

# Modelling the emergent dynamics and major metabolites of the human colonic microbiota

Helen Kettle,<sup>1\*</sup> Petra Louis,<sup>2</sup> Grietje Holtrop,<sup>3</sup>  
Sylvia H. Duncan<sup>2</sup> and Harry J. Flint<sup>2</sup>

<sup>1</sup>*Biomathematics and Statistics Scotland, Kings Buildings, Edinburgh, EH9 3JZ, UK.*

<sup>2</sup>*Microbial Ecology Group, Rowett Institute of Nutrition and Health, University of Aberdeen, Greenburn Road, Bucksburn, Aberdeen AB21 9SB, UK.*

<sup>3</sup>*Biomathematics and Statistics Scotland, Greenburn Road, Aberdeen AB21 9SB, UK.*

## Summary

**We present here a first attempt at modelling microbial dynamics in the human colon incorporating both uncertainty and adaptation. This is based on the development of a Monod-equation based, differential equation model, which produces computer simulations of the population dynamics and major metabolites of microbial communities from the human colon. To reduce the complexity of the system, we divide the bacterial community into 10 bacterial functional groups (BFGs) each distinguished by its substrate preferences, metabolic pathways and its preferred pH range. The model simulates the growth of a large number of bacterial strains and incorporates variation in microbiota composition between people, while also allowing succession and enabling adaptation to environmental changes. The model is shown to reproduce many of the observed changes in major phylogenetic groups and key metabolites such as butyrate, acetate and propionate in response to a one unit pH shift in experimental continuous flow fermentors inoculated with human faecal microbiota. Nevertheless, it should be regarded as a learning tool to be updated as our knowledge of bacterial groups and their interactions expands. Given the difficulty of accessing the colon, modelling can play an extremely important role in interpreting experimental data and predicting the consequences of dietary modulation.**

## Introduction

The human colon is one of the most densely populated microbial ecosystems on Earth (Whitman *et al.*, 1998) with several hundred different bacterial strains co-existing at any one time (Eckburg *et al.*, 2005; Qin *et al.*, 2010). The exact composition of the bacterial community in any given person varies with host factors including secretory products, peristalsis and host genotype; and environmental factors including diet, microbial inocula and drugs (Egert *et al.*, 2006). It has become apparent recently that diet plays a particularly important role in determining microbiota composition (de Filippo *et al.*, 2010; Walker *et al.*, 2011; Wu *et al.*, 2011; Flint *et al.*, 2012a; Cotillard *et al.*, 2013; David *et al.*, 2014). The impact of diet is assumed to be mediated largely via the substrate preferences of different bacterial groups, but also depends on microbial interactions such as cross-feeding and changes in environmental variables such as pH and transit time (Stephen *et al.*, 1987; Walker *et al.*, 2005; Flint *et al.*, 2007; Louis *et al.*, 2007; Duncan *et al.*, 2009). Incomplete data and the high variability in microbial composition within the human population make fully deterministic modelling of the human gut microbial community difficult. Nevertheless, modelling can potentially play an extremely important role in the interpretation of experimental data and ultimately in predicting responses to dietary modulation.

While the human colonic community is highly diverse from a phylogenetic viewpoint, there is also a high degree of functional redundancy (Flint *et al.*, 2007; Turnbaugh *et al.*, 2009; Louis *et al.*, 2010). This makes it feasible and indeed desirable to simplify the system in order to analyse its functional interactions and dynamics in a quantitative manner. This approach has also been taken by Munoz-Tamayo and colleagues (2010) who developed a mathematical model of colonic bacteria using four functional groups and much physical detail. Our approach here uses a simpler physical environment (corresponding to a laboratory fermentor experiment), but investigates the behaviour of human colonic microbial communities that are represented by 10 bacterial functional groups (BFGs) and comprise a large number of microbial strains with stochastically varying traits. The 10 BFGs include producers of the three major short

Received 21 November 2013; accepted 14 August 2014. \*For correspondence. E-mail helen.kettle@bioss.ac.uk; Tel. +44 (0) 131 650 8740; Fax +44 (0) 131 650 4901.

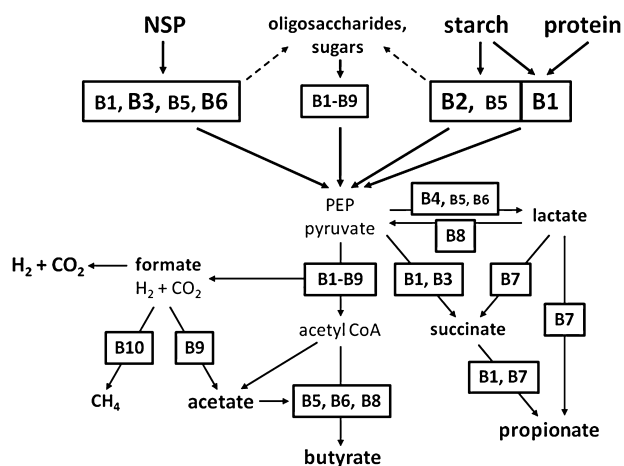
**Table 1.** Functional groups are defined by their contributions to dietary substrate (DS) breakdown and SCFA metabolism (see Appendix S2 for more detail on metabolic activity).

BFG	Name	Examples	Metabolic activity
B1	Acetate-propionate-succinate group	<i>Bacteroides</i> spp.	Grow on all DS including protein; when growing on either carbohydrate or protein they produce acetate, propionate, succinate, H <sub>2</sub> and CO <sub>2</sub> (Macfarlane and Gibson, 1997); when growing on protein, they also produce other metabolites (50%) not included in this model. Assumptions about substrate utilization are based on Salyers and colleagues (1977).
B2	Non-butyrate-forming starch degraders	<i>Ruminococcaceae</i> related to <i>Ruminococcus bromii</i> . Might also include certain <i>Lachnospiraceae</i> .	Grow on starch and (poorly) on NSP to produce acetate, CO <sub>2</sub> and H <sub>2</sub> (Moore <i>et al.</i> , 1972; Ze <i>et al.</i> , 2012).
B3	Non-butyrate-forming fibre degraders	<i>Ruminococcaceae</i> related to <i>Ruminococcus albus</i> , <i>Ruminococcus flavefaciens</i> . Might also include certain <i>Lachnospiraceae</i> .	Grow on NSP, but poorly on starch, to produce acetate and succinate (Moore <i>et al.</i> , 1972; Macfarlane and Gibson, 1997; Robert and Bernalier-Donadille, 2003).
B4	Lactate producers	<i>Actinobacteria</i> , especially <i>Bifidobacterium</i> spp, <i>Collinsella aerofaciens</i>	Grow on all DS except protein to produce acetate, lactate, formate and ethanol (Macfarlane and Gibson, 1997).
B5	Butyrate producers 1	<i>Lachnospiraceae</i> related to <i>Eubacterium rectale</i> , <i>Roseburia</i> spp.	Grow on all DS except protein. Acetate is regarded as a semi-essential substrate and growth is optimal when acetate is present, but there is no growth on acetate alone. They produce butyrate, H <sub>2</sub> and CO <sub>2</sub> (Barcenilla <i>et al.</i> , 2000; Duncan <i>et al.</i> , 2006; Duncan and Flint, 2008; Louis <i>et al.</i> , 2010).
B6	Butyrate producers 2	Certain <i>Ruminococcaceae</i> , in particular <i>Faecalibacterium prausnitzii</i>	Grow on all DS except protein, but with limited ability to utilize starch (Lopez-Siles <i>et al.</i> , 2012). Acetate is regarded as a semi-essential substrate and growth is optimal when acetate is present, but there is no growth on acetate itself. They produce butyrate, lactate, formate and CO <sub>2</sub> (Duncan <i>et al.</i> , 2002).
B7	Propionate producers	<i>Veillonellaceae</i> e.g. <i>Veillonella</i> spp., <i>Megasphaera elsdenii</i>	Representatives are assumed to be able to grow on NSP, starch and sugars to produce acetate, propionate and CO <sub>2</sub> . Alternatively, they can grow on lactate to produce acetate, propionate and CO <sub>2</sub> (Prins, 1977; Gottschalk, 1979; Seeliger <i>et al.</i> , 2002)
B8	Butyrate producers 3, able to utilize lactate.	<i>Lachnospiraceae</i> related to <i>Eubacterium hallii</i> , <i>Anaerostipes</i> spp.	Representatives are assumed to be able to grow on all the DS except protein to produce acetate, butyrate, formate, H <sub>2</sub> and CO <sub>2</sub> . If acetate is present, they can also grow on lactate to produce butyrate, H <sub>2</sub> and CO <sub>2</sub> (Duncan <i>et al.</i> , 2004a).
B9	Acetogens	Certain <i>Lachnospiraceae</i> , e.g. <i>Blautia hydrogenotrophica</i>	Representatives are assumed to be able to grow on all DS except protein to produce acetate, or they can grow on H <sub>2</sub> and CO <sub>2</sub> (and formate) to produce acetate (Ragsdale and Pierce, 2008), or they can convert formate while growing on carbohydrates to produce acetate, H <sub>2</sub> and CO <sub>2</sub> (Wolin <i>et al.</i> , 2003).
B10	Methanogens	Methanogenic archaea, e.g. <i>Methanobrevibacter smithii</i>	Can grow on H <sub>2</sub> and CO <sub>2</sub> to produce CH <sub>4</sub> . Alternatively they can grow on formate to produce CO <sub>2</sub> , and CH <sub>4</sub> (Liu and Whitman, 2008).

