



Cite this article: Kettle H, Louis P, Flint HJ. 2022 Process-based modelling of microbial community dynamics in the human colon. *J. R. Soc. Interface* 19: 20220489. <https://doi.org/10.1098/rsif.2022.0489>

Received: 5 July 2022

Accepted: 16 September 2022

Subject Category:

Life Sciences—Mathematics interface

Subject Areas:

systems biology, biomathematics, computational biology

Keywords:

microbial modelling, human gut, colon, colonic microbiota, modelling

Author for correspondence:

Helen Kettle

e-mail: helen.kettle@bioss.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6216295>.

Process-based modelling of microbial community dynamics in the human colon

Helen Kettle¹, Petra Louis² and Harry J. Flint²

¹Biomathematics and Statistics Scotland, James Clerk Maxwell Building, Peter Guthrie Tait Road, Edinburgh EH9 3FD, UK

²Gut Health Group, Rowett Institute, University of Aberdeen, Aberdeen, UK

HK, 0000-0002-6960-9733

The human colon contains a dynamic microbial community whose composition has important implications for human health. In this work, we build a process-based model of the colonic microbial ecosystem and compare with general empirical observations and the results of *in vivo* experiments. Our model comprises a complex microbial ecosystem along with absorption of short chain fatty acids (SCFA) and water by the host through the gut wall, variations in incoming dietary substrates (in the form of ‘meals’ whose composition varies in time), bowel movements, feedback on microbial growth from changes in pH resulting from SCFA production and multiple compartments to represent the proximal, transverse and distal colon. We verify our model against a number of observed criteria, e.g. total SCFA concentrations, SCFA ratios, mass of bowel movements, pH and water absorption over the transit time; and then run simulations investigating the effect of colonic transit time, and the composition and amount of indigestible carbohydrate in the host diet, which we compare with *in vivo* studies. The code is available as an R package (*microPopGut*) to aid future research.

1. Introduction

The human colon harbours a dense and diverse community of microbiota whose interactions with the host can have a profound effect on human health (e.g. [1,2]). Owing to the location of this community within its host, data collection and experimentation are problematic. Information on this system mostly comes from volunteer experiments in which diet and stool samples are monitored or from laboratory experiments using the microbes found in stool samples. Another approach is to put current knowledge into a mathematical framework and run simulations of the system to test our understanding and identify knowledge gaps. To this end, a number of mathematical models of this system have been developed (e.g. [3–7]).

When developing a model, a number of assumptions about the system are made in order to reduce complexity/dimensionality so that the model is easier to parametrize, run and analyse. Some modellers choose to reduce the microbial complexity and focus on the physics of the gut (e.g. [3,4]), some try to achieve a balance of both (e.g. [6]) and some choose to develop the microbial community (e.g. [7]). The model described here focuses on the microbial community dynamics and on interactions with the host, with a fairly simple model of the colon. We include the simulation of ‘meals’ (of random composition and size) arriving at the colon and look at the effects of bowel movements, both of which, as far as we are aware, have not been previously incorporated into such models. Having developed a complex model of human gut microbiota in a fermentor system [8], and publicly available software (*microPop*—an R package for modelling microbial communities [9]) we now incorporate this 10-group microbial ecosystem model (table 1) into a model of the human gut in order to simulate the effects of diet and host on the microbial composition and subsequent short chain fatty acid (SCFA) production.

Table 1. Microbial functional groups included in the model (and the R package microPop [9]) and described by Kettle *et al.* [8]. Users should be aware that the parameter values given in the data frames in the software will almost certainly change with increasing knowledge of gut microbiota and in some cases are simply a ‘best guess’.

microPop Name	abbr.	description	examples
bacteroides	B	acetate–propionate–succinate group	<i>Bacteroides</i> spp., <i>Prevotella</i> spp., <i>Akkermansia muciniphila</i> (Verrucomicrobia)
NoButyStarchDeg	NBSD	non-butyrate-forming starch degraders	Ruminococcaceae related to <i>Ruminococcus bromii</i> . Also includes certain Lachnospiraceae
NoButyFibreDeg	NBFD	non-butyrate-forming fibre degraders	Ruminococcaceae related to <i>Ruminococcus albus</i> , <i>Ruminococcus flavefaciens</i> . Also includes certain Lachnospiraceae
LactateProducers	LP	lactate producers	Actinobacteria, especially <i>Bifidobacterium</i> spp., <i>Collinsella aerofaciens</i>
ButyrateProducers1	BP1	butyrate producers	Lachnospiraceae related to <i>Eubacterium rectale</i> , <i>Roseburia</i> spp.
ButyrateProducers2	BP2	butyrate producers	certain Ruminococcaceae, in particular <i>Faecalibacterium prausnitzii</i>
PropionateProducers	PP	propionate producers	Veillonellaceae, e.g. <i>Veillonella</i> spp., <i>Megasphaera elsdenii</i>
ButyrateProducers3	BP3	butyrate producers	Lachnospiraceae related to <i>Eubacterium hallii</i> , <i>Anaerostipes</i> spp.
acetogens	A	acetate producers	certain Lachnospiraceae, e.g. <i>Blautia hydrogenotrophica</i>
methanogens	M	methanogenic archaea	<i>Methanobrevibacter smithii</i>

Since approximately 95% of the SCFAs produced by the microbes during growth are absorbed by the host through the gut wall this represents a strong interaction between the microbes and the host. Indeed the ratio of the three main SCFAs (acetate, butyrate and propionate) is known to have a significant effect on human health [1,10]. Thus, we prioritize information on the values of these ratios in our model verification. Similarly approximately 90% of the water flowing into the colon is absorbed. Changes in the volume of water have a significant effect on the concentration of the molecules in the colon which in turn affects pH which then affects microbial growth, all of which are included in our model.

Owing to its shape within the body, the colon is commonly divided into three different regions—the proximal, transverse and distal sections running from beginning to end (figure 1a,b). The availability of substrate, microbial growth and hence pH vary along the colon; therefore, although our model is not spatial we simulate these three regions explicitly, with flow from one to another. Furthermore, as well as incorporating varying substrate inflow in the form of meals we also add in the release of mucins along the length of the colon which can be microbially broken down to release proteins and carbohydrates, allowing for further microbial growth away from the beginning of the colon where the substrates enter. A graphical summary of the model is shown in figure 2, the microbial functional groups are shown in table 1, and the model state variables are summarized in table 2.

We use the following criteria to verify our model captures the main features established for the system:

1. Total SCFA (TSCFA) concentration in the proximal, transverse and distal compartments should be around 123, 117 and 80 mM, respectively, according to sudden death human autopsies [11].
2. Acetate : propionate : butyrate ratios are similar (around 3 : 1 : 1) in all regions of the colon and around 60 : 20 : 20 mM [11].
3. Over 95% of SCFAs are absorbed by the host [12].

4. Approx. 90% of incoming water is absorbed by the host [13].
5. pH in the proximal, transverse and distal compartments should be around 5.7, 6.2 and 6.6, respectively [11] (and telemetry data from [14,15]).
6. Normal daily faecal output in Britain is 100–200 g d⁻¹ of which 25–50 g is solid matter (i.e. 50–175 g d⁻¹ is water). Bacteria make up about 55% of the solid matter, i.e. 14–28 g d⁻¹ of microbes emitted [16].
7. TSCFA concentration decreases with transit time [17].

After model verification, we examine the effects of including meals, bowel movements and fixed/varying pH into the model. We then use the model to look at how carbohydrate composition (based on the fractions of resistant starch (RS) and non-starch polysaccharides (NSP)) and total carbohydrate affect the microbial community and SCFA composition. The simulations are then compared with *in vivo* data from human volunteer experiments.

Although gut microbiota are highly complex and not fully understood, here we show that it is nonetheless possible to develop predictive models of key components of this ecological system. An important goal of our modelling is to aid and inform the interpretation of data obtained, mostly from faecal samples, in studies on diet and health in humans. Our results show promise and we believe this model represents a significant step forward in analysing this highly complex system. We refer to the model as ‘microPopGut’ and to aid future research the code is available as an R package on github (<https://github.com/HelenKettle/microPopGutCode>) and instructions on how to use the package are given in the electronic supplementary material, file ‘getStartedWithMicroPopGut.pdf’.

2. Results

2.1. Standard model

The model settings which give the best fit to our criteria are shown in table 4 (colon parameters and dietary inflow).

