Figure 1B and C), which is biologically coherent. Polysaccharide/oligosaccharide breakdown demands coordinated synthesis, and potentially export, of often more than one type of specialised enzymes, a process that *should* guarantee that the involved organism(s) has access to simpler molecules it can then utilise as nutrients (86).

A second part of our analysis focuses on the metabolic profiles large bowel strains can produce when utilising the range of monosaccharides considered in our assessment. Our results showed that the majority of large bowel strains (542 of 598 tested) can adopt metabolic states that allow them to grow above a minimum threshold while producing and exporting at least one of the 11 metabolites. Of these, only 12 strains reported a metabolite production profile that varied for different monosaccharides, whilst the vast majority (530 strains) presented a consistent metabolite production profile across the tested monosaccharides (Supplementary file 2.6). We identified 129 unique metabolic combinations in total, far exceeding the phylogenetic variation represented in AGORA strains (69). Many metabolic profiles were unique to a particular strain, and only 29 metabolite production profiles were shared among at least five strains. The top five largest groups of strains sharing the same production profile ranged from 48 to 20 members that could produce four to nine of the examined metabolites (Figure 1E). Interestingly, there was substantial overlap between the metabolic profiles of members of the *Bacteroidaceae* and *Prevotellaceae* families. Our analysis predicts that on average 316 strains can produce each of the tested metabolites, however only 82 of the large bowel strains tested can produce butyrate (Figure 1D), considered a key SCFA for intestinal health (87-89), in the tested environment..

Carbohydrate utilisation capabilities suggest evolutionary trade-offs among prevalent gut taxa

Delineation of the carbohydrate degradation capabilities among prevalent human gut taxa is a necessary foundation for rational community design. Figure 2 summarises utilisation capabilities (by number/percentage of carbohydrates in a category) among strains from seven representative families both in the human gut and stable large bowel colonisers in AGORA. Based on these, three patterns were observed. First, we find families where their utilisation potential is comparatively skewed towards more complex carbohydrate categories. These are *Bacteroidaceae*, *Bifidobacteriaceae* and *Prevotellaceae*. *Bifidobacteriaceae* and *Prevotellaceae* present a high polysaccharide utilisation median percentage compared to the rest of the families examined, while *Bacteroidaceae* presents the largest range in the group with several strains with a high polysaccharide utilisation capability (Figure 2A). *Bifidobacteriaceae* also stand out as prominent oligosaccharide utilisers, compared to the rest of the families. The opposite occurs with the *Clostridiaceae* and *Enterobacteriaceae*, which display higher compatibility for simpler carbon sources, compared to other families. This pattern is particularly evident in the *Enterobacteriaceae* family, which displays a notably low average for polysaccharide utilisation and a comparatively high monosaccharide utilisation average. Finally, the *Lachnospiraceae* and *Oscilospiraceae* display were broadly able to use a range of oligosaccharides and monosaccharides. Overall, these findings suggest that a tendency exists at the family level that may influence the ecological roles of strains in terms of carbohydrate utilisation.

Figure 2 Carbohydrate utilisation profiles among prominent human gut microbiome families represented in AGORA. A) Boxplot showing distribution of the percentage of nutrients from each structural category utilised by a bacteria family. Family median is shown in bold. B-E) t-SNE analysis (Euclidean distance) based on individual strain metabolite production/carbohydrate utilisation profile (B: Metabolite production, C: Monosaccharides, D: Oligosaccharides, E: Polysaccharides) suggests more conserved metabolic capabilities at the family level in terms of metabolite production profiles. Datapoint sizes correspond to the number of strains with an overlapping location on a plot, and therefore, and identical metabolic profile. Only strains that were inferred to utilise/produce one or more molecules in a group are included.

While family-level patters were identified, it is important to highlight that clear variability exists at the strain level (Figure 2B-E). We find that metabolite production profiles are relatively similar within members of the *Bacteroidaceae*, *Prevotellaceae*, *Bifidobacteriaceae* and *Enterobacteriaceae*, which occupy analogous plot locations in respect to other family members (Figure 2B). However, as nutritional complexity increases degradation patterns among family members show less consistency.

Design of an inulin-degrading butyrate-producing community

Systematic characterisation of the carbohydrate degrading capabilities within large bowel strains shows that community design for the delivery of rational interventions requires a hierarchical approach as the production of desirable outcomes can be disrupted at various trophic levels (27, 86, 90). Applying the functional data that we have generated here, we follow a process to rationally design an intervention – one that revolves around a purpose-based community that can be sustained by, and produce target metabolites from, a selected DRC. In addition to being purpose-based, communities at the centre of the interventions we aim to design should be a) feasible – with a realistic number of members; b) resilient – members should be capable of covering others auxotrophies through cross-feeding during changing nutritional conditions in the gut; and c) function-guided – members should actively contribute to the target metabolite production process, while avoiding detrimental outputs (e.g. hydrogen sulphide production). We define feasible communities as those with 4 up to 33 members, according to existing defined microbial communities in the context of human gut microbiome interventions (91). Based on these characteristics, we aim to provide proof-of-principle for the design of a community of gut commensals capable of optimising the production of butyrate from the degradation of inulin (Figure 3A). Inulin is a commonly used and broadly studied polysaccharide (85, 92, 93), while butyrate has been extensively described as a metabolite with a positive impact to gut health (87, 94).

Figure 3 Steps for the design of an inulin-degrading, butyrate producing microbial community. We rely on reverse ecology-derived functional data (step 1 and diagram A), microbial community modelling (steps 2 and 4) and network analysis (step 3 and diagram B) to identify ideal candidates for the implementation of a rational intervention.

We start the consortia design process by using the data generated in the analyses above, identifying every large bowel strain represented in AGORA also capable of contributing to the process of converting inulin into butyrate (Figure 3 step 1 and diagram A; function-guided selection). Based on the metabolic reactions AGORA GEMs can perform, inulin is broken into 29 units of fructose and one unit of glucose or into 25 units of fructose and one unit of kestopentaose, a fructose-oligosaccharide (FOS), which when degraded produces four units of fructose and one unit of glucose. Therefore, we initially selected strains that were identified as inulin and kestopentaose degraders. Additionally, strains capable of generating butyrate from the utilisation of glucose and/or fructose were included. This resulted in an initial selection of 151 strains (Table 1). In doing so, a total of seven different phenotypes were identified, from specialists (inulin or FOS degraders, butyrate producers) to generalists (those that can contribute at every level of the network).

Table 1 Predicted functional roles in members of a function-guided 151-strain community. Functional characterisation of AGORA strains revealed that 151 strains presented desirable traits to implement our hypothetical intervention. Conversely, community modelling of this group of strains identified members that displayed adverse functional traits in the tested environmental conditions: butyrate consumption (instead of production) and hydrogen sulphide (H_2S) production.

Phenotype groups	**No of strains**	**Butyrate consumers**	**Produce H_2S**
Butyrate producers	60	6	11
FOS degraders	24	0	4
Inulin degraders	33	2	8
FOS degraders / butyrate producers	8	1	3
Inulin degraders / butyrate producers	6	0	1
FOS-inulin degraders	16	1	5
Inulin-FOS degraders / butyrate producers (generalists)	4	0	0

To continue with the community refinement process we modelled the community resulting from the 151 selected strains (Figure 3, step 2, Supplementary file 2.7), utilising MICOM, a Python package

for the modelling of GEMs as microbial communities (65). The media utilised for modelling was our universal growth media (cfUDM) supplemented with inulin. While predictions showed that every strain grew in this context, results revealed that 10 strains consumed butyrate and that 32 produced hydrogen sulphide (H_2S, Table 1), a metabolite associated with detrimental impacts on gut health (20, 21, 95).

The implementation of a 151-strain community is not feasible by current standards, and it is likely that its complexity will produce undesired, hard to predict, outcomes. Therefore, we relied on a network analysis, based on the nutritional exchanges predicted in the modelling of such a community, to identify strains that can better guarantee a resilient community (Supplementary file 2.8). We calculated total in- and out-degree edges (shared metabolites either imported or exported) within the network for each strain in the 151-strain community. The number of interactions was then normalised by the number of strains in the network resulting in an average of 4 imports and 4 exports per strain. We theorise that strains that are highly interconnected in the 151-strain community are hub strains, crucial for community resilience in the tested nutritional context (cfUDM + inulin) (96). We additionally aimed to identify strains with a balanced degree coefficient (exports/imports $\approx$ 1) to prevent the addition of '*cheaters*', strains that take advantage of nutrients available in the media but do not reciprocate accordingly (high imports/low exports). This is supported by the literature, which reports that stable communities display balance between auxotrophies and prototrophies between members, which translates into balanced competitive and collaborative interactions (97-99). Based on these aims, we devised criteria to: a) identify hub strains, those with a high imports and exports sum (> 6.7, fourth decile); and b) strains with a balanced degree coefficient (exports/imports = 0.9-1.1) to account for auxotrophy/prototrophy balance (Figure 3, step 3 and diagram B; resilience selection).