The putative relationships of BFGs to phylogenetic groupings are based on information from cultured representatives allied to 16S rRNA-based enumeration (e.g. Aminov *et al.*, 2006, Flint *et al.*, 2007, Walker *et al.*, 2011, Flint *et al.*, 2012b). These associations must be regarded as first approximations. A recent survey suggests that although the most dominant bacterial species have been isolated from the human colon, many less abundant bacterial phylotypes have not been studied in culture (Walker *et al.*, 2011).

chain fatty acids detected in faecal samples [acetate (Ace), butyrate (But) and propionate (Pro)] together with utilizers of Ace, lactate (Lac), succinate (Suc), formate (For) and hydrogen (H<sub>2</sub>) (see Table 1). It is worth emphasizing that a small number of highly abundant bacterial species dominate the system, including the But

producers, *Faecalibacterium prausnitzii* and *Eubacterium rectale* (Flint *et al.*, 2012b) and that many of these have now been cultured (Tap *et al.*, 2009; Walker *et al.*, 2011); in a recent study, more than 60% of phylotypes that individually accounted for > 0.5% of sequences were found to have been cultured (Walker *et al.*, 2011).



**Fig. 1.** Summary of functional groups among human colonic bacteria. For detailed assumptions see Table 1, Supplementary Tables S1 and S9, and Appendix S2. CoA, coenzyme A; PEP, phosphoenolpyruvate.

Our assumptions are largely based on studies with these dominant cultured bacteria.

This paper describes the development of a theoretical model, which simulates the human colonic microbial community and its major metabolic products. The model is designed to simulate a wide range of different microbial communities in an attempt to reproduce the kind of diversity that can be found in the human population. However, before attempting to produce realistic population scale results (future work), we compare the model to results generated in more tightly controlled environments. Thus in this paper, we describe comparisons with continuous flow fermentor experiments on the effects of pH and peptide supply on human faecal microbiota from healthy volunteers (Walker *et al.*, 2005). The pH of the colon is potentially a key factor in determining the species composition and metabolic outputs of the human intestinal microbiota because of the relative sensitivity of most Gram-negative anaerobes to slightly acidic pH (Duncan *et al.*, 2009).

The model (described in detail in Appendix S1, with Tables S1–S9) has been developed as a framework within which to test our current understanding of gut microbial interactions and to study the effects of diversity on system function. The metabolic pathways are easy to change (they are input files), the parameter ranges can be as wide or narrow as the user wishes, and the model can be run with any subset of the 10 BFGs or even for just one strain in one BFG. Thus, the model can also be used by researchers wishing to understand and parameterise relatively simple interactions. In this paper, however, we aim to demonstrate the potential that the model has at a wider level; for example, to simulate variation between microbial communities at the (human) population scale and to

investigate the role of diversity within each microbial community. We hope that our model, which we refer to as ‘AMES’ for Anaerobic Microbial Ecosystem Simulator, will prove useful for all of these different requirements and the source code (FORTRAN 90) which runs on Linux and Windows is available in Appendix S4. However, because this model is part of an ongoing research project, we strongly recommend that researchers wishing to use the code for their own research should contact the lead author for the latest version.

## Results

### Model description and assumptions

Our aim here is to develop a model that simplifies the enormous complexity of the human intestinal microbiota by focussing on major functional groups and their metabolic products. The 10 BFGs shown in Table 1 were chosen to represent functional groups that can be defined based on our existing knowledge of the substrate preferences and fermentation products (specifically short chain fatty acids) of numerically predominant species. In order to relate bacterial growth and substrate utilization to metabolic outputs, it was essential to specify stoichiometries and these were derived from the major fermentation pathways described for representatives of each BFG (shown in detail in Appendix S2). Some BFG (e.g. B5 and B6 But producers, B10 methanogenic archaea) correspond closely to well-defined phylogenetic groups, whereas others include representatives of several phylogenetic groups.

Growth substrates available in the large intestine are divided into four categories: protein, non-starch polysaccharides (NSP), resistant starch and sugars (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: Ace, Pro, But, Lac, Suc, For, H<sub>2</sub>, carbon dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>) and ethanol. Six of these metabolites (Ace, Lac, Suc, For, H<sub>2</sub> and CO<sub>2</sub>) are also considered as substrates, because they are known to be consumed by some BFGs. The relationships among substrates, BFGs and metabolites are summarized in Table 1 and Fig. 1. The assumed stoichiometries for metabolite production and uptake, which are derived from the schemes shown in Appendix S2, are given in Table S1. For each substrate and BFG combination, we have assumed a maximum growth rate  $\mu^M$  and yield  $Y$ , as detailed in Tables S2 and S3. These values are in part empirically derived, but most must be regarded as crude approximations. A

further aim of the model was to simulate the responses of different BFGs within the microbial community to physiologically relevant pH changes. The impact of pH on growth was based on available empirical evidence for representatives of each BFG (Duncan *et al.*, 2009). In the model, each BFG 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 [Table S4 and Appendix S1 (A2.4)]. The pH limitation is incorporated into the model by scaling the bacterial growth rate by a pH limitation factor  $H$  (where  $0 \leq H \leq 1$ ), which is given a trapezoidal shape [see Appendix S1 (A2.4)].

In the interests of brevity, the governing equations for the model are included in Appendix S1 and only a brief overview is given here. Essentially, we model the rate of bacterial growth using Monod kinetics and assume that from 1 g of resource,  $Y$  g of biomass is produced (Table S3). We then assume that any of the resource that is taken up, but is 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.