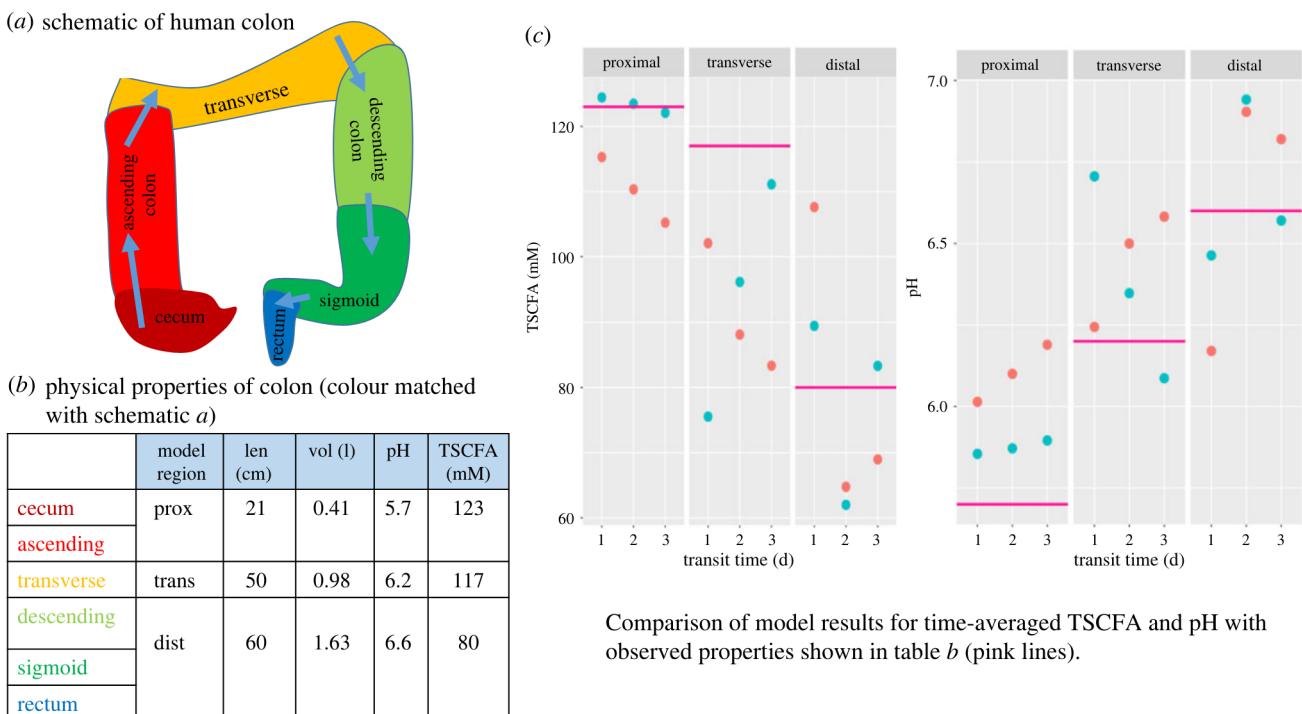


Figure 1. (a) Colon schematic, (b) table of typical values for physical properties (length, volume, pH and TSCFA) and (c) plots of summarized model simulations for average TSCFA and pH for comparison with typical values. Red dots show results from meals-inflow averaged over four random seeds; blue dots show results from continuous substrate inflow.

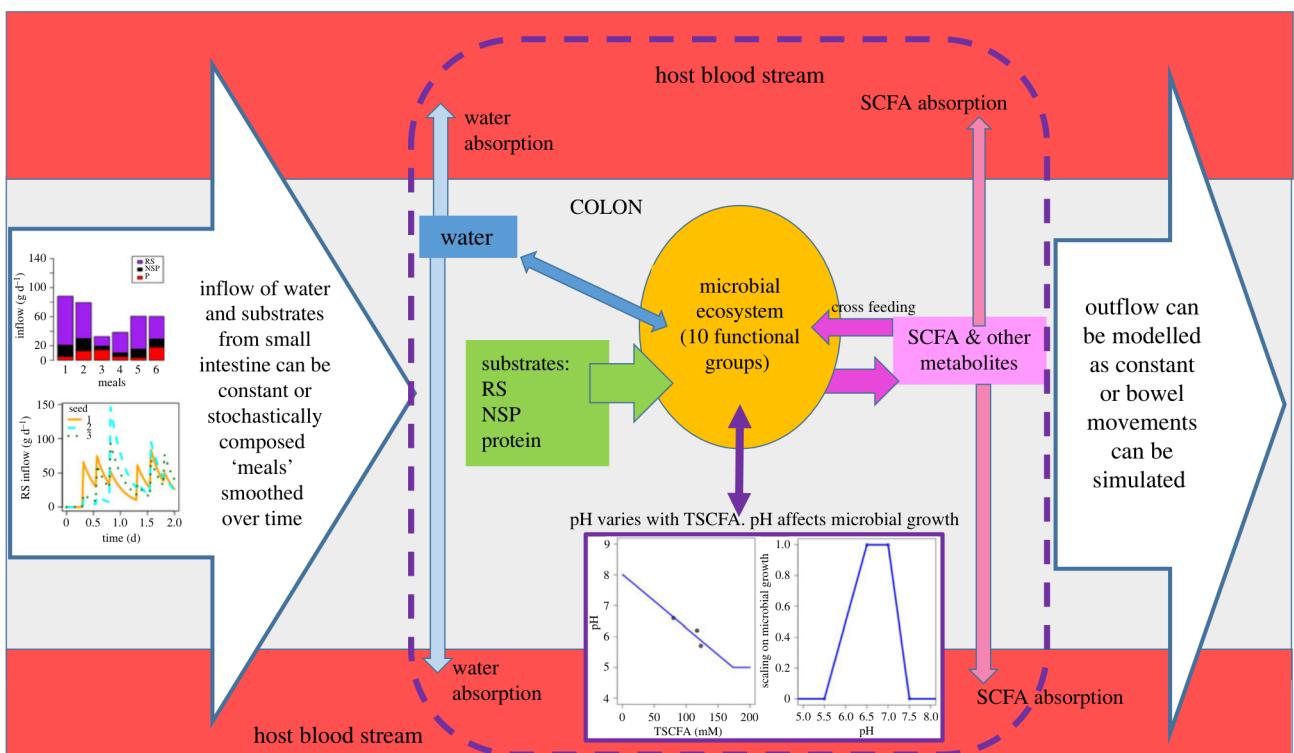


Figure 2. Model system with the microbial ecosystem comprising 10 microbial functional groups (table 1) which consume substrates (RS, NSP and protein) and water. The microbes produce metabolites some of which are consumed by other MFGs (cross-feeding). SCFA and water are absorbed through the colon wall (at different specific rates). The system shown within the dashed line is repeated in each of the modelled regions of the colon (proximal, transverse, and distal) with the contents of the previous region flowing into the next. The first compartment (proximal) has inflow from the small intestine—this can be constant inflow or simulated meals whose composition varies randomly in time. The third model compartment (distal) has outflow to stool which can be constant or evacuation via bowel movements can be simulated. pH varies with the TSCFA concentration and affects the rate of microbial growth differently for each MFG.

The microbial group parameters are listed in electronic supplementary material, §S3. These define our default model. From this we investigate the effects of different model configurations, e.g. with/without bowel movements, meals and

variable pH, for a range of transit times. Simulations with meals have a random component and therefore the model is run for a number of different starting seed values. Owing to the random fluctuations these simulations will not reach

Table 2. State variables included in the model. They are all in units of mass (g; with the exception of pH) and they are computed for each model compartment (e.g. prox., trans. and dist.). They are derived automatically from the substrates and metabolites specified for each microbial functional group (MFG) in the input file/dataframe to the R package microPop [9].

name	details
microbial biomass	computed for each of the 10 MFGs (table 1)
water	from dietary intake or from microbial metabolism
protein	from dietary intake or mucin
resistant starch (RS)	from dietary intake or mucin
non-starch polysaccharides (NSP)	from intake or mucin
pH	computed from TSCFA (equation (4.5))
acetate	metabolite
butyrate	metabolite
propionate	metabolite
formate	metabolite
carbon dioxide	metabolite
methane	metabolite
ethanol	metabolite
lactate	metabolite
succinate	metabolite
hydrogen	metabolite

steady state; therefore the summary values are taken as the mean from day 7 (to remove the effect of the initial conditions) to the end of the simulation (28 d) and are averaged over multiple seeds.

Table 3 gives summary results of the model simulations without bowel movements but with varying pH for each bowel region. Figure 3 shows results from more simulations but for the distal colon only. Figure 3a shows that although bowel movements make a difference to the total biomass and the TSCFA they do not have a large effect on the community composition or the SCFA ratios. Thus in the interests of model simplicity we decide to not include bowel movements in later simulations. However, varying pH with TSCFA can be seen to make a large difference to the microbial community (figure 3b) and also improves the SCFA ratios with respect to our verification criteria. The addition of meals makes a significant difference which increases with increasing transit time (figure 3c). In figure 4, the time series output from the model shows how the meals-inflow allows the community to experience large shifts over time (on a much longer time scale than the variations in the input), as opposed to the fixed state approached using a constant substrate inflow.

Figure 1c shows the average pH and TSCFA for the proximal, transverse and distal compartments. A decrease in TSCFA (and concomitant increase in pH) with longer transit time is predicted in the proximal colon both for meal inflow and continuous input and this is in broad agreement with experimental findings [17]. In the electronic supplementary material, §S2, we suggest a mathematical explanation for

this based on the supposition that the specific rate of absorption of water through the gut wall is slower than that for SCFA.

Regarding table 3, for some criteria, e.g. pH, the continuous inflow setting gives results closer to our verification values, but in other cases, e.g. A : B : P in distal colon, simulating meals gives closer results. Note that we consider a transit time of 1 d the most typical of the three transit times, and the one that should be compared with our verification criteria; the others are included to show the variation in results. Ideally TSCFA should be 123, 117 and 80 mM for prox., trans. and dist. but the best match we have to this is for a 3 d transit time and continuous inflow. This is most likely due to the fact that our model has fixed rates of specific absorption of SCFA and water throughout the colon. However, our TSCFA values are within a reasonable range and display the general trend of decreasing TSCFA from the proximal to distal colon. The microbe output, i.e. the outflow of faecal microbes, is steady at around 20 g d⁻¹ in all cases which fits well in the verification range (14–28 g d⁻¹). The water fraction is the ratio of the rate of faecal water over the rate of water flowing into the colon, and since 90% of water is absorbed this should be 0.1. This is approximately correct for our 1 d simulations (0.14) but, as expected, when transit time increases this decreases significantly. In summary, comparing these simulation results with our list of model verification criteria shows that in general our model is fit for purpose, and that the inclusion of meals-inflow and varying pH improve our simulations.

2.2. Model experiments

We now use our model to simulate two scenarios—firstly, the effects of decreasing total carbohydrate intake and secondly, the effects of changing carbohydrate composition (whilst keeping total intake fixed) on the microbial community and associated SCFA production. Comparing our simulations with data from human volunteer experiments is not straightforward since in order to run our model, ingested food must be translated to non-digestable substrates reaching the colon. This is problematic due to unknown water consumption and transit times and uncertainties associated with the absorption rates of the ingested carbohydrate and protein higher up the digestive tract. Thus we do not attempt to reproduce human experiments exactly but rather we run simulations based on variations to our standard model set up which are qualitatively similar, and then compare our results with the trends in the available data.