Primary degraders occupy a privileged and central niche in the community and can drastically shape the nutritional context (86, 90, 99-101). Above, we identified that *Bacteroidaceae*, *Bifidobacteriaceae* and *Prevotellaceae* families display traits that make them prone to occupy primary degrader roles compared to other prevalent gut families. This was indeed the case in our 151-strain inulin-to-butyrate community (Supplementary file 2.7), where every member of these families occupied niches based on

inulin and/or FOS degradation. To better understand how these strains influence the broader nutritional context, we focused on their network analysis scores. Notably, *Bacteroidaceae* strains consistently showed a degree coefficient with a higher number of imports than exports, while *Bifidobacteriaceae* and *Prevotellaceae* displayed the opposite (Table 2). We further analysed fluxes from these families in the amino acid dimension in the 151-strain MICOM simulation. We found that there were only two instances were *Bacteroidaceae* exported amino acids to the media, while *Bifidobacteriaceae* and *Prevotellaceae* contributed with a median of 4 and 2.5 amino acids per strain, respectively (Table 2). Strains that export a greater number of amino acids have a greater potential to support other members in a microbial consortium. Interestingly, we also identified nine of the 13 members of the *Bacteroidaceae* in this consortium as H_2S producers, whereas no members of the other two families presented such trait.

Table 2 Members of three prevalent bacteria families in the human gut consistently displayed inulin-FOS degradation roles in a function-guided community of 151 strains. Degree coefficient (exports/imports) and amino acid uptake/export profiles were assessed, showing that members of the Bacteroidaceae *adopt a more "selfish" profile (consistent tendency for importing a higher number of molecules from the media than exporting them; and similar amino acid uptake but low export median) compared to the other two families.*

Family (number of strains)	**Degree coefficient (exports/imports)**		**Amino acids imported**		**Amino acids exported**	
	Range	**Average**	**Range**	**Median**	**Range**	**Median**
Bacteroidaceae (13)	0.3-0.8	0.6	8-13	11	0-1	0
Bifidobacteriaceae (11)	0.9-1.7	1.3	11-15	13	2-7	4
Prevotellaceae (12)	0.6-2.3	1.2	9-18	13	1-4	2.5

A total of 13 strains were identified to fulfil our network analysis criteria, which were subsequently modelled as a community to assess their ecological roles (Figure 3, step 4; final filter). Community modelling showed that two potential butyrate producers were not predicted to do so in this context, and five strains consumed butyrate and/or produced H_2S (Supplementary file 2.9). Exclusion of these strains resulted in a purpose-based, six-strain community (Figure 3, step 5).

The resultant community consisted of a primary degrader of inulin (*Prevotella intermedia* 17), four secondary degraders that can produce butyrate (*Clostridium sordellii* ATCC 9714, *Clostridium* sp. SY8519, *Ruminococcaceae bacterium* D16 and *Peptostreptococcus anaerobius* DSM 2949) and a generalist (*Roseburia intestinalis* L1-82) (Figure 4). Modelling of our 6-strain purpose-based community (Supplementary file 2.10) confirmed the ecological roles of each strain while also predicting a butyrate production rate (mmol/gDW/h) with a percent change of 30.2% higher than what was predicted for the 151-strain community. A higher inulin to butyrate conversion efficiency was also reported, as the 6-strain community produced 20.2 mmol of butyrate per mmol of inulin degraded (Supplementary file 2.10), while the 151-strain community produced 11.1 mmol of butyrate per mmol of inulin degraded (Supplementary file 2.7).

Figure 4 Community resilience in amino acid limitation. A: Modelling of our final inulin-to-butyrate 6-strain community in amino acid restricted media shows that members can cover auxotrophies for the rest of the community while also presenting a degree of functional redundancy. During a knock-one-out analysis (B to F) we identified that the community is resilient to the absence of individual member strains, other than Peptostreptococcus anaerobius *DSM 2949 (knockout not shown). The knocked-out strain is shown in grey in panels B to F. Metabolic fluxes for each knock out condition are shown in Supplementary file 3*

Testing resilience in a minimal, purpose-based community

To test the resilience of our purpose-based community, we modelled it in media deprived of proteinogenic amino acids (Figure 4, Supplementary file 2.11). Predictions reported that every strain could grow, supported by complementary prototrophies. Figure 4A shows that amino acid prototrophy functional redundancy is scarce within the proposed community and that *Peptostreptococcus anaerobius* DSM 2949 plays a central role in this nutritional environment as the sole contributor of five amino acids. While every predicted butyrate producer adopts a butyrate consuming phenotype in this extreme scenario, butyrate production remains higher than overall community consumption, however at a percent change of 65.7% lower than the predicted rate in original circumstances. The overall positive flux is explained by the phenotype adopted by *Roseburia intestinalis* L1-82, which remains a butyrate producer. Conversion efficiency was also lower, with 6.2 mmol of butyrate produced per mmol of inulin consumed.

We additionally tested network resilience by adopting a knock-one-out approach in community modelling. Importantly, no strain was essential under standard conditions (cfUDM + inulin, Supplementary file 2.12). The overall community also showed resilience to missing one of five of its members under conditions of proteinogenic amino acid restriction (Figure 4B-F, Supplementary file 2.13). Yet, supporting the role of *Peptostreptococcus anaerobius* DSM 2949, it was shown that none of the other five strains were able to grow in amino acid restricted media in the absence of this strain. Together, these results highlight the use of modelling to design a resilient community, but also to test for weaknesses within the community, with the potential to make targeted additions of either strains or nutritional supplements to protect against such weaknesses.

To further support the rationale behind our network analysis selection criteria, we designed a community with a similar structure, however using six strains that did not fit our criteria; four butyrate producers (*Anaerofustis stercorihominis* DSM 17244, *Odoribacter splanchnicus* DSM 20712, *Alcaligenes faecalis* subsp. *Faecalis* NCIB 8687 and *Butyricimonas synergistica* DSM 23225) with the lowest export/import score among the butyrate producing ecological role in our sample. Meanwhile, the selected inulin degrader (*Prevotella intermedia* ATCC 25611) and a generalist

(*Roseburia inulinivorans* DSM 16841) also did not meet our requirements regarding reciprocation within a network, and these examples were selected due to their close phylogenetic relationship with the two primary degraders in the original 6-strain community (Figure 3B, poor metabolic contribution to the community). Community modelling in standard media showed that only two strains produce butyrate in this context (one butyrate producer and the community generalist) while one adopts a butyrate consuming phenotype (Supplementary file 2.14). Community butyrate production rate was predicted at a percent change of 60.6% lower than our original 6-strain purpose-based community, while only 12.9 mmol of butyrate was produced per mmol of inulin degraded. In testing the resilience of this network to amino acid limitation, we found that the community had no resilience for methionine restriction, as the amino acid was inferred to be essential for five of the strains. The sixth additional strain (*Alcaligenes faecalis* subsp. *Faecalis* NCIB 8687) whilst not a methionine prototroph, requires alanine from external sources while also relying on primary degraders for access to simpler carbohydrates (Supplementary file 2.15 and 2.16). These findings suggest that the criteria we have devised to identify key members within a complex community can effectively guide the design process for resilient, minimal communities.

Discussion

Here, we leverage genome-scale modelling and reverse ecology principles to characterise 816 strains with potential to occur in the human gut in terms of their capabilities to utilise a representative group of carbohydrates and the resultant metabolites generated. This has provided novel insights into gut microbial ecology as well as intervention and consortia design. The identified metabolic capabilities of individual strains inform the rational design of a purpose-based microbial community. The framework we propose here has potential to guide future efforts for rational microbiome interventions.

By dividing the carbohydrate dimension into structural categories, our analysis identifies functional trends and trade-offs among prevalent gut taxa. First, we find a funnel-like effect where a lower number of strains are compatible with carbohydrates with increased structural complexity, and vice versa. This is coherent as complex polysaccharide utilisation often requires coordinated synthesis of a

range of carbohydrate-active enzymes (102-105), demanding high resource and energy investment (106). This capability may stem from a long-standing evolutionary relationship with the host and its diet as polysaccharide utilisation seems to be a highly specialised trait that provides crucial leverage to select commensals over other organisms (107-109). However, it may come at a cost, as our findings suggest that families that have a tendency to primary degrader roles are compatible with a narrower set of simpler metabolites compared to other families. Notably, while we observe a predisposition among gut families towards specific degrader niches, variation at the strain level needs to be accounted for.