The model takes an 'emergent behaviour' approach, meaning that each simulation starts not with just 10 strains (one for each BFG), but with 100 (10 per BFG) strains with traits stochastically assigned within the limits of the respective BFG, thus allowing the system to respond through strain competition. Because we assign strain traits stochastically each time we run a simulation, the microbial community will be slightly different. We will henceforth refer to each model run as a 'realization'. The importance of this inbuilt strain diversity in producing meaningful simulations will be explained and illustrated in more detail once we have considered the application of the model to some real experimental data.

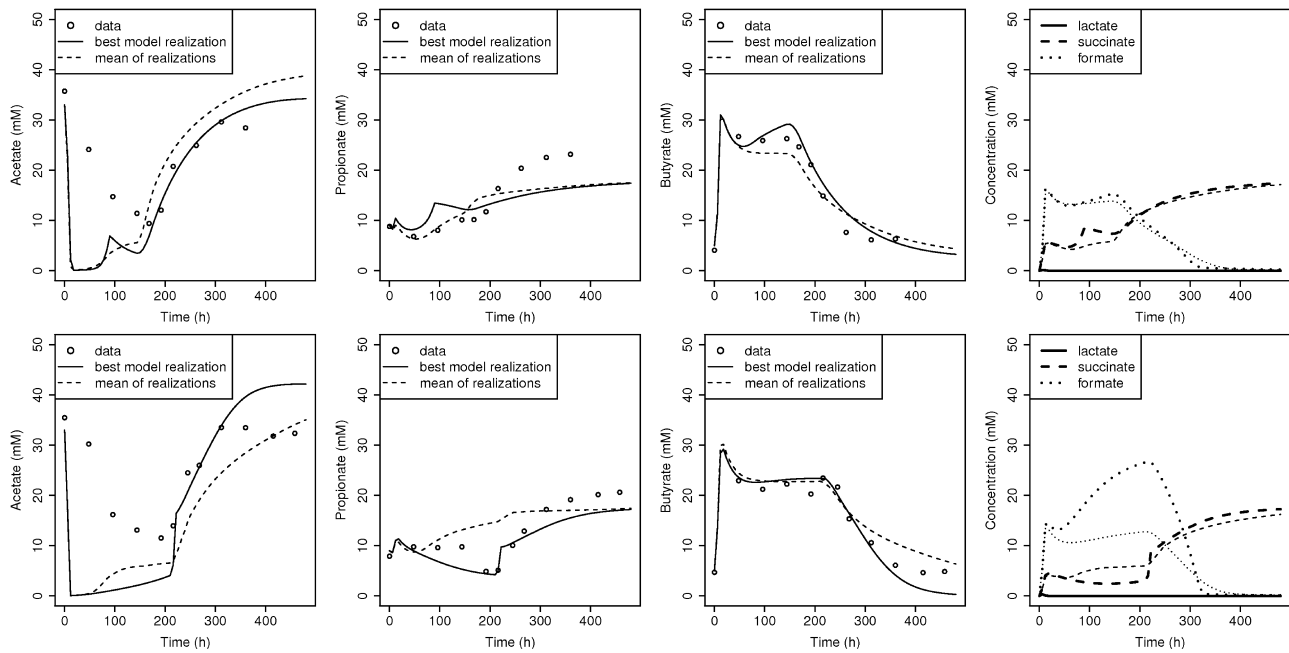
#### *Application of the model to the human gut microbiota in vitro continuous culture*

In order to explore the potential of the model, we applied it to earlier experiments that determined the effect of different controlled pHs and peptide inputs upon microbial community composition and net short chain fatty acid (SCFA) formation for human colonic microbiota in anaerobic continuous culture. These experiments used two parallel single-stage anaerobic fermentors of 250 ml of volume that were simultaneously inoculated with bacteria from a faecal sample ('donor 1') and then supplied continuously with fresh medium at the rate of one turnover per day (Walker *et al.*, 2005). The supply medium contained a mixture of dietary polysaccharides (total 0.74% w/v) as

energy sources, together with either 0.1% peptide or 0.6% peptide. The experiment was repeated using a faecal inoculum from a second donor ('donor 2'). The pH in each vessel was controlled at 5.5 for 6 days (144 h; donor 1) or 9 days (216 h; donor 2) and then increased over 2 days to reach 6.5 at which it was then fixed for 9 days (Walker *et al.*, 2005). The allocation of the constituents of the fermentor medium to our four model substrate groups is shown in Table S5. Because fermentor medium was maintained under a steady stream of  $\text{CO}_2$ , we added this to the inflow in the model at  $1 \text{ g d}^{-1}$  to ensure that the system was not  $\text{CO}_2$  limited. In the experiments, fluorescence in situ hybridization (FISH) microscopy was used to monitor the absolute populations of different phylogenetic groups of bacteria (Walker *et al.*, 2005). The phylogenetic groups detected by the FISH probes correspond well with our assumed BFGs for bacterial groups B1 (Bac303), B4 (Bif164), B5 (Rrec584), B6 (Fprau645) and B7 (Prop853) (Table S6). On the other hand, B2 and B3 are expected to include a subset of species recognized by the Erec482 (*Lachnospiraceae*) and Rfla729/Rbro730 (*Ruminococcus* spp.) probes and B8 and B9 comprised a subset of species recognized by Erec482; estimation of their numbers must therefore be regarded as a crude approximation (see Table S6).

We performed 100 realizations of the model for each donor, with 10 strains per BFG, and then found the realization that was the closest [in terms of root mean squared error (RMSE)] to the measured Ace, Pro and But concentrations and BFG counts from the FISH probes (Table S6). The results for the high (0.6%) peptide input are shown for SCFA in Fig. 2 and for BFGs in Fig. 3. First, it is clear that the SCFAs from the two donors have a very similar response to the change in pH and our model captures the experimental data well, especially for Pro and But. The model tends to overpredict the rate of uptake of Ace at pH 5.5, however, probably reflecting the fact that Ace is both produced and consumed by different BFGs (Flint *et al.*, 2012b), which makes it more difficult to accurately simulate its concentration changes.

The model was very successful in predicting the switchover from a more butyrogenic to a more propionigenic fermentation between pH 5.5 and pH 6.5. This was largely explained by the greater ability of the B5 (and in the model, B8) groups of But producers to compete for carbohydrate substrates at pH 5.5, whereas these groups were outcompeted by B1 (*Bacteroides*) at pH 6.5 (Fig. 3). In addition, some new experimental evidence shown in Fig. 4 reveals that pH can alter the fermentation stoichiometries for representatives of the B5 group, changing in favour of increased But production and Ace utilization as the pH becomes more acidic. For this reason, the simulations presented in this paper assumed the stoichiometry for group B5 that is observed at pH 5.5



**Fig. 2.** Comparing SCFA concentrations from the 'best-fit' model realization and the mean of 100 model realizations with the Walker and colleagues (2005) fermentor experiment data for peptide at 0.6% for Donor 1 (top row) and Donor 2 (bottom row). In the column on the far right, the bold lines are from the 'best-fit' model and the thin lines are the mean of the 100 model realizations. Observed concentrations of lactate, succinate and formate (not shown) did not exceed 2 mM in the fermentor experiment.