2.2.1. Effects of total dietary carbohydrate

In this model experiment, we investigate the effects of decreasing carbohydrate on the microbial community. Here we compare our results qualitatively with the human dietary study of Duncan *et al.* [18] which explored the impacts of carefully controlled decreases in carbohydrate intake upon weight loss and microbial fermentation products in obese subjects using three diets—a maintenance (M) diet, a high protein, moderate carbohydrate diet (HPMC) and a high protein, low carbohydrate diet (HPLC) (see figure 5 for details). This is, of course, the composition for ingested food, which is not easily translated into substrate concentrations entering the colon. However, we can look at the general trends in SCFA and microbial composition with changing colonic

Table 3. Summary of model results (for comparison with our list of criteria) for three different transit times, with meals or continuous inflow and with pH varying with TSCFA. Microbe output is the mass of microbes leaving the colon per day and the water fraction is amount of water leaving the colon per day divided by the amount entering. All simulations were run for 28 d and the results shown are the average over days 7–28. The results for the simulations with meals are averaged over four random seeds. ‘A : B : P dist’ refers to the acetate : butyrate : propionate ratio (mM) in the distal colon.

transit time	meals			continuous inflow		
	1 d	2 d	3 d	1 d	2 d	3 d
TSCFA prox (mM)	115.3	110.3	105.2	124.4	123.5	122.1
TSCFA trans (mM)	102.1	88.1	83.4	75.5	96.1	111.1
TSCFA dist (mM)	107.6	64.8	69.0	89.5	62.0	83.3
A : B : P dist (mM)	62 : 28 : 17	31 : 23 : 10	34 : 23 : 12	56 : 27 : 7	38 : 18 : 6	59 : 17 : 7
pH prox	6.0	6.1	6.2	5.9	5.9	5.9
pH trans	6.2	6.5	6.6	6.7	6.4	6.1
pH dist	6.2	6.9	6.8	6.5	6.9	6.6
microbe output (g d^{-1})	20.2	20.1	20.1	20.0	20.1	20.0
water fraction	0.14	0.04	0.02	0.14	0.04	0.02

Table 4. Summary of default values used in the model. Parameter values for the microbial groups are given in the electronic supplementary material, §S3.

symbol	description	default value
T_t	transit time through colon	1.25 d
\dot{P}_{diet}	protein inflow rate	10 g d^{-1}
\dot{C}_{diet}	carbohydrate inflow rate	50 g d^{-1}
\dot{W}_{diet}	water inflow rate	1100 g d^{-1}
M	mucin inflow rate	5 g d^{-1}
K_M	half saturation constant for mucin breakdown	0.5 g l^{-1}
a_W	rate of water absorption by host	3 d^{-1}
a_Z	rate of SCFA absorption by host	9.6 d^{-1}

carbohydrate intake rate. Thus, in these model experiments we keep protein inflow to the colon at 10 g d^{-1} (our default value) and then increase inflowing non-digestable (ND) carbohydrate from 10 g d^{-1} to 60 g d^{-1} in 10 g d^{-1} intervals. To include the effects of different ND carbohydrate composition, we run the model for an RS fraction of either 0.2 or 0.78 (the default value), with NSP making up the remaining ND carbohydrate in each case. Although subject to large uncertainties, we estimate the RS fractions for the [18] experiments of 0–0.6 (M diet), 0–0.68 (HPMC) and 0–0.12 (HPLC) (based on RS is 0–20% of ingested starch [20] and bio-available NSP is 75% of ingested NSP [21]). Owing to the low fibre nature of many of these simulations, we run the model with a slightly longer transit time of 1.5 d and for both continuous inflow and meals.

Figure 6 shows the SCFA results from our model experiment, and figure 5 shows the results from the *in vivo* experiment. It is clear from both the model and *in vivo* results that the proportion of butyrate increases as the amount of ND carbohydrate in the diet increases. Furthermore, both

model and *in vivo* results show an increase in TSCFA with ND carbohydrate intake rate. Since Duncan *et al.* [18] also looked at the relationship between butyrate concentration and grams of carbohydrate eaten per day, we plot butyrate against carbohydrate entering the colon (figure 7) to compare with their fig. 1. In both cases, butyrate concentration increases with incoming carbohydrate. Furthermore, as seen in both the model and the data, the percentage of butyrate increases with carbohydrate intake (figure 7). Analysis of 10 human studies involving 163 subjects has shown a highly significant increase in percentage butyrate with increasing total SCFA concentration in faecal samples [22].

In terms of microbial composition, figure 6 shows the results from our simulations are reasonably consistent across inflow type (meals or continuous), with B dominating at low carbohydrate intake. When the RS fraction is low (i.e. when ND carbohydrate is made up of 80% NSP) then NBSD increase with increased C intake. Whereas when C is mostly RS then NBSD and BP1 increase with C. In both cases, BP2 increase with increasing C intake.

2.2.2. Effects of carbohydrate composition

Here we use the model to simulate the effects of changing carbohydrate composition on the microbial community composition by changing the ratio of RS to NSP while keeping the same amount of total incoming carbohydrate. Figure 8 shows a summary of the model results. Although there are differences between the continuous inflow/meals, and also for the different transit times (1 d and 3 d), the modelled trends are generally similar, showing a significant shift in community as the fraction of RS increases, an increase in TSCFA and changes in the SCFA ratios. We compare our results with a human dietary study ([19,23] and references therein) examining the impact of switching the major type of ND carbohydrate from wheat bran (NSP) to resistant starch. Volunteers were provided successively with a maintenance diet, diets high in RS or NSPs and a reduced carbohydrate weight loss (WL) diet, over 10 weeks (figure 5).

There are large discrepancies between the SCFA predicted by our model (figure 8) and the measured SCFA data

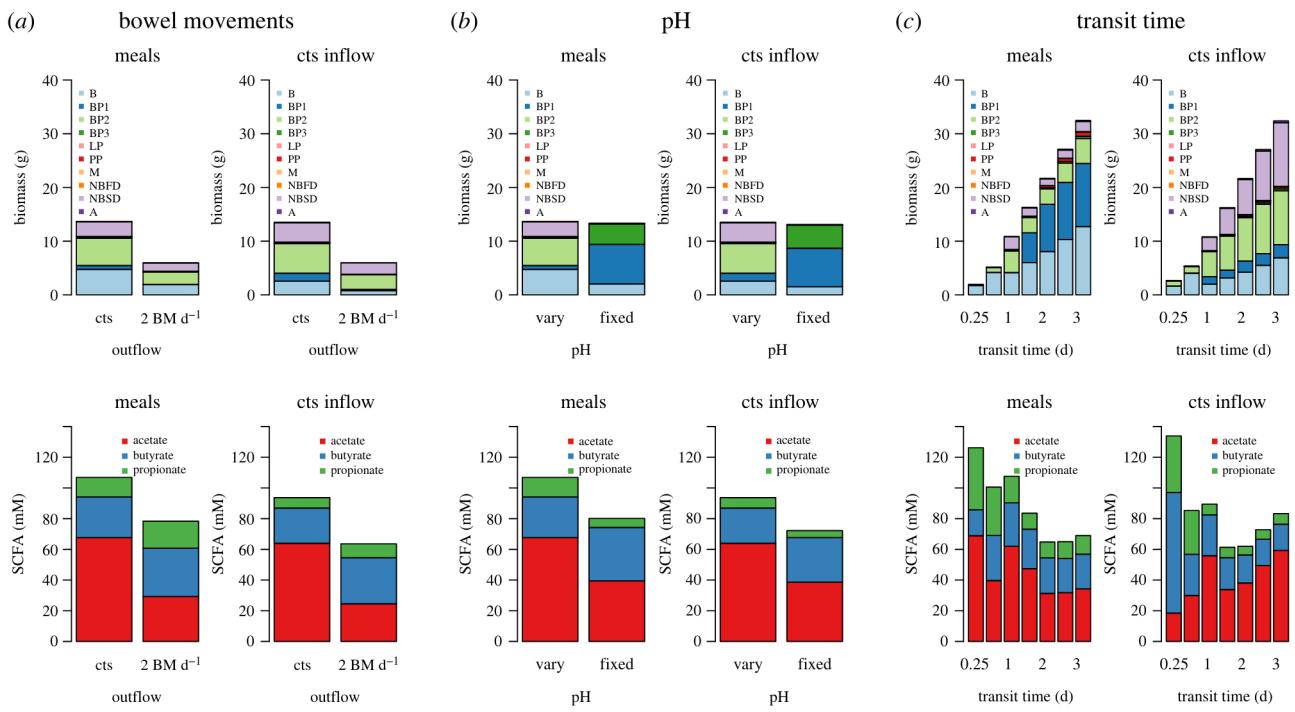


Figure 3. Summary results (averaged over days 7–28 and over random seeds) for the distal compartment for continuous inflow or fluctuating inflow (i.e. ‘meals’) for continuous outflow from colon or for two bowel movements per day (2 BM d^{-1}). The RS fraction is 0.78 (i.e. 78% of the dietary carbohydrate is RS and 22% is NSP) and the transit time is 0.93 d for (a), 1.25 d for (b) and at 0.25, 0.5, 1, 1.5, 2, 2.5 and 3 d for (c). The top row shows the biomass of each group, the bottom row shows the SCFA.