Primary degraders occupy key nutritional niches in community contexts. We found these roles to be primarily occupied by members of the *Bacteroidaceae*, *Prevotellaceae* and *Bifidobacteriaceae* families. Myriad studies have suggested that the families display competitive interactions due to overlapping ecological niches and energy extraction from diet (4, 10, 12, 110-113). Our analysis shows that *Bacteroidaceae* and *Prevotellaceae* strains present similar carbohydrate utilisation patterns, which compared to other families, favour more complex carbohydrates over monosaccharides. Further, members of both families were inferred to share the similar metabolite production profiles. These findings offer mechanistic grounding for the competitive nature between the two families. Through community modelling we find that members of the *Bacteroidaceae* display a less collaborative phenotype towards the host, by producing H_2S, and towards other strains in the community, by exporting a comparatively low number of amino acids. Alternatively, members of the *Bifidobacteriaceae* were predicted to export the highest number of amino acids among the three families we identified as likely to occupy a primary degrader role. This is interesting in the context of inulin supplementation, which consistently promotes *Bifidobacterium* abundance, while it decreases *Bacteroides* relative abundance, despite both groups encompassing a similar number of strains with the ability to utilise this substrate within AGORA (92). It is possible that the more supportive phenotype displayed by the *Bifidobacteriaceae* better enables the establishment of trophic networks that reciprocate to the family, allowing them to thrive.

Polysaccharide utilisation in the gut is an uncommon phenomenon, at best, that occurs at a strain-level, and yet polysaccharides are considered a key carbon and energy source for microbes in the larger bowel (13). This implies that polysaccharide utilisers act as energy gatekeepers for the rest of the community. Understanding how polysaccharide utilisation traits are distributed in an individual's microbiome will enable us to rationally modulate representative portions of the community. Similarly, our findings suggest that while metabolite production in large bowel strains is largely consistent across monosaccharides, metabolite production profiles exceed phylogeny, with more than 100 different metabolite production profiles being shared by only four or fewer strains. Hence, community metabolic outputs are likely more sensitive to microbiome composition, at the strain level, than monosaccharide dietary content per se. Exploration of the specific nutritional conditions that favour accumulation of different metabolites will improve our ability to modulate metabolic outputs within the human gut microbiome.

Together, these findings further underline the need for improved interventions, where DRCs are selected to rationally match microbial community metabolic capabilities and hierarchical structure. As a proof of principle, we envision the design of an inulin-degrading, butyrate-producing community that is better fitted to improve metabolic outputs. The framework we propose for the design of purpose-based communities is replicable, traceable and, together with the functional dataset we have generated as part of this study, can be applied for alternative substrates and metabolic objectives. A similar approach could further aim to explore media conditions or strain combinations that minimise the production of metabolites with a negative connotation. For example, the six-strain community is also predicted to produce succinate, which has been linked to local inflammation when produced in excess (114, 115). In addition to considering interactions in the carbohydrate dimension, we also consider requirements for other nutrient dimensions, namely, amino acids. The community we design is shown to be resilient in terms of amino acid auxotrophies, which are mostly covered by *P. anaerobius* DSM 2949. High reliance on a single member of the community is not necessarily a desirable trait from a resilience perspective; therefore, the addition of an extra member to distribute this burden might be beneficial. Nevertheless, it is likely that increasing the number of total

community members exponentially increases its complexity and, hence, should be carefully considered. Alternative nutritional dimensions, such as cofactor requirements should be accounted for in the design of resilient microbial communities, however we have reported on limitations when assessing this dimension in AGORA GEMs (63). Every modelling-based approach should further aim for experimental validation, which we are actively pursuing to ensure the practical applicability of our framework and further refinement. However, isolating and culturing human gut microbiome commensals remains challenging. An additional filter that could be applied to our pipeline could be to exclude uncultured strains, aiming to design a community of previously cultured members.

Our findings highlight that the actual challenges for intervention response may be found at lower levels of carbohydrate complexity. DRC compatibility is a crucial consideration, yet functional redundancy at lower levels complicates rational intervention design. For our example community, we rely on a DRC that releases fructose and glucose, two widely utilised monosaccharides in the human gut. Although the use of inulin for this exercise adequately showcases the rational process for assembly of purpose-based microbial communities, in practice, it may promote alternative populations and outcomes even following inoculation efforts of such community.

We utilise GEMNAST to infer functional traits from AGORA GEMs, and by combining community modelling and network analysis, we device a feasible, resilient microbial community capable of butyrate production from inulin degradation. Our study finds that butyrate production, a key metabolite for colonocyte health (88, 89), is an uncommon capability among human gut strains represented in AGORA. This implies that efforts for the promotion of gut health need to target these strains, and nutritional contexts that promote butyrate production, in particular. This further underscores the need of supporting fibre supplementation with a rationally designed consortia that can better guarantee the delivery of intended metabolic outputs and that these communities should include identified butyrate producers.

The limited count of butyrate producers in AGORA might additionally be explained or influenced by challenges related to gaps in genome annotation related to transporter gene polymorphisms and lack of experimental data, resulting in GEMs that do not encode the corresponding pathways (116, 117).

Additionally, further prediction of metabolite production rates in this instance would have provided valuable insights. However, limited experimental data on enzyme kinetics prevents reliable mapping of metabolite production rates. Nevertheless, genome-scale modelling is a rapidly growing technology with an ever-expanding repertoire of GEMs and efforts to develop GEMs-based frameworks that better capture cellular and community dynamics are ongoing (64, 118, 119); for example, the integration of metatranscriptomics and GEMs community modelling (64) or the application of AI to improve model accuracy (116, 119). Importantly, predicted relative differences and percentage changes can provide meaningful insights of the studied system (120, 121). Recently, a quantitative metabolite exchange scoring system was proposed to identify exchanged metabolites that are central in a community of GEMs (21). Meanwhile, the conceptual framework we propose here focuses on the identification of strains/GEMs, in a given network, that can better guarantee the establishment of resilient minimal communities. The versatility in GEMNAST allows its application for the study of alternative groups of molecules. Evidence suggest that proteoglycans play a similar role to that of DRCs in the gut (122-125) and future efforts could be centred around discerning metabolic capabilities in terms of this dimension of molecules. AGORA2, an extended version of AGORA, was recently released (126), and includes close to 8,000 GEMs for strains that can inhabit the human gut. For this study we rely on its predecessor due to its extensive application and validation in the field (63, 127-129). While the development of GEMs, as a technology, is ongoing, they are ideally placed for analyses such as the one we perform here, where we granularly inspect cellular metabolism and inter-strain interactions for community design, difficult to perform through alternative modalities. With technology evolving at a faster pace than ever before, and GEMs rapidly gaining momentum, studies like this are necessary to inform the field of microbial ecology and establish the foundations of consortia and intervention design.

Conclusions

By combining genome-scale modelling and reverse ecology we offer a yet untapped perspective into gut ecology, personalised intervention, and consortia design. Our functional analysis enables novel insights into metabolic capabilities distribution and community hierarchies in the human gut and sets

the foundation for the framework we propose for the design of purpose-based communities. Such framework further shows promising results to guide the design of resilient, minimal communities capable of supporting intervention outcomes. In so doing, our work paves the way to the development of holistic strategies to improve gut health.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

All data generated or analysed during this study are included in this published article and its supplementary information files. The datasets generated and/or analysed during the current study are available in the purpose-based community design repository, https://bitbucket.csiro.au/scm/~mol131/purpose-based-community-design.git

Competing interests

The authors declare that they have no competing interests.

Funding

The preparation of this manuscript was supported through funding from CSIRO Microbiomes for One Systems Health (MOSH)-Future Science Platform. It was also supported by the Environment Business Unit, CSIRO Australia. This work was initially supported by the University of Sydney's Centre for Advanced Food Engineering. J.M. acknowledges a PhD scholarship from the Faculty of Engineering at the University of Sydney. E.S. acknowledges financial support from the à Beckett Cancer Research Trust (University of Sydney Fellowship).

Authors' contributions

J.M., M.R., E.S. conceived of the presented idea. J.M. reviewed the literature, generated and analysed GEMs data and performed networks analyses. J.M., M.R., D.M., and E.S. were major contributors in writing the manuscript with input from all authors. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Lambeth SM, Carson T, Lowe J, Ramaraj T, Leff JW, Luo L, et al. Composition, diversity and abundance of gut microbiome in prediabetes and type 2 diabetes. Journal of Diabetes and Obesity. 2015;2(3):1.
2. Everard A, Lazarevic V, Derrien M, Girard M, Muccioli GG, Neyrinck AM, et al. Responses of gut microbiota and glucose and lipid metabolism to prebiotics in genetic obese and diet-induced leptin-resistant mice. diabetes. 2011;60(11):2775-86.
3. Cani PD, Bibiloni R, Knauf C, Waget A, Neyrinck AM, Delzenne NM, Burcelin R. Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet–induced obesity and diabetes in mice. Diabetes. 2008;57(6):1470-81.
4. Vieira-Silva S, Falony G, Darzi Y, Lima-Mendez G, Garcia Yunta R, Okuda S, et al. Species–function relationships shape ecological properties of the human gut microbiome. Nature microbiology. 2016;1(8):1-8.