(Table S1) although an alternative stoichiometry closer to that observed at pH 6.5 in *E. rectale* ('Alt. B5' in Table S1; Duncan and Flint, 2008) did not greatly change the findings (data not shown). The third group of But producers (B6) did not achieve large populations either in these simulations or in the experiments of Walker and colleagues (2005). This is explained by the limited ability of *F. prausnitzii* strains (B6) to utilize the dietary polysaccharide substrates (mainly starch) supplied here (Lopez-Siles *et al.*, 2012). Results from all model realizations are shown in the box and whisker plots in Figs 5 (SCFA) and 6 (BFGs).

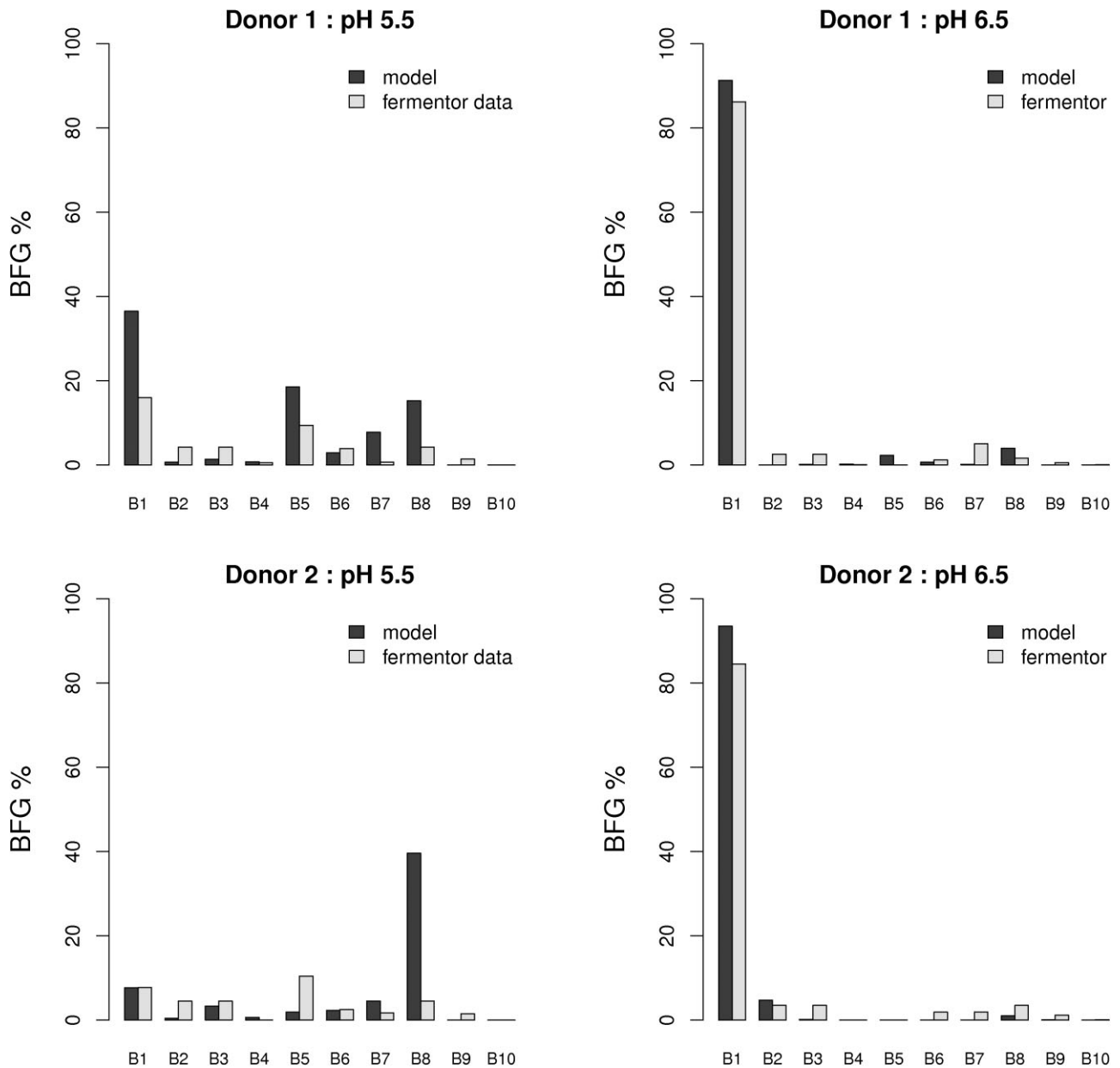
The model also broadly captured the responses to pH observed at the 0.1% peptide input, at least for donor 1 where the proportion of B1 bacteria again increased markedly with the pH change from 5.5 to 6.5, although this was not seen for donor 2 at 0.1% peptide (Fig. 6). Generally, the simulations covered the data (i.e. the data points fall within the whiskers of the box plots) (Figs 5 and 6). Total SCFA concentrations were similar for the two peptide input levels both in the model prediction and experimental data (Fig. 5), whereas the observed bacterial count (Eub probe) at 0.1% peptide was only about 20% of that observed at 0.6% peptide (Walker *et al.*, 2005) (Fig. 6). Thus, SCFA production was maintained despite the decreased growth rates, reflecting limitation by the nitrogen supply rather than the carbohydrate energy source. As a result, the predicted biomass at 0.1%

peptide was higher than the observed biomass and adjustments of the model would clearly be needed to predict absolute bacterial numbers whenever the carbohydrate energy supply is not the main growth-limiting factor.

The measured concentrations of Lac, Suc and For were generally close to zero in the 0.6% peptide condition although showing some transient increases in the 0.1% peptide condition. The model overestimated both For and Suc, which suggests that either their consumption via cross-feeding is being underestimated or their production is being overestimated (for reasons that will be considered in the Discussion).

#### *Investigation of competition and cross-feeding by 'knockout' of single BFGs*

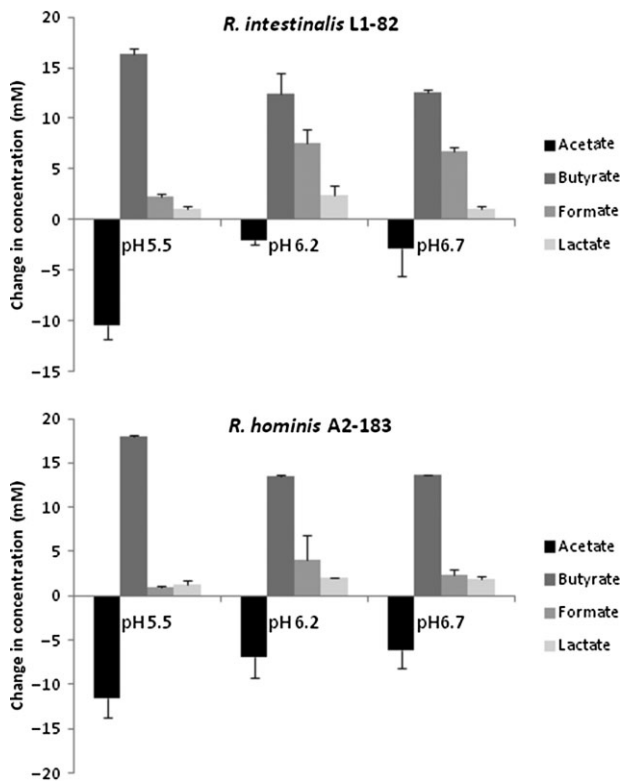
An influence of input microbiota composition is apparent from the different predicted outcomes for BFG proportions and SCFA concentrations for donors 1 and 2 (Figs 3, 4). It is of considerable interest therefore to ask how complete elimination of any one of the 10 functional groups would affect the remaining BFGs and the metabolite profile of the residual community. For simplicity, we assumed an even spread of the initial biomass between groups for these simulations, but retained a pH switch after 8 days, as in the Walker and colleagues (2005) experiments. Figs 7 and 8 show the