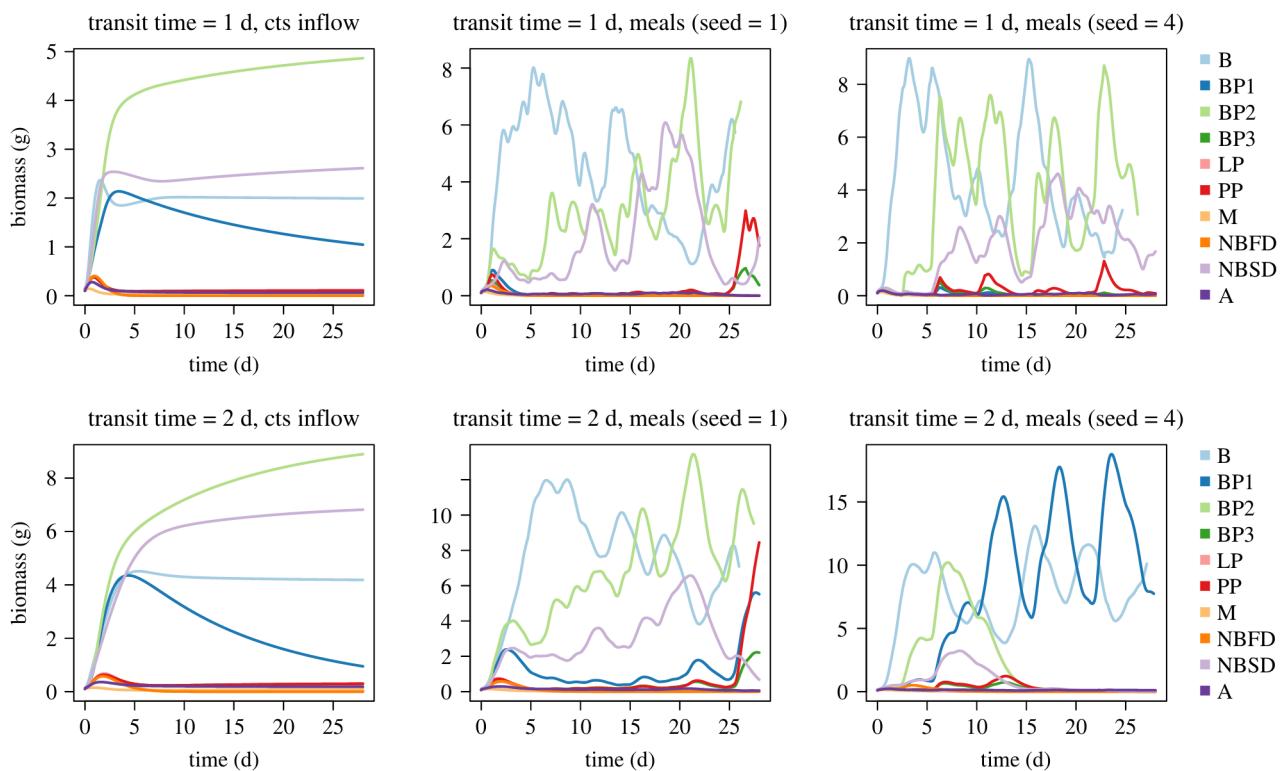


Figure 4. Simulation results for the distal compartment for continuous inflow (first plot in each row) or fluctuating inflow (i.e. ‘meals’) for transit times of 1 d (top row) and 2 d (bottom row) and for 2 random seeds. Modelled pH varies with TSCFA and the RS fraction is 0.78. There are no bowel movements (i.e. outflow is continuous). See table 1 for microbial groups.

(figure 5). Our model predicts an increase in TSCFA as proportion of RS increases whereas total faecal SCFA were significantly lower for the RS and WL diets compared to the other two diets (in which NSP is higher). One possible explanation is that fermentation of RS occurs in more proximal regions of the colon compared with NSP fibre fermentation,

such that there is greater absorption of the SCFA products. A second possibility, also likely, is that transit times were longer for the RS diet than for the NSP diet, which we predict would result in decreased SCFA concentrations. In our model the effect of the RS fraction on TSCFA is greater than the effect of transit time so we do not see this in figure 8.

diet	PI (g d ⁻¹)	CI (g d ⁻¹)	SI (g d ⁻¹)	NSPI (g d ⁻¹)
Duncan				
M	94	399	187	28
HPMC	127	164	95	12
HPLC	120	24	2.7	6.1
Walker				
M	103	427	230	28
high NSP	102	427	138	42
high RS	109	434	275	13
WL	144	201	110	22

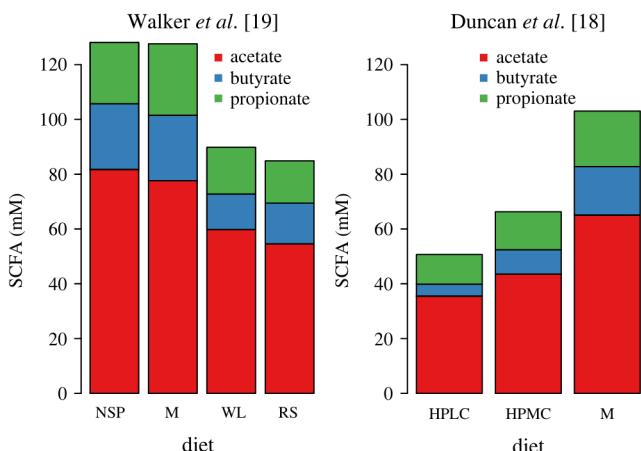


Figure 5. Table on left shows the dietary intake for two human studies [18,19]. PI, CI, SI and NSPI refer to ingested dietary protein, carbohydrate, starch and NSP. Note that starch value for the high RS diet in the [19] study included 26 g commercial RS. Bar plots show SCFA data from these studies. The bars in the plots have been ordered to show increasing RS fraction (estimated by SI/CI) for the Walker study (for comparison with figure 8) and increasing carbohydrate for the Duncan study (for comparison with figure 6).

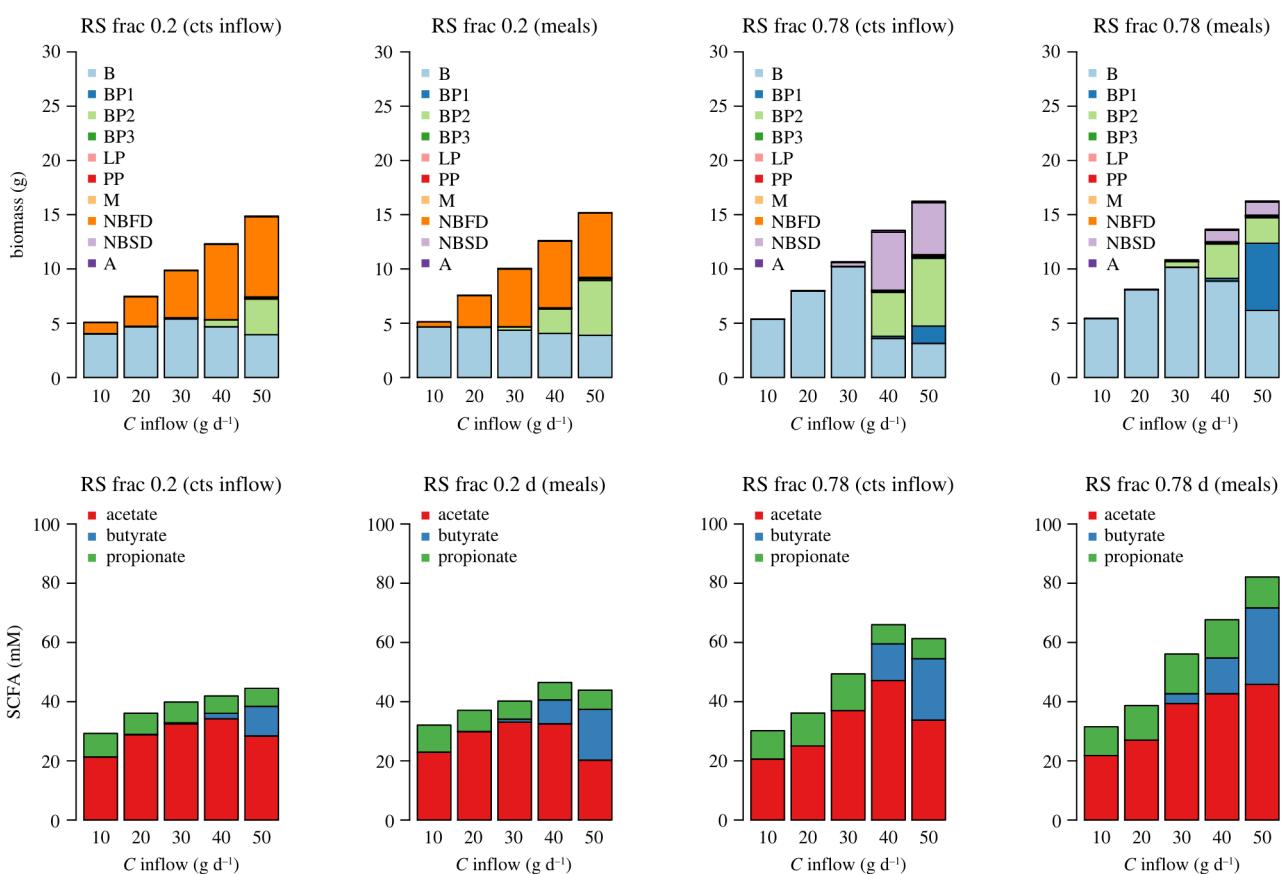


Figure 6. Simulated biomass and SCFA results for increasing carbohydrate inflow. Simulations are run with continuous substrate inflow (cts) and with 'meals' for a transit time of 1.5 d. The results are the average over the last three weeks of a 28 d simulation and 'meals' is the average over four stochastically generated simulations.

The human study also included detailed compositional analysis of the faecal microbiota [19,23] that revealed specific responses mainly by different groups of Firmicutes bacteria to the RS and NSP diets. This information was particularly important for the phylogenetic assignments to the functional groups used here and previously in the model of [8]. Our modelling predicts striking shifts in the microbial community, especially involving the NBSD, NBFD and butyrate-producing groups, with changing proportions of RS and NSP fibre (figure 8). We should also note that in the volunteer experiments many bacterial species were not significantly

altered by the RS–NSP switch *in vivo* [19] indicating that many may be generalists, able to switch quickly between energy sources.

3. Discussion

The development of a complex model of the microbial community in the human colon, whose simulations compare well with data, represents a significant step forward. Previous models have been based on simpler microbial models (e.g.