5. Valles-Colomer M, Menni C, Berry SE, Valdes AM, Spector TD, Segata N. Cardiometabolic health, diet and the gut microbiome: A meta-omics perspective. Nature Medicine. 2023;29(3):551-61.
6. Birt DF, Boylston T, Hendrich S, Jane J-L, Hollis J, Li L, et al. Resistant starch: promise for improving human health. Advances in nutrition (Bethesda, Md). 2013;4(6):587-601.
7. Higgins JA. Resistant starch: metabolic effects and potential health benefits. Journal of AOAC International. 2004;87(3):761-8.
8. Walker AW, Ince J, Duncan SH, Webster LM, Holtrop G, Ze X, et al. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. The ISME Journal. 2011;5(2):220-30.
9. Delannoy-Bruno O, Desai C, Castillo JJ, Couture G, Barve RA, Lombard V, et al. An approach for evaluating the effects of dietary fiber polysaccharides on the human gut microbiome and plasma proteome. Proceedings of the National Academy of Sciences. 2022;119(20):e2123411119.
10. Kovatcheva-Datchary P, Nilsson A, Akrami R, Lee YS, De Vadder F, Arora T, et al. Dietary fiber-induced improvement in glucose metabolism is associated with increased abundance of *Prevotella*. Cell Metabolism. 2015;22(6):971-82.
11. Sun Y, Zhang S, Nie Q, He H, Tan H, Geng F, et al. Gut firmicutes: Relationship with dietary fiber and role in host homeostasis. Critical reviews in food science and nutrition. 2023;63(33):12073-88.
12. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. nature. 2011;473(7346):174-80.
13. Chung WS, Walker AW, Louis P, Parkhill J, Vermeiren J, Bosscher D, et al. Modulation of the human gut microbiota by dietary fibres occurs at the species level. BMC Biol. 2016;14:3.
14. O'Donnell JA, Zheng T, Meric G, Marques FZ. The gut microbiome and hypertension. Nature Reviews Nephrology. 2023;19(3):153-67.
15. Beyer-Sehlmeyer G, Glei M, Hartmann E, Hughes R, Persin C, Böhm V, et al. Butyrate is only one of several growth inhibitors produced during gut flora-mediated fermentation of dietary fibre sources. British Journal of Nutrition. 2003;90(6):1057-70.
16. Menni C, Jackson MA, Pallister T, Steves CJ, Spector TD, Valdes AM. Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. International Journal of Obesity. 2017;41(7):1099-105.
17. Andersen V, Olsen A, Carbonnel F, Tjonneland A, Vogel U. Diet and risk of inflammatory bowel disease. Dig Liver Dis. 2012;44(3):185-94.
18. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature. 2014;505(7484):559-63.
19. Mottawea W, Chiang C-K, Mühlbauer M, Starr AE, Butcher J, Abujamel T, et al. Altered intestinal microbiota–host mitochondria crosstalk in new onset Crohn's disease. Nature communications. 2016;7(1):13419.
20. Dordević D, Jančíková S, Vítězová M, Kushkevych I. Hydrogen sulfide toxicity in the gut environment: Meta-analysis of sulfate-reducing and lactic acid bacteria in inflammatory processes. Journal of advanced research. 2021;27:55-69.
21. Marcelino VR, Welsh C, Diener C, Gulliver EL, Rutten EL, Young RB, et al. Disease-specific loss of microbial cross-feeding interactions in the human gut. Nature Communications. 2023;14(1):6546.
22. Xie L, Alam MJ, Marques FZ, Mackay CR, editors. A major mechanism for immunomodulation: Dietary fibres and acid metabolites. Seminars in Immunology; 2023: Elsevier.
23. Thorburn AN, McKenzie CI, Shen S, Stanley D, Macia L, Mason LJ, et al. Evidence that asthma is a developmental origin disease influenced by maternal diet and bacterial metabolites. Nature Communications. 2015;6(1):1-13.
24. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science. 2011;334(6052):105-8.

750 25. Zhang C, Yin A, Li H, Wang R, Wu G, Shen J, et al. Dietary modulation of gut microbiota
751 contributes to alleviation of both genetic and simple obesity in children. EBioMedicine.
752 2015;2(8):968-84.
753 26. Koay YC, Wali JA, Luk AW, Macia L, Cogger VC, Pulpitel TJ, et al. Ingestion of resistant starch
754 by mice markedly increases microbiome-derived metabolites. The FASEB Journal. 2019;33(7):8033-
755 42.
756 27. Van den Abbeele P, Verstraete W, El Aidy S, Geirnaert A, Van de Wiele T. Prebiotics, faecal
757 transplants and microbial network units to stimulate biodiversity of the human gut microbiome.
758 Microb Biotechnol. 2013;6(4):335-40.
759 28. Culp EJ, Goodman AL. Cross-feeding in the gut microbiome: Ecology and mechanisms. Cell
760 Host & Microbe. 2023;31(4):485-99.
761 29. Jung D-H, Park C-S. Resistant starch utilization by Bifidobacterium, the beneficial human gut
762 bacteria. Food Science and Biotechnology. 2023;32(4):441-52.
763 30. Carlson RP, Beck AE, Phalak P, Fields MW, Gedeon T, Hanley L, et al. Competitive resource
764 allocation to metabolic pathways contributes to overflow metabolisms and emergent properties in
765 cross-feeding microbial consortia. Biochemical Society Transactions. 2018;46(2):269-84.
766 31. Fuhrer T, Fischer E, Sauer U. Experimental identification and quantification of glucose
767 metabolism in seven bacterial species. J Bacteriol. 2005;187(5):1581-90.
768 32. Laverde Gomez JA, Mukhopadhya I, Duncan SH, Louis P, Shaw S, Collie-Duguid E, et al.
769 Formate cross-feeding and cooperative metabolic interactions revealed by transcriptomics in
770 co-cultures of acetogenic and amylolytic human colonic bacteria. Environmental Microbiology.
771 2019;21(1):259-71.
772 33. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience
773 of the human gut microbiota. Nature. 2012;489(7415):220.
774 34. Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, et al. Individuality in gut microbiota
775 composition is a complex polygenic trait shaped by multiple environmental and host genetic factors.
776 Proceedings of the National Academy of Sciences. 2010;107(44):18933-8.
777 35. Bourriaud C, Robins R, Martin L, Kozlowski F, Tenailleau E, Cherbut C, Michel C. Lactate is
778 mainly fermented to butyrate by human intestinal microfloras but inter-individual variation is
779 evident. Journal of Applied Microbiology. 2005;99(1):201-12.
780 36. Turnbaugh PJ, Quince C, Faith JJ, McHardy AC, Yatsunenko T, Niazi F, et al. Organismal,
781 genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins.
782 Proceedings of the National Academy of Sciences. 2010;107(16):7503-8.
783 37. Bonazzi E, Bretin A, Vigué L, Hao F, Patterson AD, Gewirtz AT, Chassaing B. Individualized
784 microbiotas dictate the impact of dietary fiber on colitis sensitivity. Microbiome. 2024;12(1):1-14.
785 38. Lopez-Santamarina A, Mondragon AdC, Cardelle-Cobas A, Santos EM, Porto-Arias JJ, Cepeda
786 A, Miranda JM. Effects of unconventional work and shift work on the human gut microbiota and the
787 potential of probiotics to restore dysbiosis. Nutrients. 2023;15(13):3070.
788 39. Derrien M, Veiga P. Rethinking Diet to Aid Human-Microbe Symbiosis. Trends Microbiol.
789 2017;25(2):100-12.
790 40. McFarland LV, Dublin S. Meta-analysis of probiotics for the treatment of irritable bowel
791 syndrome. World journal of gastroenterology: WJG. 2008;14(17):2650.
792 41. Vinelli V, Biscotti P, Martini D, Del Bo' C, Marino M, Meroño T, et al. Effects of dietary fibers
793 on short-chain fatty acids and gut microbiota composition in healthy adults: a systematic review.
794 Nutrients. 2022;14(13):2559.
795 42. Wagenaar CA, van de Put M, Bisschops M, Walrabenstein W, de Jonge CS, Herrema H, van
796 Schaardenburg D. The effect of dietary interventions on chronic inflammatory diseases in relation to
797 the microbiome: A systematic review. Nutrients. 2021;13(9):3208.
798 43. Healey GR, Murphy R, Brough L, Butts CA, Coad J. Interindividual variability in gut microbiota
799 and host response to dietary interventions. Nutrition reviews. 2017;75(12):1059-80.