**Fig. 3.** Comparing bacterial communities from the 'best-fit' model realization with our estimates of the BFGs from Walker and colleagues (2005) probe data where the fermentor pH is changed from 5.5 to 6.5 at 144 h (Donor 1) and 216 h (Donor 2) for peptide at 0.6%. The plots show the community composition at the end of the pH 5.5 regime and at the end of the experiment (360 h for Donor 1, and 458 h for Donor 2). The 'model' values are a percentage of the total biomass at the end of the simulation; the 'fermentor data' values are a percentage of the total biomass in our allocated BFGs at the end of the fermentor experiment.

consequences on the mean SCFA concentrations and the mean microbial community composition (over 50 realizations) if any one group was removed at a time. When all groups were present, B1 (*Bacteroides*) dominated the 6.5 pH regime in the model community; at pH 5.5 the community is more evenly spread (Fig. 8). When B1 is removed, B2 dominates the pH 6.5 regime. With regard to SCFA, removing B7 had multiple effects in decreasing Pro and also allowing But to increase in the

pH 5.5 regime. Interestingly, removing B1 had a large effect in maintaining But concentrations at pH 6.5. The reason for this must be that by removing B1 (which dominates the pH 6.5 regime and does not produce But) the But producers (B5, B6 and B8) are given greater access to the available resources. Removing B6 significantly decreases the uptake of Ace in the pH 5.5 regime. The resulting Ace profile is similar to the fermentor results (Fig. 2) and implies that the overestimation of Ace uptake



**Fig. 4.** Effect of pH upon observed formation stoichiometries for two butyrate producers (representatives of group B5). Data represent the means of triplicate cultures. A shift towards greater butyrate production and acetate consumption is evident for both strains at pH 5.5.

by the model could be due to an overestimate of B6. Furthermore, removing B10 allows a large build-up of For at pH 5.5, which implies that the high For concentrations shown in Fig. 2 would be reduced (and therefore more similar to the fermentor measurements) if the proportion of B10 was increased in the model.

It is also interesting to consider how many BFGs are actually required to accurately represent the system. For the input medium used here (Table S5), knocking out B4 and B9 makes very little difference to the metabolite concentrations (Fig. 7), so for this particular situation, these BFGs could be omitted and the system could be adequately modelled with eight groups. However, for a different input medium, this may not be the case.

#### *Emergent behaviour and adaptation to environmental change (resilience)*

A very important feature of our model, which should now be explained in more detail, is that we have adopted an 'emergent behaviour' approach. In practice, this means that each simulation starts not just with 10 strains (one

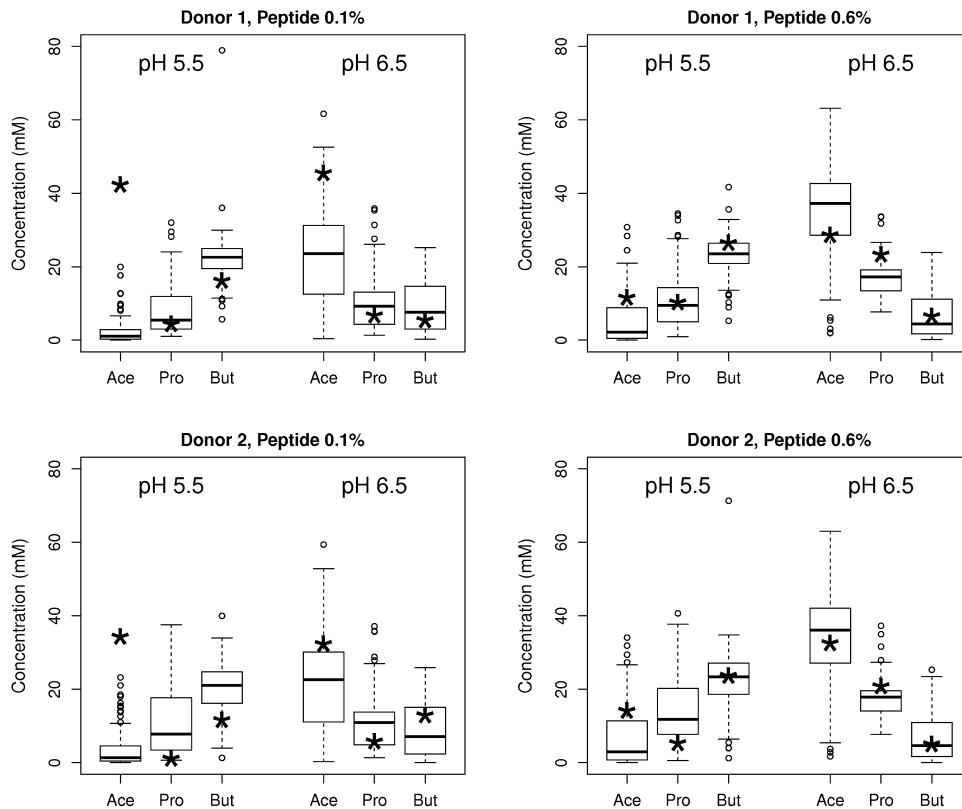
for each BFG), but with 100 (10 per BFG) or more (this is simply a model parameter,  $N_s$ ). The multiple strains assigned into each of these BFGs have stochastically generated parameter values [from ranges appropriate to the given BFG; see Appendix S1 (A5)] for the following traits: maximum growth rate  $\mu^M$ , half-saturation constant  $K$ , yield  $Y$  and pH preference  $H$ . Although the strain traits are stochastically generated at the start of the simulation, they are fixed at these values throughout the duration of the simulation, and it is the competition between these strains that then allows the community to adapt as a whole. The generation of these traits is described in detail in Appendix S1 (A5), the parameter values are given in Table S8 and the relative growth rates for each BFG are given in Table S2. The philosophy behind this approach is that given a certain environment, a large, stochastically generated bacterial community will self-organize until it reaches a stable community composition. In a low resource situation, the strain with the highest value of  $\mu^M/K$  is expected to outcompete the other strains, whereas in a high resource situation,  $\mu^M$  is more important.

In order to demonstrate how the model allows the mean trait values of each BFG to adapt through strain succession, we will briefly consider how the mean pH trait of just one BFG, B1, changes as the environmental pH changes. We define the mean trait as a weighted average based on the concentration of each strain,  $X_i$ , such that for a trait value,  $z_i$ , the mean trait value,  $z^{av}$ , for the group is given by

$$z^{av} = \frac{\sum_{i=1}^{N_s} X_i z_i}{\sum_{i=1}^{N_s} X_i} \quad (1)$$

(Norberg *et al.*, 2001). Fig. 9A shows how the biomass of B1 varies in response to a sinusoidally daily varying environmental pH when  $N_s$  is changed from 1 (i.e. only one strain in the group) to 25. It is clear that the ability of the group to adapt has a big effect on maintaining bacterial growth – the group biomass is consistently higher when there are 25 strains present initially rather than one (mean concentration of 2.9 g l<sup>-1</sup> compared to 2.1 g l<sup>-1</sup> – an increase in biomass of almost 40%). The result that increased diversity leads to greater production is a well-known result, sometimes referred to as the 'insurance hypothesis' (Yachi and Loreau, 1999).