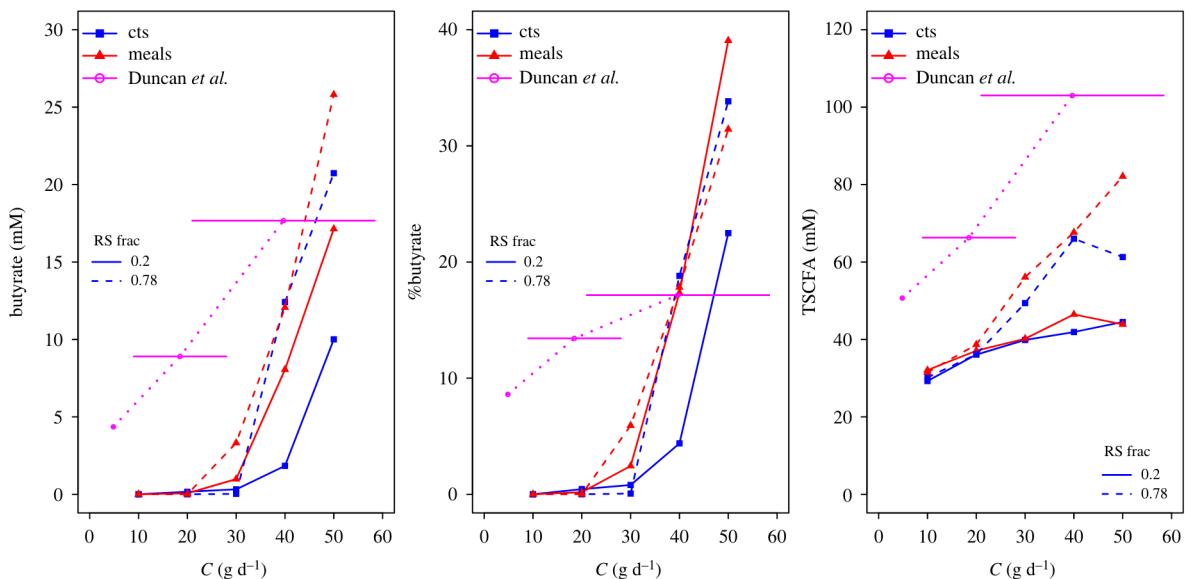


Figure 7. Plot of modelled butyrate, %butyrate and TSCFA against grams of carbohydrate entering the colon each day. Data from [18] are shown in magenta—due to uncertainties in converting ingested starch to RS entering the colon there are large error bars on the amount of C (g d^{-1}). Error bars show C estimated by the sum of 75% of ingested NSP plus 0–20% of ingested starch.

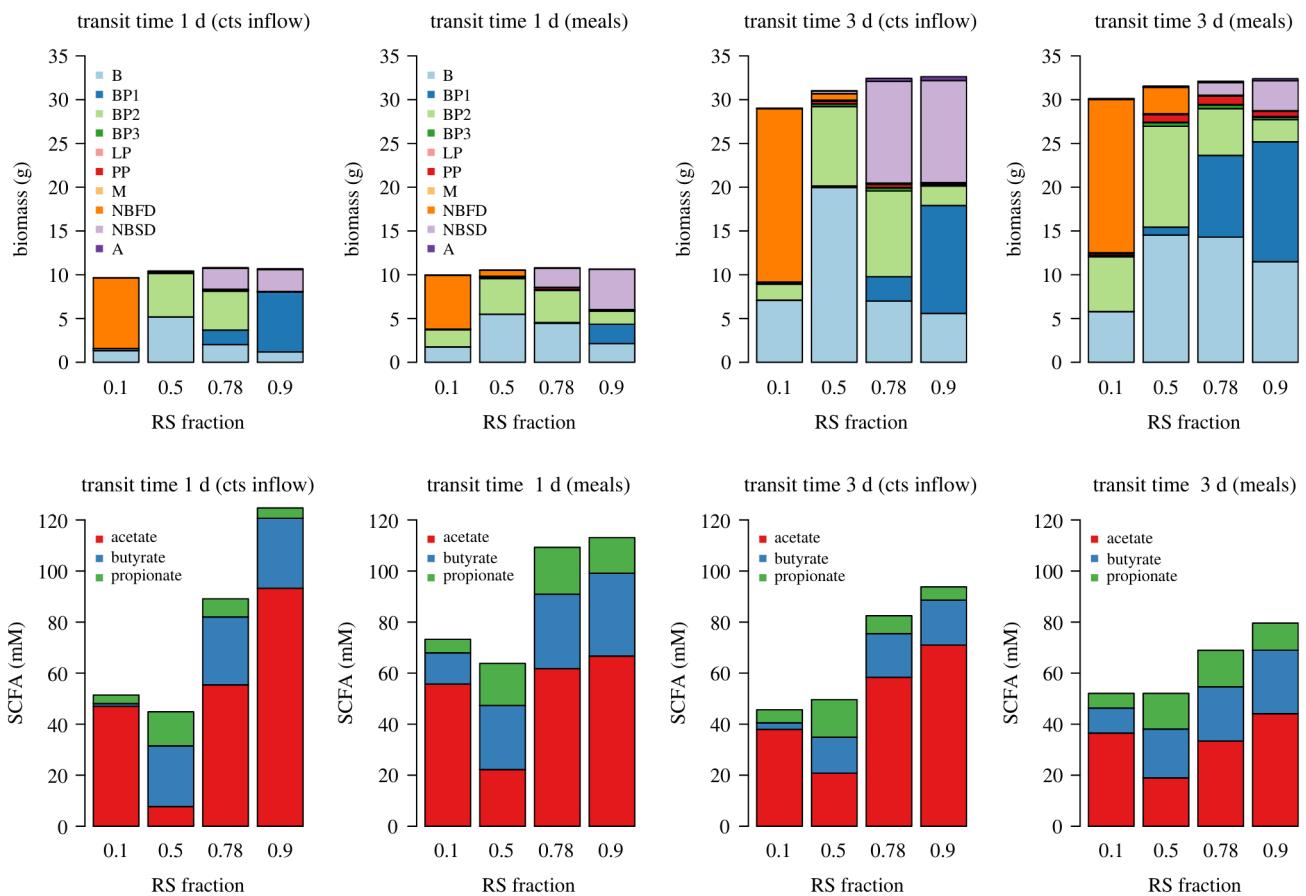


Figure 8. Biomass and SCFA results for changing the RS fraction of inflowing carbohydrate with continuous substrate inflow (cts inflow) and with ‘meals’. Protein and carbohydrate inflows are 10 and 50 g d^{-1} , respectively. The results are the average over the last three weeks of a 28 d simulation and ‘meals’ is the average over four stochastically generated simulations.

[3,5,6]), or have not shown such a good agreement with data (e.g. [7]). Our previous complex model community consisted of 10 functional groups, but the model was designed only to simulate continuous culture conditions in a chemostat [8]. Translating this 10-group model into an *in vivo* setting has required introducing multiple gut compartments, and the absorption of water and SCFA, followed by comparison

with generally observed characteristics of the system. We were then able to use this model to examine the predicted impact of changes in the amount and type of non-digestible carbohydrate (fibre) present in human diets upon concentrations of fermentation products (SCFA) in different gut compartments and in stool. At the same time, we predict the likely impact of dietary changes and variations in gut

transit upon microbiota composition and fermentation products. The model must be regarded as work in progress particularly with respect to microbiota composition. Predictions can however become improved and refined as more information becomes available in time.

Assignments of microbial taxa to our 10 functional groups were based initially on evidence from cultured isolates. These assignments have since been supported and greatly extended by analysis of genes diagnostic for different fermentation pathways within genomes and metagenomes [24] and by molecular detection of species enriched within the community by defined growth substrates in chemostat experiments [25] and dietary intervention studies [23]. Nevertheless, these assignments inevitably remain provisional and incomplete and we do not claim that the model predictions can be made precise at a phylogenetic level. More emphasis is placed in our model on the prediction of metabolic outputs based on microbial transformations and interactions. While there is relatively little phylogenetic overlap for example between producers of propionate and butyrate [24,26], there are many cases where individual species are known to use multiple alternative substrates as energy sources, which complicates assignments. For this reason, more weight was given to fermentation pathways than to substrate preferences in defining the functional groups. However, it would also be possible to define completely different groupings that relate to other outputs (e.g. bile acid metabolism, or vitamin/micronutrient supply) in order to address specific questions. Furthermore, it may well be worthwhile to increase the number of functional groups in the future. The large B group for example currently includes members of the Bacteroidetes phylum, but its characteristics are mainly based on well-studied members of the *Bacteroides* genus. We know that *Prevotella* is another highly abundant genus of Bacteroidetes in the human colon, but the two genera tend not to co-occur at high levels in the same individuals [27,28]. Less is known about human colonic *Prevotella*, for which there are relatively few cultured representatives, making it premature to create a separate grouping, but this would clearly be desirable in the future as their prevalence is reported to affect health and responses to dietary intervention. In future, it should become possible to define the relative abundance of functional groups (MFGs) and their relationship to phylogeny directly from genomic and metagenome analysis, by examining genes diagnostic for particular pathways and functions (e.g. [24]).

The parameter values for the microbial groups used in our model are from the intrinsic data frames in the microPop package (the only changes are to LactateProducers). Although the work presented here did not attempt to fit particular parameters to data, as we focused on expanding the scope of the model (i.e. changing the environment from fermentor to colon), these values are easy to alter; e.g. Wang *et al.* [29] changed many of these parameters to achieve a better model fit to their data. As well as adjusting the parameters for each group to represent inter-individual variation, groups can also be easily added or removed from the model through the input argument ‘microbeNames’. Furthermore, it is also possible to include any number of strains (with varying parameter values) within each functional group in order to add more variation in outcome [8] but we did not do this here in the interests of computational time. It should also be noted that the parameter values are

highly uncertain in many cases and within each of our functional groups there will be large variability due to adaptation and evolution. Given this, we do not claim that the model response is necessarily representative of what may happen in an individual’s gut, rather it can be used as an aid to gain insight into the relative importance of the different processes we are currently aware of and potentially to highlight those we are not.