800 44. Oliver A, Chase AB, Weihe C, Orchanian SB, Riedel SF, Hendrickson CL, et al. High-fiber,
801 whole-food dietary intervention alters the human gut microbiome but not fecal short-chain fatty
802 acids. Msystems. 2021;6(2):10.1128/msystems. 00115-21.
803 45. Han S, Lu Y, Xie J, Fei Y, Zheng G, Wang Z, et al. Probiotic gastrointestinal transit and
804 colonization after oral administration: A long journey. Frontiers in cellular and infection microbiology.
805 2021;11:609722.
806 46. Laursen MF, Laursen RP, Larnkjær A, Michaelsen KF, Bahl MI, Licht TR. Administration of two
807 probiotic strains during early childhood does not affect the endogenous gut microbiota composition
808 despite probiotic proliferation. BMC microbiology. 2017;17(1):1-9.
809 47. Zmora N, Zilberman-Schapira G, Suez J, Mor U, Dori-Bachash M, Bashiardes S, et al.
810 Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique
811 Host and Microbiome Features. Cell. 2018;174(6):1388-405 e21.
812 48. Turroni F, Milani C, Duranti S, Mancabelli L, Mangifesta M, Viappiani A, et al. Deciphering
813 bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics
814 approach. The ISME journal. 2016;10(7):1656-68.
815 49. Manichanh C, Reeder J, Gibert P, Varela E, Llopis M, Antolin M, et al. Reshaping the gut
816 microbiome with bacterial transplantation and antibiotic intake. Genome research.
817 2010;20(10):1411-9.
818 50. Wang T, Goyal A, Dubinkina V, Maslov S. Evidence for a multi-level trophic organization of the
819 human gut microbiome. PLoS computational biology. 2019;15(12):e1007524.
820 51. Zengler K, Zaramela LS. The social network of microorganisms - how auxotrophies shape
821 complex communities. Nat Rev Microbiol. 2018;16(6):383-90.
822 52. Chung WSF, Walker AW, Vermeiren J, Sheridan PO, Bosscher D, Garcia-Campayo V, et al.
823 Impact of carbohydrate substrate complexity on the diversity of the human colonic microbiota. FEMS
824 microbiology ecology. 2019;95(1):fiy201.
825 53. Gonçalves OS, Creevey CJ, Santana MF. Designing a synthetic microbial community through
826 genome metabolic modeling to enhance plant–microbe interaction. Environmental microbiome.
827 2023;18(1):81.
828 54. Li YF, Costello JC, Holloway AK, Hahn MW. "Reverse ecology" and the power of population
829 genomics. Evolution. 2008;62(12):2984-94.
830 55. Arevalo P, VanInsberghe D, Elsherbini J, Gore J, Polz MF. A reverse ecology approach based on
831 a biological definition of microbial populations. Cell. 2019;178(4):820-34. e14.
832 56. Levy R, Borenstein E. Reverse ecology: from systems to environments and back. Evolutionary
833 systems biology: Springer; 2012. p. 329-45.
834 57. Altamirano Á, Saa PA, Garrido D. Inferring composition and function of the human gut
835 microbiome in time and space: A review of genome-scale metabolic modelling tools. Computational
836 and Structural Biotechnology Journal. 2020;18:3897-904.
837 58. King ZA, Lu J, Dräger A, Miller P, Federowicz S, Lerman JA, et al. BiGG Models: A platform for
838 integrating, standardizing and sharing genome-scale models. Nucleic Acids Research.
839 2016;44(D1):D515-D22.
840 59. Thiele I, Palsson BO. A protocol for generating a high-quality genome-scale metabolic
841 reconstruction. Nat Protoc. 2010;5(1):93-121.
842 60. Ottman N, Davids M, Suarez-Diez M, Boeren S, Schaap PJ, Martins dos Santos VAP, et al.
843 Genome-scale model and omics analysis of metabolic capacities of *Akkermansia muciniphila* reveal a
844 preferential mucin-degrading lifestyle. Applied and Environmental Microbiology.
845 2017;83(18):e01014-17.
846 61. Sánchez BJ, Zhang C, Nilsson A, Lahtvee PJ, Kerkhoven EJ, Nielsen J. Improving the phenotype
847 predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints.
848 Molecular Systems Biology. 2017;13(8):935.

849 62. Zhang C, Sánchez BJ, Li F, Eiden CWQ, Scott Jr WT, Liebal UW, et al. Yeast9: A Consensus Yeast
850 Metabolic Model Enables Quantitative Analysis of Cellular Metabolism By Incorporating Big Data.
851 bioRxiv. 2023:2023.12. 03.569754.
852 63. Molina Ortiz JP, Read MN, McClure DD, Holmes A, Dehghani F, Shanahan ER. High throughput
853 genome scale modeling predicts microbial vitamin requirements contribute to gut microbiome
854 community structure. Gut Microbes. 2022;14(1):2118831.
855 64. Zampieri G, Campanaro S, Angione C, Treu L. Metatranscriptomics-guided genome-scale
856 metabolic modeling of microbial communities. Cell Reports Methods. 2023;3(1).
857 65. Diener C, Gibbons SM, Resendis-Antonio O. MICOM: metagenome-scale modeling to infer
858 metabolic interactions in the gut microbiota. Msystems. 2020;5(1).
859 66. Shoaie S, Nielsen J. Elucidating the interactions between the human gut microbiota and its
860 host through metabolic modeling. Front Genet. 2014;5:86.
861 67. Heinken A, Ravcheev DA, Baldini F, Heirendt L, Fleming RM, Thiele I. Systematic assessment
862 of secondary bile acid metabolism in gut microbes reveals distinct metabolic capabilities in
863 inflammatory bowel disease. Microbiome. 2019;7:1-18.
864 68. Molina Ortiz JP, McClure DD, Shanahan ER, Dehghani F, Holmes AJ, Read MN. Enabling
865 rational gut microbiome manipulations by understanding gut ecology through experimentally-
866 evidenced in silico models. Gut Microbes. 2021;13(1):1965698.
867 69. Magnusdottir S, Heinken A, Kutt L, Ravcheev DA, Bauer E, Noronha A, et al. Generation of
868 genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nature
869 Biotechnology. 2017;35(1):81-9.
870 70. Noronha A, Modamio J, Jarosz Y, Guerard E, Sompairac N, Preciat G, et al. The Virtual
871 Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and
872 disease. Nucleic acids research. 2019;47(D1):D614-D24.
873 71. Ebrahim A, Lerman JA, Palsson BO, Hyduke DR. COBRApy: constraints-based reconstruction
874 and analysis for python. BMC Systems Biology. 2013;7(1):1-6.
875 72. Gurobi Optimization L. Gurobi Optimizer Reference Manual 2023 [Available from:
876 https://www.gurobi.com.
877 73. Asnicar F, Leeming ER, Dimidi E, Mazidi M, Franks PW, Al Khatib H, et al. Blue poo: impact of
878 gut transit time on the gut microbiome using a novel marker. Gut. 2021;70(9):1665-74.
879 74. Ghosh OM, Good BH. Emergent evolutionary forces in spatial models of luminal growth and
880 their application to the human gut microbiota. Proceedings of the National Academy of Sciences.
881 2022;119(28):e2114931119.
882 75. The Human Microbiome Project C, Huttenhower C, Gevers D, Knight R, Abubucker S, Badger
883 JH, et al. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486:207.
884 76. Chen T, Yu W-H, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. The Human Oral
885 Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and
886 genomic information. Database. 2010;2010.
887 77. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial Community
888 Variation in Human Body Habitats Across Space and Time. Science. 2009;326(5960):1694-7.
889 78. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, et al. Topographical and
890 temporal diversity of the human skin microbiome. science. 2009;324(5931):1190-2.
891 79. Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, et al. NCBI
892 Taxonomy: a comprehensive update on curation, resources and tools. Database. 2020;2020.
893 80. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn:
894 Machine Learning in Python. Journal of Machine Learning Research. 2011;12:2825-30.
895 81. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0:
896 fundamental algorithms for scientific computing in Python. Nature methods. 2020;17(3):261-72.
897 82. Waskom ML. seaborn: statistical data visualization. Journal of open source software.
898 2021;6(60):3021.