Fig. 9B shows how the mean pH trait of the group (Eq. 1) containing 25 strains changes from the centre of the trait space at the start towards the lower end of the trait space as those strains with the lowest pH range outcompete the others. Thus, the initial intra-group diversity is important in allowing selection to occur, but this diversity decreases with time as the best strain will often



**Fig. 5.** Comparing the three main SCFA [acetate (Ac), butyrate (Bu) and propionate (Pr)] for different peptide and pH for two donors at the end of each pH period. One hundred model realizations are represented by box and whisker plots where the mean is shown as a line and the box covers 50% of the simulations and the whiskers extend to 95%. The observed data are shown by stars.

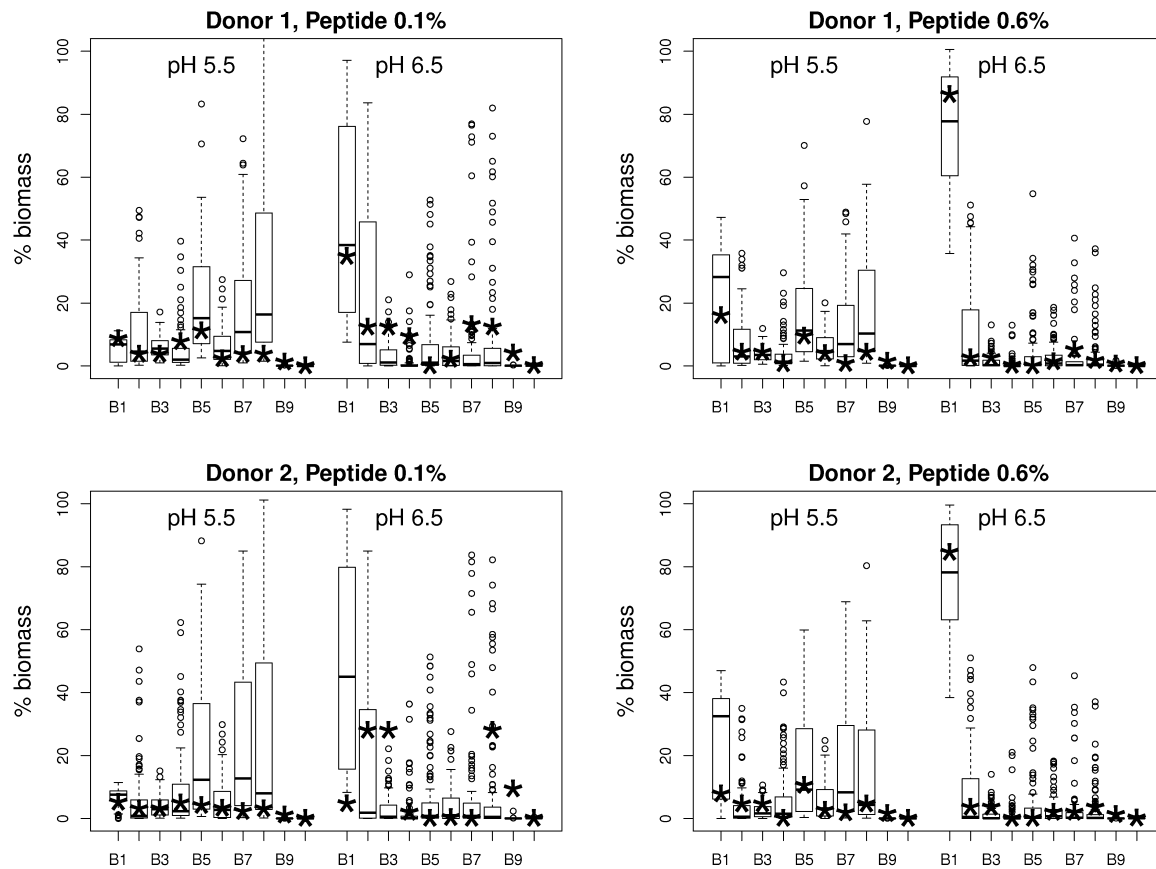
dominate the group to the exclusion of all other strains. This is beginning to happen in our simulation (Fig. 9C) where currently two strong strains have begun to dominate, and as time goes on, the top strain will outcompete all others (competitive exclusion). We recognize that intra-group diversity is high within the colonic microbiota *in vivo*, which means that there must be mechanisms that counteract this tendency towards competitive exclusion (e.g. see Kettle et al., 2014 for a possible explanation).

*An inverse relationship between microbiota diversity and variability within the human population*

The model can also be used to consider the question of variability of microbial communities within the human population by regarding each model realization as representing the microbiota of a different person. Individual microbial communities can of course be made to differ by using different random seeds to generate the stochastic strain traits or by omitting some of the BFGs; however, we can also consider the influence of microbiota diversity simply by varying the number of strains per group ( $N_s$ ). When this was done, we found that as  $N_s$  increases,

the differences between the microbial communities in metabolite outputs (exemplified by Ace) emerging from different realizations (or individual microbiotas) decreases, leading to a less varied population of (human) individuals. This is because the model allows adaptation through strain succession (as described in the previous section) and the higher the value of  $N_s$ , the more likely it is that the best strain for the environmental conditions will be very similar from one model realization ('person') to the next. On the other hand, if  $N_s$  is small, it is more likely that the randomly allocated strain traits will differ between people. Fig. 10 shows a simulation over 20 days with pH held at 6.5 for 50 'people' for 1, 10 and 30 strains per group where the initial biomass is spread evenly between all BFGs. Looking at Fig 10A–C and the summary of these in Fig. 10D, it is clear that a small number of strains per group results in more variation with respect to net Ace formation between the microbial communities in our modelled 'human' population than high strain diversity. This is important to note if one wishes to use the model to simulate response statistics for a human population – e.g. if wishing to simulate the average effects of changes in diet on gut microbiota.





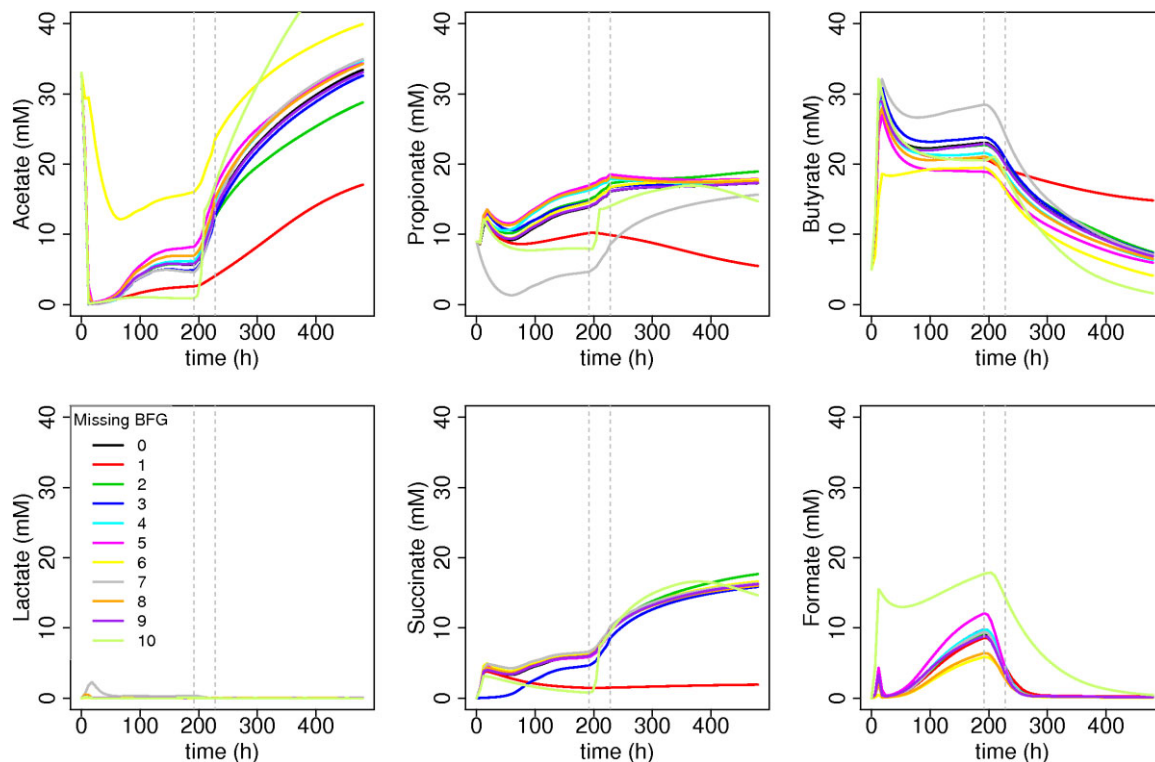
**Fig. 6.** Comparing the bacterial community composition (as percentage of the total biomass at the end of the simulation) for different peptide and pH for two donors. One hundred model realizations are represented by box and whisker plots where the mean is shown as a line and the box covers 50% of the simulations and the whiskers extend to 95%. The data are indicated by stars.