In addition to this, it must be noted that the default diet chosen here with 10 g of protein and 50 g of carbohydrate fibre reaching the colon each day could be revised for any given population. However, converting from quantities of ingested food to substrate inflow to the colon is highly uncertain with large variations between studies, as well as technical issues with measuring this accurately. With more time, it would be interesting to investigate a larger range of typical diets but this was beyond the scope of the current work.

In summary, although performing reasonably well, the model has the potential to be considerably improved simply by altering the parameter values and existing settings; however, more fundamental changes such as those listed below could also be investigated in future work:

- Adding more functional groups or pathway switches in the existing functional groups. For example at present only the *Bacteroides* group can use protein but it is now known that some butyrate producers can also use amino acids [26].
- Our pH relation with TSCFA is very simplistic and could potentially be improved, although host secretions mean this is not necessarily straightforward.
- Currently, we set the transit time for the colon and then this is split between the three model compartments based on their relative sizes. An interesting addition would be to alter transit time based on the composition of the various substrates entering the colon. For example, increasing residence time for high protein and/or low fibre diets. Owing to variation in individual response this may need to include significant uncertainty ranges.
- Related to this is changing the absorption rate of water through the gut wall based on the diet, for example, more water could remain in the gut on a high fibre diet.
- A longer term goal would be to model the processes in the gastrointestinal tract preceding the colon in order to simulate how substrates entering the colon relate to dietary intake. This would allow more accurate prediction of microbial metabolite production based on diet.

To conclude, our model helps to explain some important, but poorly understood, relationships that have been reported in human studies, including the increase in butyrate proportion with increasing total faecal SCFA [22]. This phenomenon has important implications in view of the claimed benefits of butyrate supply for colorectal cancer prevention and the health of the colonic mucosa [10,30]. The model also predicts increasing total faecal SCFA with greater fibre intake and more rapid gut transit. Gut transit is also shown to have potentially important consequences for microbiota composition and gut metabolism. In addition, the model confirms that the amount and type of non-digestible carbohydrate in the diet has the potential to cause major changes in microbiota composition. The nature of such changes is, however, predicted to be influenced by patterns

of meal feeding and by any effects of dietary components (e.g. dietary fibre) upon gut transit. Human studies suggest that they will also depend on the initial microbiota composition. There is potential to use the model to explore how the presence of particular functional groups (such as lactate-users [29]) within an individual's microbiota can influence their gut metabolism and response to dietary intervention. This may indeed be one of the most intriguing and fruitful applications of such modelling approaches in the future.

4. Material and methods

4.1. Software

To facilitate continued research and future model development by other researchers, we provide all model code on github (<https://github.com/HelenKettle/microPopGutCode>). The R package microPopGut is contained in the file microPopGut_1.0.tar.gz. This can be downloaded and installed in R using install.packages ('microPopGut_1.0.tar.gz'). Furthermore instructions on how to use the package are given in the electronic supplementary material, file 'gettingStartedWithMicroPopGut.pdf'.

4.2. Microbial model

The microbial functional group model is based on the model described by Kettle *et al.* [8] and implemented using the R package microPop [9]. The microbial groups include producers of the three major SCFAs detected in faecal samples (acetate, butyrate and propionate) together with utilizers of acetate, lactate, succinate, formate and hydrogen (see table 1 for a summary, or refer to Kettle *et al.* [8] for more detail). The model and its equations are described in detail by Kettle *et al.* [8,9] so only a brief overview is given here. The microbial groups are defined as data frames within the R package and these are shown in the electronic supplementary material, §S3. Although this application uses the microbial parameters (e.g. maximum growth rates, yields, etc.) that are in the package's intrinsic data frames, these can be easily changed by either modifying the dataframe in R or by providing a new dataframe—either as an input csv file or by creating one in R. One of the input arguments to the function *microPopGut()* is *microbeNames* which allows the user to also enter other microbial groups.

The growth substrates available in the large intestine are divided into four categories: protein (*P*), NSP, RS and sugars (and oligosaccharides and sugar alcohols); for simplicity, all carbohydrate units are regarded as being hexoses. NSP comprise major components of dietary fibre including the structural polysaccharides of the plant cell wall (cellulose, xylan, pectin), whereas RS refers to the fraction of dietary starch that resists digestion in the small intestine. We consider 10 major metabolites that arise from substrate fermentation: acetate, propionate, butyrate, lactate, succinate, formate, hydrogen, carbon dioxide, methane and ethanol. Six of these metabolites (acetate, lactate, succinate, formate, hydrogen and carbon dioxide) are also considered as substrates, because they are known to be consumed by some groups (cross-feeding). It is well known that pH affects growth rate and therefore each group is assigned a preferred range of pH within which it can reach its maximum growth rate, but outside of which, its growth is reduced or zero. We model the rate of bacterial growth using Monod kinetics and assume that from 1 g of resource, Y g of biomass is produced. We assume that resource that is taken up by microbes, but not used to produce biomass, is converted to metabolites. If not all of the resource is converted to biomass or to the metabolites represented in our model, it is discarded. This applies, for example, to many diverse fermentation products of proteins (e.g. phenols, amines) that are not among the 10 major products covered by the model. Although the model was initially developed to be run

with multiple strains within each functional group, in the current work we do not do this due to the high CPU time associated with multiple compartments.

4.3. Inflow to colon

4.3.1. Incoming substrates and water

The main sources of nutrient for microbiota in the colon are complex dietary carbohydrates that are not absorbed higher up the digestive tract. We use a default value of 50 g d⁻¹ of carbohydrate, *C*, in our model and we vary the proportion of this which is NSP or RS using the RS fraction (i.e. RS/(RS + NSP) where RS + NSP = *C*). Based on [3] and references therein, about 15 g of bio-available NSP and 30–40 g of RS enter the colon per day which gives us an RS fraction of 0.67–0.9 with average value of 0.78 which we use as our default value. According to [31] less is known regarding dietary proteins, *P*, that escape digestion to reach the large intestine, although it is estimated that around 6–18 g *P* reaches the large intestine daily, the majority from the diet and a small proportion from endogenous origins. Given this, here we assume that 10 g d⁻¹ of undigested *P* reaches the colon from dietary intake along with a small amount from mucin degradation (approx. 1 g d⁻¹). Phillips & Giller [13] state that water enters at approximately 1.5 l d⁻¹ and about 90% of this is absorbed by the colon. Stephen & Cummings [16] state that normal faecal daily output in Britain is 100–200 g d⁻¹ of which 25–50 g d⁻¹ is solid matter and the rest (50–175 g d⁻¹) is water. Thus if 90% is absorbed, then this indicates water inflow in the range 0.5–1.75 l d⁻¹. The midpoint of this range is 110 g d⁻¹ of water outflow which, if 90% is absorbed, implies that the inflow of water is approximately 1100 g d⁻¹. This will clearly vary depending on the host's oral water intake but we use 1100 g d⁻¹ as our default value. The default inflow values are summarized in table 4.

4.3.2. Meals

The normal human diet does not consist of continuous fixed inflow of substrate; for a more realistic substrate inflow to the colon we simulate eating three meals a day with randomly varying composition. We then approximate the passage of these meals through the stomach and small intestine to obtain a smoothed time series for substrate entering the colon. Note that we are not simulating all the food ingested by the host (most of which will not reach the colon) but rather simply trying to produce a more realistic time series for the substrates that we know reach the colon.

We specify three meals per day each with a duration of 30 min. This time series is then passed through a one-compartment ordinary differential equation model representing the time spent in the stomach and small intestine (estimated to take 7 h), i.e.

$$\frac{ds(t)}{dt} = s(t)_{\text{in}} - vs(t), \quad (4.1)$$

where *v* = 3.4 d⁻¹ (inverse of 7 h transit time in days), *s(t)_{in}* is time series representing three meals a day (g d⁻¹) and *t* is time in days. The inflow to the colon (i.e. the outflow from small intestine) is given by *vs(t)*. The composition (in terms of *P*, NSP, RS and water (*W*)) of these meals varies randomly around the mean of each component (table 4) for each meal. To generate such random fluctuations, we draw samples for each meal from a gamma distribution (since this is always above zero) defined by a scale parameter (γ_s) and the daily average inflow of the substrate (g d⁻¹). We assume the magnitudes of the substrate fluctuations are proportional to the mean value. Preliminary simulations showed that γ_s equal to half the mean value of each substrate gave a good variation for *P*, RS and NSP, and for water variation we assumed γ_s was one-tenth of the incoming daily flow. The distributions and flow patterns are shown in figure 9.