83. Hunter JD. Matplotlib: A 2D graphics environment. Computing in science & engineering. 2007;9(03):90-5.
84. Ge Z-B, Zhai Z-Q, Xie W-Y, Dai J, Huang K, Johnson DR, et al. Two-tiered mutualism improves survival and competitiveness of cross-feeding soil bacteria. The ISME journal. 2023;17(11):2090-102.
85. Moens F, Verce M, De Vuyst L. Lactate-and acetate-based cross-feeding interactions between selected strains of lactobacilli, bifidobacteria and colon bacteria in the presence of inulin-type fructans. International Journal of Food Microbiology. 2017;241:225-36.
86. Cockburn DW, Koropatkin NM. Polysaccharide degradation by the intestinal microbiota and its influence on human health and disease. Journal of molecular biology. 2016;428(16):3230-52.
87. Siddiqui MT, Cresci GA. The immunomodulatory functions of butyrate. Journal of inflammation research. 2021;14:6025.
88. Gasaly N, Hermoso MA, Gotteland M. Butyrate and the fine-tuning of colonic homeostasis: implication for inflammatory bowel diseases. International journal of molecular sciences. 2021;22(6):3061.
89. Salvi PS, Cowles RA. Butyrate and the intestinal epithelium: modulation of proliferation and inflammation in homeostasis and disease. Cells. 2021;10(7):1775.
90. Cerqueira FM, Photenhauer AL, Pollet RM, Brown HA, Koropatkin NM. Starch digestion by gut bacteria: crowdsourcing for carbs. Trends in Microbiology. 2020;28(2):95-108.
91. Bayer G, Ganobis CM, Allen-Vercoe E, Philpott DJ. Defined gut microbial communities: promising tools to understand and combat disease. Microbes and Infection. 2021;23(6-7):104816.
92. Le Bastard Q, Chapelet G, Javaudin F, Lepelletier D, Batard E, Montassier E. The effects of inulin on gut microbial composition: a systematic review of evidence from human studies. European Journal of Clinical Microbiology & Infectious Diseases. 2020;39:403-13.
93. Moens F, De Vuyst L. Inulin-type fructan degradation capacity of Clostridium cluster IV and XIVa butyrate-producing colon bacteria and their associated metabolic outcomes. Beneficial microbes. 2017;8(3):473-90.
94. Machiels K, Joossens M, Sabino J, De Preter V, Arijs I, Eeckhaut V, et al. A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. Gut. 2014;63(8):1275-83.
95. Beaumont M, Andriamihaja M, Lan A, Khodorova N, Audebert M, Blouin J-M, et al. Detrimental effects for colonocytes of an increased exposure to luminal hydrogen sulfide: the adaptive response. Free Radical Biology and Medicine. 2016;93:155-64.
96. Thébault E, Fontaine C. Stability of ecological communities and the architecture of mutualistic and trophic networks. Science. 2010;329(5993):853-6.
97. Zhang Z, Geng J, Tang X, Fan H, Xu J, Wen X, et al. Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota. ISME J. 2014;8(4):881-93.
98. Kost C, Patil KR, Friedman J, Garcia SL, Ralser M. Metabolic exchanges are ubiquitous in natural microbial communities. Nature Microbiology. 2023:1-9.
99. Rakoff-Nahoum S, Foster KR, Comstock LE. The evolution of cooperation within the gut microbiota. Nature. 2016;533(7602):255-9.
100. Liu H, Liao C, Wu L, Tang J, Chen J, Lei C, et al. Ecological dynamics of the gut microbiome in response to dietary fiber. The ISME journal. 2022;16(8):2040-55.
101. Mahowald MA, Rey FE, Seedorf H, Turnbaugh PJ, Fulton RS, Wollam A, et al. Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. Proc Natl Acad Sci U S A. 2009;106(14):5859-64.
102. Wardman JF, Bains RK, Rahfeld P, Withers SG. Carbohydrate-active enzymes (CAZymes) in the gut microbiome. Nature Reviews Microbiology. 2022:1-15.
103. Sheridan PO, Martin JC, Lawley TD, Browne HP, Harris HM, Bernalier-Donadille A, et al. Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic Firmicutes. Microb Genom. 2016;2(2):e000043.

949 104. El Kaoutari A, Armougom F, Gordon JI, Raoult D, Henrissat B. The abundance and variety of
950 carbohydrate-active enzymes in the human gut microbiota. Nat Rev Microbiol. 2013;11(7):497-504.
951 105. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-
952 active enzymes database (CAZy) in 2013. Nucleic Acids Research. 2013;42(D1):D490-D5.
953 106. Horn SJ, Sikorski P, Cederkvist JB, Vaaje-Kolstad G, Sørlie M, Synstad B, et al. Costs and
954 benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. Proceedings of the
955 National Academy of Sciences. 2006;103(48):18089-94.
956 107. Flint HJ, Scott KP, Duncan SH, Louis P, Forano E. Microbial degradation of complex
957 carbohydrates in the gut. Gut Microbes. 2012;3(4):289-306.
958 108. Cantarel BL, Lombard V, Henrissat B. Complex carbohydrate utilization by the healthy human
959 microbiome. PLoS One. 2012;7(6):e28742.
960 109. Martens EC, Kelly AG, Tauzin AS, Brumer H. The devil lies in the details: how variations in
961 polysaccharide fine-structure impact the physiology and evolution of gut microbes. Journal of
962 molecular biology. 2014;426(23):3851-65.
963 110. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, et al. Linking long-term
964 dietary patterns with gut microbial enterotypes. Science. 2011;334(6052):105-8.
965 111. Chen T, Long W, Zhang C, Liu S, Zhao L, Hamaker BR. Fiber-utilizing capacity varies in
966 Prevotella-versus Bacteroides-dominated gut microbiota. Scientific reports. 2017;7(1):1-7.
967 112. Roager HM, Licht TR, Poulsen SK, Larsen TM, Bahl MI. Microbial enterotypes, inferred by the
968 prevotella-to-bacteroides ratio, remained stable during a 6-month randomized controlled diet
969 intervention with the new nordic diet. Applied and environmental microbiology. 2014;80(3):1142-9.
970 113. Ley RE. Prevotella in the gut: choose carefully. Nature reviews Gastroenterology &
971 hepatology. 2016;13(2):69-70.
972 114. Connors J, Dawe N, Van Limbergen J. The role of succinate in the regulation of intestinal
973 inflammation. Nutrients. 2018;11(1):25.
974 115. Tannahill G, Curtis A, Adamik J, Palsson-McDermott E, McGettrick A, Goel G, et al. Succinate
975 is an inflammatory signal that induces IL-1β through HIF-1α. Nature. 2013;496(7444):238-42.
976 116. Chen C, Liao C, Liu Y-Y. Teasing out missing reactions in genome-scale metabolic networks
977 through hypergraph learning. Nature Communications. 2023;14(1):2375.
978 117. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, et al. Creation and analysis
979 of biochemical constraint-based models using the COBRA Toolbox v. 3.0. Nature Protocols.
980 2019;14(3):639-702.
981 118. Paklao T, Suratanee A, Plaimas K. ICON-GEMs: integration of co-expression network in
982 genome-scale metabolic models, shedding light through systems biology. BMC bioinformatics.
983 2023;24(1):492.
984 119. Chen Y, Gustafsson J, Tafur Rangel A, Anton M, Domenzain I, Kittikunapong C, et al.
985 Reconstruction, simulation and analysis of enzyme-constrained metabolic models using GECKO
986 Toolbox 3.0. Nature Protocols. 2024:1-39.
987 120. Melkonian C, Zorrilla F, Kjærbølling I, Blasche S, Machado D, Junge M, et al. Microbial
988 interactions shape cheese flavour formation. Nature Communications. 2023;14(1):8348.
989 121. Ginatt AA, Berihu M, Castel E, Medina S, Carmi G, Doron-Faigenboim A, et al. A metabolic
990 modeling-based framework for predicting trophic dependencies in native rhizobiomes of crop plants.
991 bioRxiv. 2023:2023.12. 11.571117.
992 122. Paone P, Cani PD. Mucus barrier, mucins and gut microbiota: the expected slimy partners?
993 Gut. 2020;69(12):2232-43.
994 123. Tailford LE, Crost EH, Kavanaugh D, Juge N. Mucin glycan foraging in the human gut
995 microbiome. Front Genet. 2015;6:81.
996 124. Berry D, Stecher B, Schintlmeister A, Reichert J, Brugiroux S, Wild B, et al. Host-compound
997 foraging by intestinal microbiota revealed by single-cell stable isotope probing. Proceedings of the
998 National Academy of Sciences. 2013;110(12):4720-5.

125. Van Herreweghen F, De Paepe K, Roume H, Kerckhof F-M, Van de Wiele T. Mucin degradation niche as a driver of microbiome composition and *Akkermansia muciniphila* abundance in a dynamic gut model is donor independent. FEMS Microbiology Ecology. 2018;94(12):fiy186.
126. Heinken A, Hertel J, Acharya G, Ravcheev DA, Nyga M, Okpala OE, et al. Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine. Nature Biotechnology. 2023:1-12.
127. Heinken A, Hertel J, Thiele I. Metabolic modelling reveals broad changes in gut microbial metabolism in inflammatory bowel disease patients with dysbiosis. NPJ systems biology and applications. 2021;7(1):1-11.
128. Heinken A, Thiele I. Systematic interrogation of the distinct metabolic potential in gut microbiomes of inflammatory bowel disease patients with dysbiosis. bioRxiv. 2019:640649.
129. Aden K, Rehman A, Waschina S, Pan W-H, Walker A, Lucio M, et al. Metabolic functions of gut microbes associate with efficacy of tumor necrosis factor antagonists in patients with inflammatory bowel diseases. Gastroenterology. 2019;157(5):1279-92. e11.

Supplementary material

Supplementary file 1

Description: Supplementary tables and figures.

Format: Word file (.docx)

Supplementary file 2

Description: Single strain and community modelling data results and data analysis.