## Discussion

We have developed a theoretical model of the human colonic microbiota that is based on BFGs with the long-term aim of relating dietary intake, microbiota composition and metabolic outputs. The ‘emergent behaviour’ approach to modelling that we have adopted proved important in several ways. First, it allowed the introduction of variability and uncertainty into the model, meaning that knowledge of precise parameter values was not a crucial requirement and there was no need to overly constrain the model at the outset. Because species succession sorts the strains into a viable community structure (through resource competition) for the given environment, this means that after sufficient time has elapsed, the biomass-weighted average parameter set for any group is determined to a large extent by the environment rather than being fixed *a priori*. Second, the approach mirrors the *in vivo* situation in which the gut microbiota displays a high degree of strain diversity and dynamic competition. Third, it provides a mechanism for simulating the bacterial community of a large number of

human individuals; thus, 100 model runs could be interpreted as representing data from 100 different people by initializing the model with different bacterial communities. Multiple model runs can then be analysed to identify the characteristics of this simulated human population. Thus, by assigning the individual strain parameters stochastically for each model run, we can incorporate both the uncertainty in our knowledge of the system and the natural variation between people. An important conclusion from the modelling reported here is that a system comprising a larger number of strains with variable input values gave rise to more reproducible behaviour than one comprising a smaller number of strains. This implies that high strain diversity within the microbiota will tend to decrease variation in functional responses between people.

The model was used here to simulate the observed effects of controlled perturbations of the human colonic microbial community in anaerobic continuous flow fermentors. The results show that our model is able to reproduce many of the observed dynamics of an experimental continuous flow model of human colonic

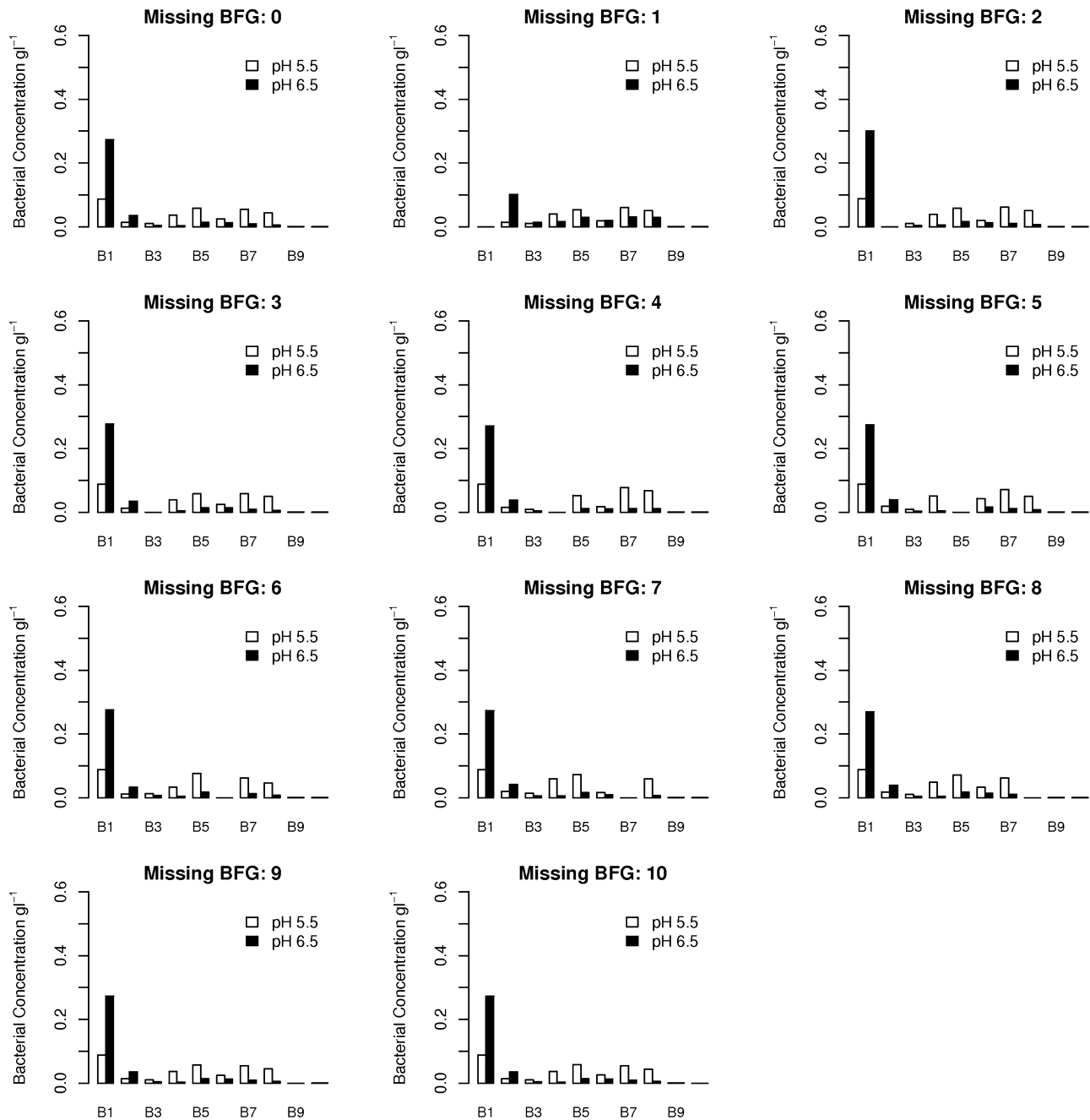


**Fig. 7.** The effect of removing one group at a time on the mean SCFA concentrations over 50 model realizations. The model was set up with biomass equally distributed between the 10 BFGs. The line marked with '0' shows the response when no groups are removed (i.e. the full community response). The model is run for a change in pH from 5.5 to 6.5 after 192 h (as in Walker *et al.*, 2005). All other conditions and settings are as given in Table S8.

microbiota in response to an imposed pH change. In particular, But concentrations were very close to the observed values, presumably reflecting our relatively detailed knowledge of bacterial groups involved in But production based on isolation and culture (Barcenilla *et al.*, 2000; Duncan *et al.*, 2004b; Louis *et al.*, 2010). For example, the ability of many of these bacteria to take up Ace and produce But is built into the model and was reflected in the initial changes in Ace and But. Another key feature appears to be the impact of slightly acidic pH in conferring a competitive advantage to certain groups of But producer (notably B5). Other aspects of SCFA metabolism were less accurately modelled, however, suggesting more limited understanding of the phylogenetic groups responsible and their physiology. Pro is formed by *Bacteroides* spp. and certain groups of *Firmicutes*, but the partitioning of flux between Pro and Suc in *Bacteroides* spp. is known to differ with the type of growth limitation and gas phase (Macfarlane and Gibson, 1997). Nevertheless, the dominant *Bacteroides* populations at pH 6.5 (both in the model and experimentally) apparently produced less Pro plus Suc per unit cell mass in the mixed system than is predicted by the assumed stoichiometries. Recent evidence

also suggests that very few *Firmicutes* produce Suc, contrary to the assumption made here for B3 (Reichardt *et al.*, 2014).