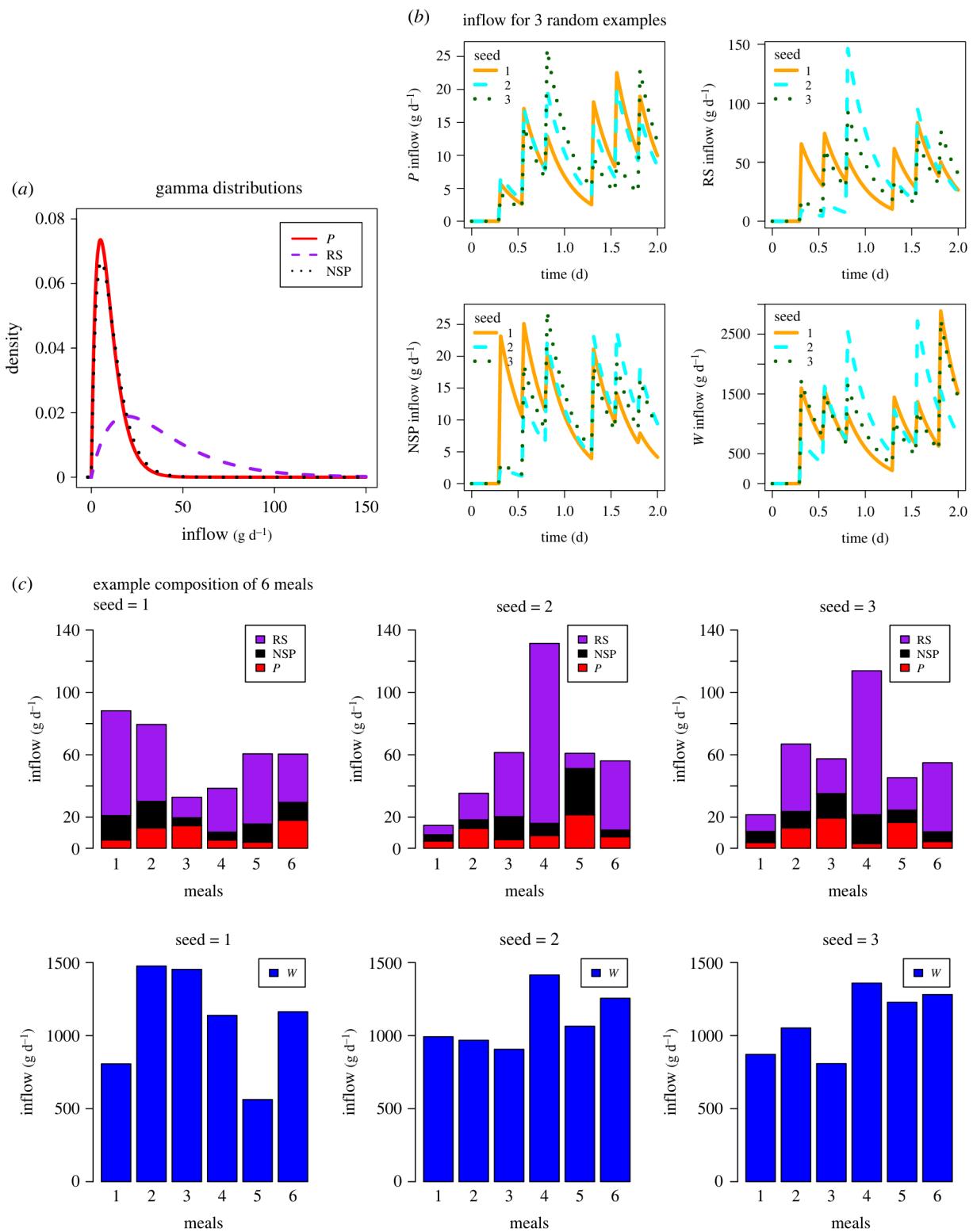


Figure 9. (a) Gamma distribution from which random values are drawn to generate the composition of each meal (note water is not shown due to the large difference in magnitude between water and dietary substrates). (b) The substrate inflow time series to the proximal colon after passing through the small intestine. Examples shown for three stochastic simulations starting with different seeds. (c) Barplots showing the composition of six meals over 2 days for three different stochastic simulations.

4.3.3. Mucin

There is a further input of protein and carbohydrate from the host via the breakdown of host-released mucin by many strains in the B group [32] and in our NBFD group [33]. It is estimated that $2.7\text{--}7.3 \text{ g d}^{-1}$ of mucin, denoted \dot{M} , is secreted into the colon [34]; therefore we take the midpoint value 5 g d^{-1} . We assume our mucin degraders break down 1 g of mucin into 0.05 g sulfate, 0.2 g P and 0.75 g C, based on [35], but consider their yield on mucin to be negligible compared with

growth on other substrates. We split C equally between NSP and RS—this arbitrary choice did not affect model results since C from mucin (3.75 g d^{-1} maximum) is much less than dietary C (50 g d^{-1}), but this should be revised if considering very different dietary drivers. Since the compartments of the colon are not equal-sized we assume that the rate of mucin entering the colon is divided through the model compartments proportional to their relative volumes. We assume this enters the colon at a fixed, continuous rate and mucin-derived P and C

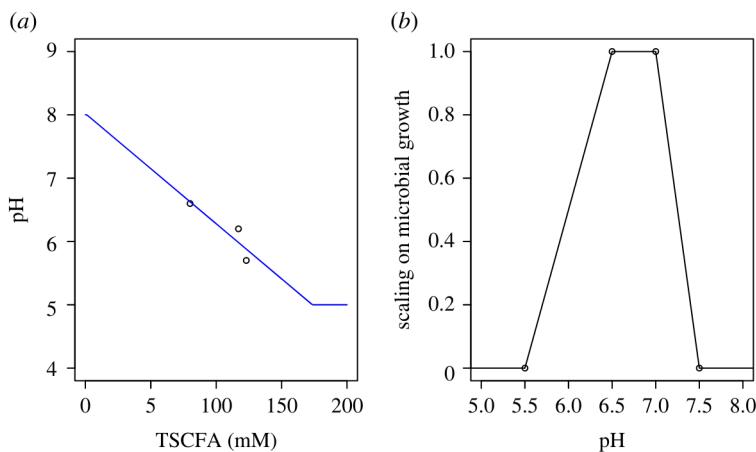


Figure 10. (a) Relating pH to TSCFA using equation (4.5) and data from [11]. (b) Example of microbial tolerance to pH. A pH tolerance function of this form is specified individually for each microbial group in our model.

are a function of the mass of mucin degraders, D_M (B and NBFD), such that

$$C(t) = 0.75 \frac{D_M(t)}{D_M(t) + K_M W_v(t)} \dot{M} \quad (4.2)$$

and

$$P(t) = 0.2 \frac{D_M(t)}{D_M(t) + K_M W_v(t)} \dot{M}, \quad (4.3)$$

where C , P , D_M are in mass units and the overdot indicates a rate (e.g. g d^{-1}), t is time and W_v is the volume of water in the model compartment. K_M (g l^{-1}) is chosen such that if $D_M \ll K_M W_v$ then there is minimal breakdown of mucin and if $D_M \gg K_M W_v$, there is maximal breakdown. The smaller the value of K_M the more breakdown there will be at low concentrations of mucin degraders. We set $K_M = 0.5 \text{ g l}^{-1}$ based on preliminary model simulations.

4.4. Absorption by host

SCFA and water are both absorbed by the host through the gut wall; over 95% of SCFA [12] and approximately 90% of incoming water is absorbed [13]. Experiments by Ruppin *et al.* [36] found the absorption rates of SCFA to be approximately 0.4 h^{-1} (i.e. 9.6 d^{-1}) with little difference in rates between the different SCFAs [12,36].

We can estimate mathematically the specific water absorption rate required to give 90% absorption of inflowing water for a given number of compartments in the colon (N) and a given transit time, T_t , using

$$a_W = \frac{16.95 - 9.72N + 1.77N^2}{T_t} \quad (4.4)$$

(see electronic supplementary material, §S1.3 for the derivation). As a rough estimation, a three-compartment model with a transit time 1–1.5 d gives $a_W \approx 3 \text{ d}^{-1}$ (electronic supplementary material, figure S1a). Given this will not be significantly affected by the microbial model (microbial uptake/production of water is small) this is a robust estimation.

To estimate the value of the specific absorption rate of SCFA, a_Z , we used a simple model (see electronic supplementary material, §§S1.1 and S1.4). Estimating the value of the specific absorption rate of SCFA based on the values of SCFA given in the verification criteria and given our estimate for a_W we found that it was necessary for the specific absorption rate to change along the colon (see electronic supplementary material, §S1.4). The best estimates were given by a_Z values of 25.2, 4.2 and 9.2 d^{-1} in the proximal, transverse and distal colon respectively.

However, in the interests of a robust model (i.e. the fewer parameter values, the better) we made the decision to use one value for a_Z . Since the experimental value of 9.6 d^{-1} compares well with our estimate in the distal colon we set $a_Z = 9.6 \text{ d}^{-1}$ throughout. It should be noted though that our model results could potentially be improved by varying a_Z between model compartments.

4.5. pH

Calculating pH in our model is not straightforward due to a lack of necessary state variables as well as pH buffering via secretions from the host. However, observations tell us the pH in the colon goes from 5.7 in the proximal, 6.2 in the transverse and 6.6 in the descending colon and TSCFA in these regions is around 123 mM, 117 mM and 80 mM, respectively [11]. Therefore, an approximate approach is to simply make pH a function of TSCFA. Fitting a line through the above points gives us the following relationship:

$$\text{pH} = 8.02 - 0.0174 \times \text{TSCFA}, \quad (4.5)$$

which we further limit by setting the minimum and maximum pH values at 5 and 8, respectively, i.e. if the TSCFA values give predicted pH outside of this range (figure 10).

The impact of pH on microbial growth is modelled via a pH limitation function whereby there is a range over which there is no limit on growth but outside of this range the growth rate decreases linearly to reach zero at the specified outer limits. Thus, there are four parameters used to describe the pH tolerance—two for the inner range where there is no limit on growth and two for the outer range outside which there is no growth—an example is shown in figure 10. The pH tolerance range for each microbial group is specified under the entry ‘pHcorners’ in the data frame for each group and shown in electronic supplementary material, §S3.

4.6. Faecal outflow

Faecal outflow (g d^{-1}) at time, t , is given by $m_d(t)V_d$ where $m_d(t)$ is the mass in the distal colon (i.e. microbes, unconsumed substrate, microbial metabolites and water) and V_d is the specific wash out rate from the colon (the inverse of the time spent in the distal colon). For continuous outflow (as is used in most gut models) we compute the specific wash out rate from each compartment by assuming the fraction of time spent in compartment i is proportional to its volume fraction, thus

$$T_i = \frac{v_i}{v_{\text{colon}}} T_t, \quad (4.6)$$

where v_i is the volume of compartment i and v_{colon} is the total volume of the colon. The specific wash out rate is then $V_i = 1/T_i$.

If we introduce bowel movements then, assuming the distal colon is approximately emptied for each bowel movement, the total transit time is given by

$$T_t = \sum_{i=1}^2 T_i + \frac{1}{N_{BM}}, \quad (4.7)$$

where N_{BM} is the number of bowel movements per day. For example, using volume measurements (figure 1b) and assuming a total transit time of 1 d would mean about 45% of the transit time is spent in the proximal and transverse colon and about 55% of the day spent in the distal, which would be similar to two bowel movements per day. In model experiments where we vary the number of bowel movements per day we also change the time spent in the rest of the colon since we assume increased bowel movements are indicative of a general increase in passage rate. We estimate the wash out rate from the colon during a bowel movement, V_{BM} , by

$$V_{BM} = -\frac{\ln(f_d)}{\Delta t_{BM}}, \quad (4.8)$$

where f_d is the fraction of mass left in the distal colon after the bowel movement and Δt_{BM} is the time taken for the bowel movement (d). For example, if a bowel movement takes 10 min to

remove 90% of the contents of the distal colon then V_{BM} is 332 d⁻¹. This is not affected by the number of bowel movements per day.