Format: Excel file (.xlsx)

Supplementary file 3

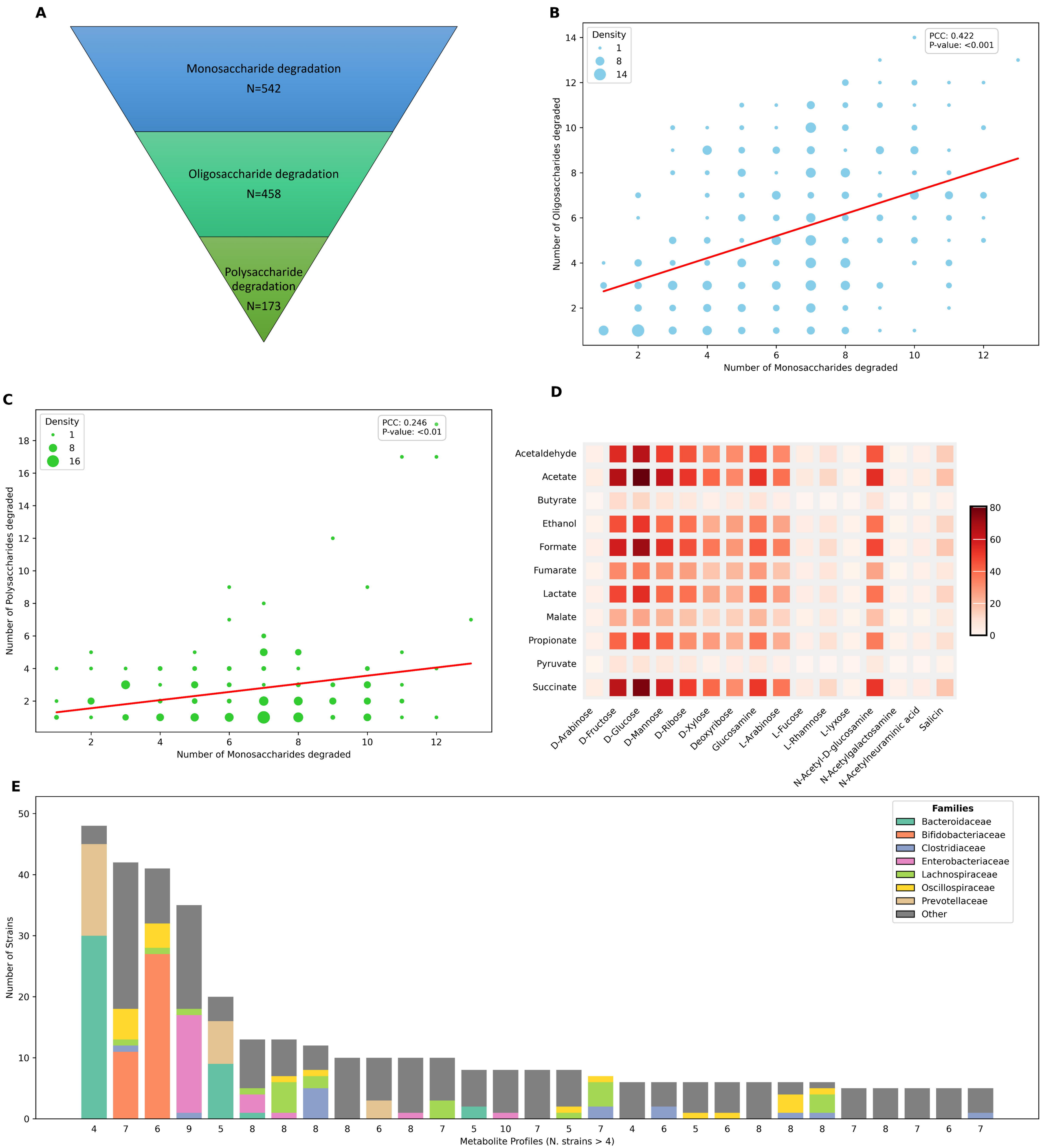
Description: Modelling fluxes from knock-one-out analysis in the 6-strain resilient community.

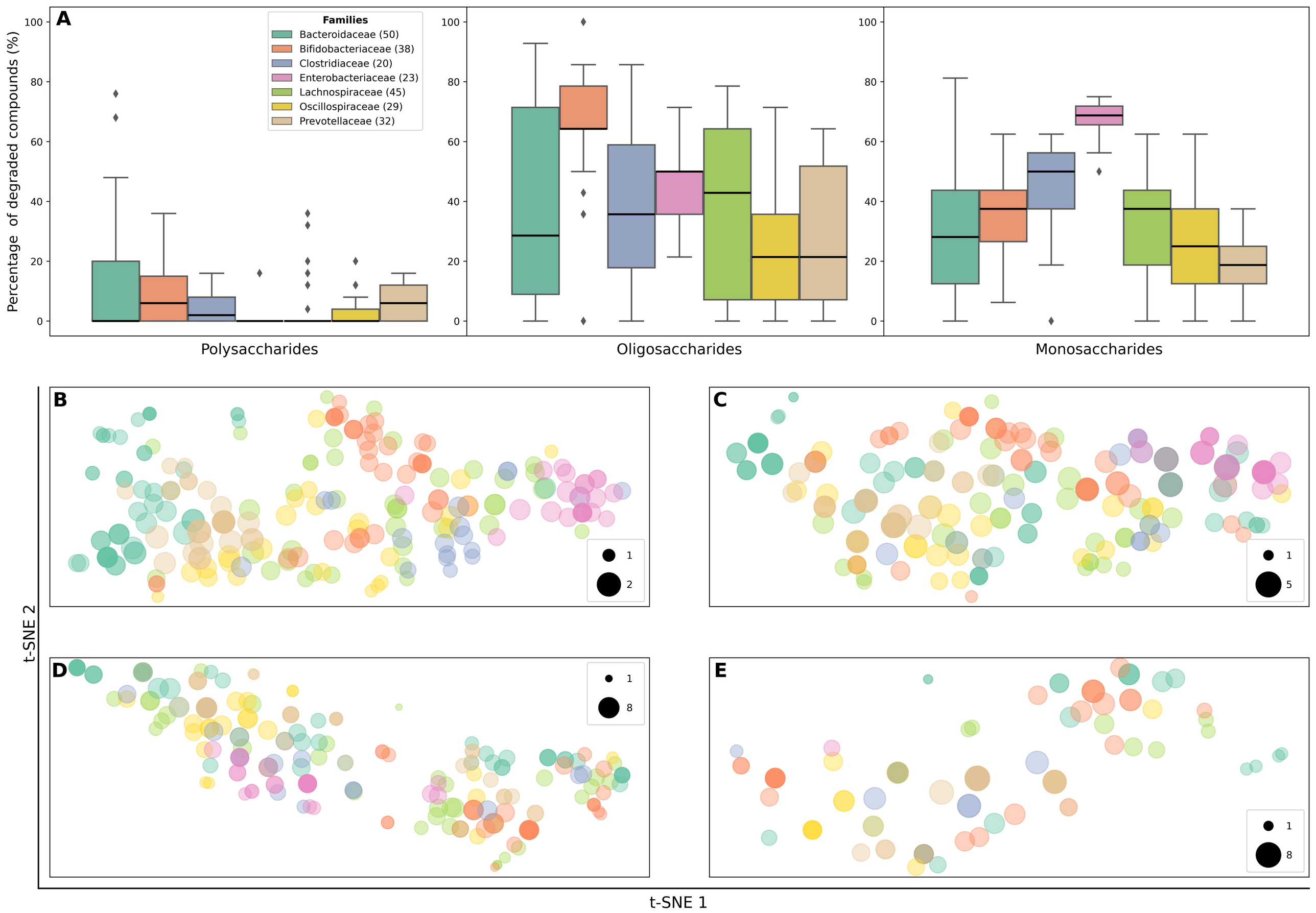
Format: Excel file (.xlsx)

Supplementary file 4

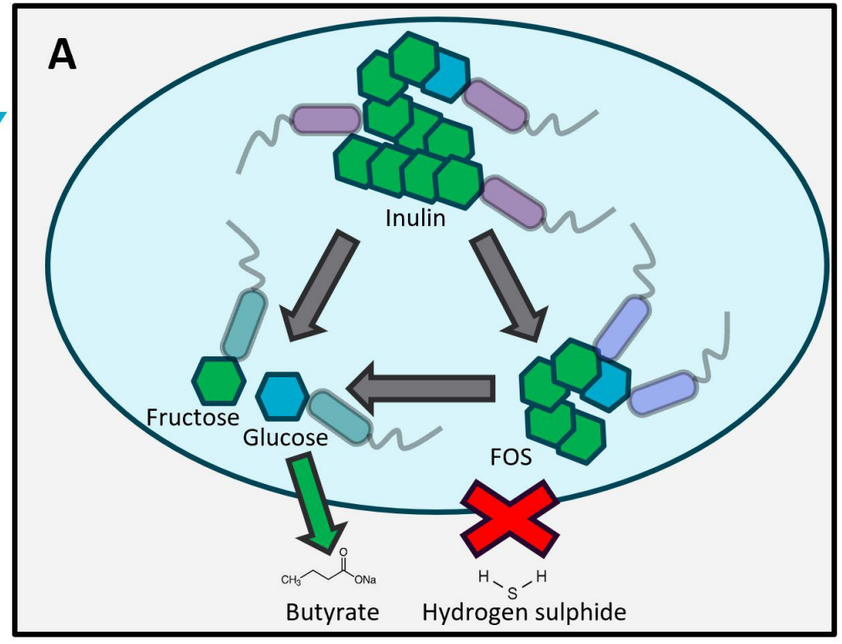
Description: Phylogeny tree file for large bowel strains in AGORA.