Data accessibility. All model code is on github (<https://github.com/HelenKettle/microPopGutCode>). The R package microPopGut is contained in the file microPopGut_1.0.tar.gz. This can be downloaded and installed in R using install.packages ('microPopGut_1.0.tar.gz'). Furthermore, instructions on how to use the package are given in the electronic supplementary material, file 'getStartedWithMicroPopGut.pdf' [37]. The model output for the simulations described in this paper are included on figshare in the file ModelRuns.tar.gz (<https://doi.org/10.6084/m9.figshare.21094558.v1>). The plotting code is provided in the github repository.

Authors' contributions. H.K.: conceptualization, formal analysis, methodology, software, validation, visualization, writing—original draft; P.L.: conceptualization, methodology, writing—original draft, writing—review and editing; H.J.F.: conceptualization, methodology, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. The Scottish Government's Rural and Environment Science and Analytical Services Division (RESAS) funded this research.

References

- Morrison DJ, Preston T. 2016 Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes* **7**, 189–200. (doi:10.1080/19490976.2015.1134082)
- Rios-Covian D, Ruas-Madiedo P, Margolles A, Gueimonde M, Salazar N. 2016 Intestinal short chain fatty acids and their link with diet and human health. *Front. Microbiol.* **7**, 185. (doi:10.3389/fmicb.2016.00185)
- Cremer J, Arnoldini M, Hwa T. 2017 Effect of water flow and chemical composition in the human colon. *Proc. Natl Acad. Sci. USA* **114**, 6438–6443. (doi:10.1073/pnas.1619598114)
- Cremer J, Segota I, Arnoldini M, Sauls J, Zhang Z, Gutierrez E, Groisman A, Hwa T. 2016 Effect of flow and peristaltic mixing on bacterial growth in a gut-like channel. *Proc. Natl Acad. Sci. USA* **113**, 11 414–11 419. (doi:10.1073/pnas.1601306113)
- Moorthy AS, Brooks SP, Kalmokoff M, Eberl HJ. 2015 Continuous model of carbohydrate digestion and transport processes in the colon. *PLoS ONE* **10**, e0145309. (doi:10.1371/journal.pone.0145309)
- Munoz-Tamayo R, Laroche B, Walter E, Dore J, Leclerc M. 2010 Mathematical modelling of carbohydrate degradation by human colonic microbiota. *J. Theor. Biol.* **266**, 189–201. (doi:10.1016/j.jtbi.2010.05.040)
- Smith N, Shorten P, Altermann E. 2021 Examination of hydrogen cross-feeders using a colonic microbiota model. *BMC Bioinf.* **22**, 3. (doi:10.1186/s12859-020-03923-6)
- Kettle H, Louis P, Holtrop G, Duncan SH, Flint HJ. 2015 Modelling the emergent dynamics and major metabolites of the human colonic microbiota. *Environ. Microbiol.* **17**, 1615–1630. (doi:10.1111/1462-2920.12599)
- Kettle H, Holtrop G, Louis P, Flint HJ. 2018 microPop: modelling microbial populations and communities in R. *Methods Ecol. Evol.* **9**, 399–409. (doi:10.1111/2041-210X.12873)
- Louis P, Hold GL, Flint HJ. 2014 The gut microbiota, bacterial metabolites and colorectal cancer. *Nat. Rev. Microbiol.* **12**, 661–672. (doi:10.1038/nrmicro3344)
- Cummings JH, Pomare E, Branch W, Naylor C, Macfarlane G. 1987 Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* **28**, 1221–1227. (doi:10.1136/gut.28.10.1221)
- Topping D, Clifton P. 2001 Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiol. Rev.* **81**, 1031–1064. (doi:10.1152/physrev.2001.81.3.1031)
- Phillips S, Giller J. 1973 The contribution of the colon to electrolyte and water conservation in man. *J. Lab. Clin. Med.* **81**, 733–746.
- Bown R, Gibson J, Sladen G, Hicks B, Dawson A. 1974 Effects of lactulose and other laxatives on ileal and colonic pH as measured by a radiotelemetry device. *Gut* **15**, 999–1004. (doi:10.1136/gut.15.12.999)
- Mikolajczyk AE, Watson S, Surma BL, Rubin DT. 2015 Assessment of tandem measurements of pH and total gut transit time in healthy volunteers. *Clin. Transl. Gastroenterol.* **6**, e100. (doi:10.1038/ctg.2015.22)
- Stephen A, Cummings J. 1980 The microbial contribution to human fecal mass. *J. Med. Microbiol.* **13**, 45–56. (doi:10.1099/00222615-13-1-45)
- Lewis S, Heaton K. 1997 Increasing butyrate concentration in the distal colon by accelerating intestinal transit. *Gut* **41**, 245–251. (doi:10.1136/gut.41.2.245)
- Duncan SH, Belenguer A, Holtrop G, Johnstone A, Flint HJ, Lobley GE. 2007 Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces. *Appl. Environ. Microbiol.* **73**, 1073–1078. (doi:10.1128/AEM.02340-06)
- Walker AW et al. 2011 Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J.* **5**, 220–230. (doi:10.1038/ismej.2010.118)
- Capuano E, Oliviero T, Fogliano V, Pellegrini N. 2018 Role of the food matrix and digestion on calculation of the actual energy content of food. *Nutr. Rev.* **76**, 274–289. (doi:10.1093/nutrit/nux072)
- Slavin JL, Brauer PM, Marlett JA. 1981 Neutral detergent fiber, hemicellulose and cellulose digestibility in human subjects. *J. Nutr.* **111**, 287–297. (doi:10.1093/jn/111.2.287)
- LaBouyer M et al. 2022 Higher total faecal short-chain fatty acid concentrations correlate with increasing proportions of butyrate and decreasing proportions of branched-chain fatty acids across multiple human studies. *Gut Microbiome* **3**, E2. (doi:10.1017/gmb.2022.1)
- Salonen A et al. 2014 Impact of diet and individual variation on intestinal microbiota composition and fermentation products in obese men. *ISME J.* **8**, 2218–2230. (doi:10.1038/ismej.2014.63)
- Reichardt N, Duncan SH, Young P, Belenguer A, Leitch CM, Scott KP, Flint HJ, Louis P. 2014 Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J.* **8**, 1323–1335. (doi:10.1038/ismej.2014.14)

25. Duncan SH, Russell WR, Quartieri A, Rossi M, Parkhill J, Walker AW, Flint HJ. 2016 Wheat bran promotes enrichment within the human colonic microbiota of butyrate-producing bacteria that release ferulic acid. *Environ. Microbiol.* **18**, 2214–2225. (doi:10.1111/1462-2920.13158)
26. Louis P, Flint HJ. 2017 Formation of propionate and butyrate by the human colonic microbiota. *Environ. Microbiol.* **19**, 29–41. (doi:10.1111/1462-2920.13589)
27. Chung WS, Walker AW, Bosscher D, Garcia-Campayo V, Wagner J, Parkhill J, Duncan SH, Flint HJ. 2020 Relative abundance of the *Prevotella* genus within the human gut microbiota of elderly volunteers determines the inter-individual responses to dietary supplementation with wheat bran arabinoxylan-oligosaccharides. *BMC Microbiol.* **20**, 283. (doi:10.1186/s12866-020-01968-4)
28. Wu G et al. 2011 Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108. (doi:10.1126/science.1208344)
29. Wang SP et al. 2020 Pivotal roles for pH, lactate, and lactate-utilizing bacteria in the stability of a human colonic microbial ecosystem. *mSystems* **5**, e00645-20. (doi:10.1128/mSystems.00645-20)
30. Hamer HM, Jonkers D, Venema K, Vanhoutvin S, Troost F, Brummer R-J. 2008 The role of butyrate on colonic function. *Aliment. Pharmacol. Ther.* **27**, 104–119. (doi:10.1111/j.1365-2036.2007.03562.x)
31. Yao CK, Muir JG, Gibson PR. 2016 Review article: insights into colonic protein fermentation, its modulation and potential health implications. *Aliment. Pharmacol. Ther.* **43**, 181–196. (doi:10.1111/apt.13456)
32. Ravcheev DA, Thiele I. 2017 Comparative genomic analysis of the human gut microbiome reveals a broad distribution of metabolic pathways for the degradation of host-synthetized mucin glycans and utilization of mucin-derived monosaccharides. *Front. Genet.* **8**, 111. (doi:10.3389/fgene.2017.00111)
33. Crost E, Tailford L, Gall GL, Fons M, Henrissat B, Juge N. 2013 Utilisation of mucin glycans by the human gut symbiont *ruminococcus gnavus* is strain-dependent. *PLoS ONE* **8**, e76341. (doi:10.1371/journal.pone.0076341)
34. Florin T, Neale G, Gibson G, Christl S, Cummings J. 1991 Metabolism of dietary sulphate: absorption and excretion in humans. *Gut* **32**, 766–773. (doi:10.1136/gut.32.7.766)
35. Sung J, Kim S, Cabatbat J, Jang S, Jin Y, Jung GY, Chia N, Kim P. 2017 Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis. *Nat. Commun.* **8**, 15393. (doi:10.1038/ncomms15393)
36. Ruppin H, Bar-Maeir S, Soergel K, Wood C, Schmitt M. 1980 Absorption of short chain fatty acids by the colon. *Gastroenterology* **78**, 1500–1507. (doi:10.1016/S0016-5085(19)30508-6)
37. Kettle H, Louis P, Flint HJ. 2022 Process-based modelling of microbial community dynamics in the human colon. Figshare. (doi:10.6084/m9.figshare.c.6216295)