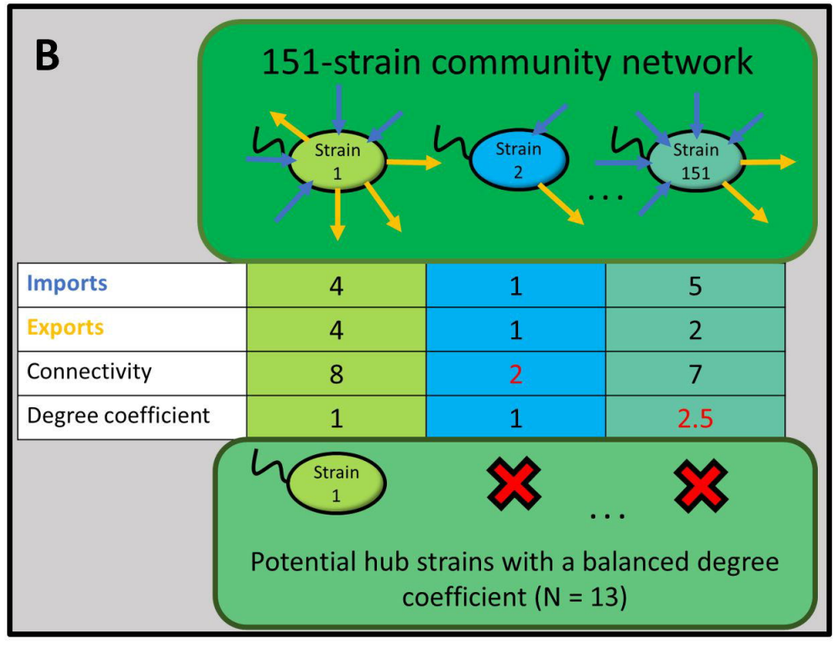
Format: Tree file (.phy)

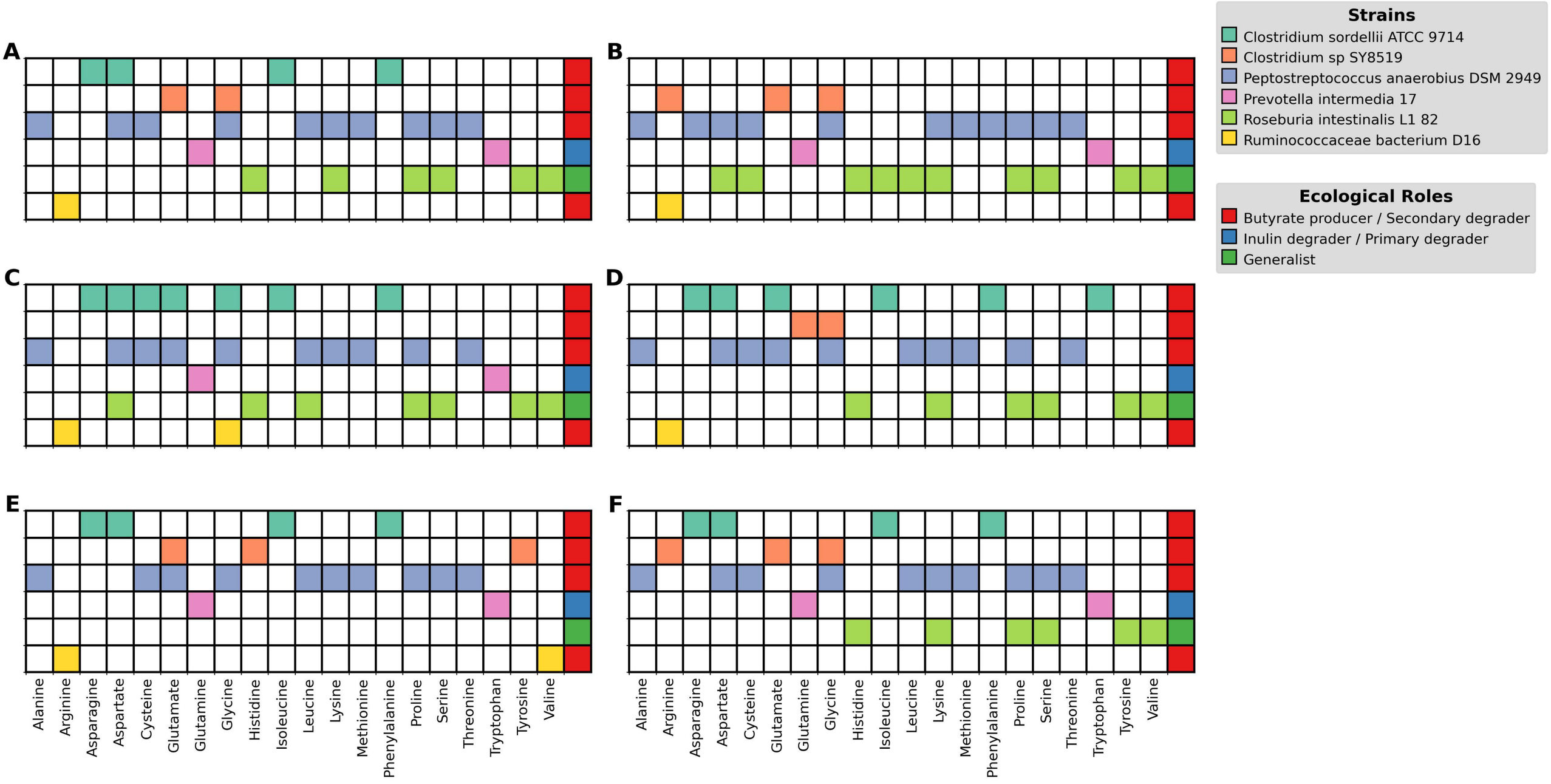

A
Monosaccharide degradation
N=542
Oligosaccharide degradation
N=458
Polysaccharide degradation
N=173
B
Density
1
8
14
PCC: 0.422
P-value: <0.001
Number of Oligosaccharides degraded
Number of Monosaccharides degraded
C
Density
1
8
16
PCC: 0.246
P-value: <0.01
Number of Polysaccharides degraded
Number of Monosaccharides degraded
D
Acetaldehyde
Acetate
Butyrate
Ethanol
Formate
Fumarate
Lactate
Malate
Propionate
Pyruvate
Succinate
D-Arabinose
D-Fructose
D-Glucose
D-Mannose
D-Ribose
D-Xylose
Deoxyribose
Glucosamine
L-Arabinose
L-Fucose
L-Rhamnose
L-lyxose
N-Acetyl-D-glucosamine
N-Acetylgalactosamine
N-Acetylneuraminic acid
Salicin
E
Families
Bacteroidaceae
Bifidobacteriaceae
Clostridiaceae
Enterobacteriaceae
Lachnospiraceae
Oscillospiraceae
Prevotellaceae
Other
Number of Strains
Metabolite Profiles (N. strains > 4)


A
Percentage of degraded compounds (%)
Families
Bacteroidaceae (50)
Bifidobacteriaceae (38)
Clostridiaceae (20)
Enterobacteriaceae (23)
Lachnospiraceae (45)
Oscillospiraceae (29)
Prevotellaceae (32)
Polysaccharides
Oligosaccharides
Monosaccharides
B
C
D
E
t-SNE 2
t-SNE 1

	598 AGORA gut microbiota strains
1. Function-guided selection: Strains with traits that can contribute to the objective of generating butyrate from inulin degradation	151 strains with desirable functional traits **(A)**
2. Preliminary assessment: primary degrader phenotype analysis; data collection for network analysis. CM with 151 strains	Community modelling (CM)
3. Resilience selection: Identification of hub strains with a balanced prototrophic/auxotrophic phenotype based on data from 151-strain CM	Network analysis **(B)**
4. Final filter: identify strains that may not contribute to the main objective (purpose-based selection) based on data from 13-strain CM	CM
5. A feasible (6 members), purpose-based community to better guarantee butyrate production from inulin degradation. Additional CM was performed to confirm resilience and ecological roles	6






A
B
C
D
E
F
Alanine
Arginine
Asparagine
Aspartate
Cysteine
Glutamate
Glutamine
Glycine
Histidine
Isoleucine
Leucine
Lysine
Methionine
Phenylalanine
Proline
Serine
Threonine
Tryptophan
Tyrosine
Valine
Strains
Clostridium sordellii ATCC 9714
Clostridium sp SY8519
Peptostreptococcus anaerobius DSM 2949
Prevotella intermedia 17
Roseburia intestinalis L1 82
Ruminococcaceae bacterium D16
Ecological Roles
Butyrate producer / Secondary degrader
Inulin degrader / Primary degrader
Generalist