The model also overestimated For concentration. Most bacteria can produce either For or H<sub>2</sub> and CO<sub>2</sub> depending on the regulation of pyruvate-metabolizing enzymes and For H<sub>2</sub> lyase activity (Gottschalk, 1979). It has been shown in *Escherichia coli* that For H<sub>2</sub> lyase is involved in cytoplasmic pH regulation in response to high For concentrations and low pH (Bagramyan and Trchounian, 2003). We may therefore either be overestimating the production of For in the mixed community, or underestimating its conversion to H<sub>2</sub> and CO<sub>2</sub> or CH<sub>4</sub> by other community members. The interactions of the main groups of H<sub>2</sub>-consuming microorganisms are of considerable interest in relation to gut health (Nava *et al.*, 2012), but more theoretical and experimental work will be needed before we can modify this model to examine the competition between acetogens (B9), methanogens (B10) and the third group (not included here) of sulfate-reducing bacteria (SRBs). It will clearly be of interest to expand the model by including SRBs, which can also use Lac as a co-substrate (Marquet *et al.*, 2009). Neither gas production nor the populations of methanogens, SRBs and

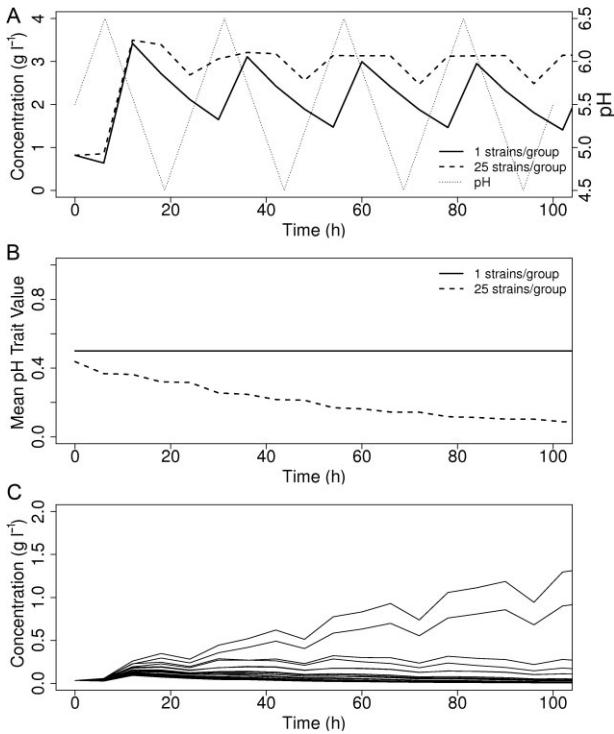


**Fig. 8.** The effect of removing one group at a time on the mean composition of the bacterial community (over 50 model realizations) at the end of the pH 5.5 regime (193 h; shown in white) and at the end of the pH 6.5 regime (480 h; shown in black). The model was set up with biomass equally distributed between the 10 BFGs. The plot marked 'Missing BFG: 0' shows the response when no groups are removed (i.e. the full community response). All other conditions and settings are as given in Table S8.

acetogens were measured by Walker and colleagues (2005).

In constructing this model, we have had to make a number of approximations that require some comment. The correspondence between functional groups (BFGs) within the model and phylogenetic groups as defined by 16S rRNA probes is often approximate, and some BFGs,

for example B8 (acetogens) (Drake *et al.*, 2008), do not fall into clearly defined phylogenetic groups. In the future, it may be better for this purpose to develop detection methods that are based on functionally relevant genes (Louis *et al.*, 2010; Flint *et al.*, 2012b) rather than relying on 16S rRNA-based methods. Assumptions about stoichiometry were based largely on *in vitro* culturing of a



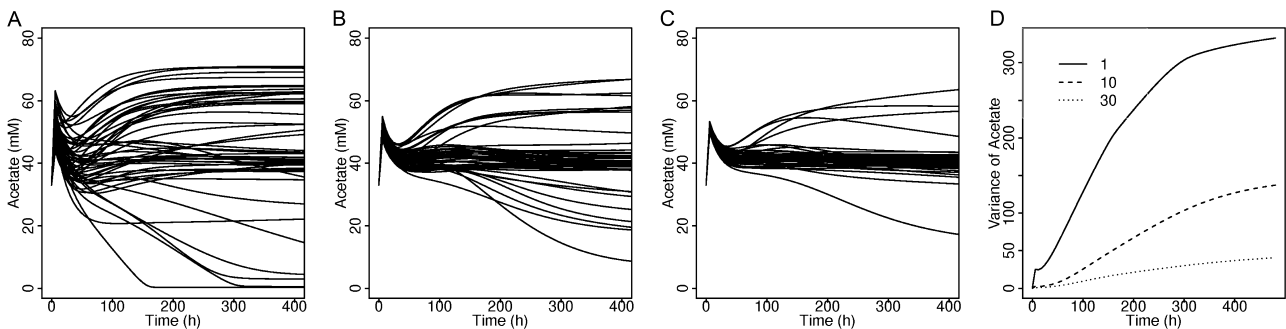
**Fig. 9.** The effect of the number of strains per group,  $N_s$ , on the group's ability to adapt when subjected to a daily varying environmental pH. The model is set up for just one group – B1 with pH limitation,  $H$ , equal to 1 (i.e. no limitation) for  $\text{pH} > 6.35$ , and decreasing linearly to 0 (i.e. no growth) as pH decreases to 6.1 ( $H = 0$  for  $\text{pH} < 6.1$ ). Two model realizations were performed:  $N_s = 1$  and  $N_s = 25$ .  
 A. Shows the variation in the total bacterial concentration for the group and the pH.  
 B. Shows the adaptation of the group's mean pH trait.  
 C. Shows the concentration for each strain for the  $N_s = 25$  realization.

small number of isolated bacteria. It is known that stoichiometries will vary in the mixed community as a result of microbial interactions and cross-feeding (Latham and Wolin, 1977; Flint *et al.*, 2007) and can be influenced

by substrates, growth rates and nutrient supply (Louis *et al.*, 2007). It is therefore remarkable how closely the model fits the observations for certain products (especially But) and BFGs, and this degree of fit should improve markedly as the discrepancies prompt research to provide the necessary information. Another approximation is that, for simplicity, a single mass per bacterial cell has been assumed throughout, independent of phylogenetic group or environmental conditions.

The model includes substrate affinities, but assumes simultaneous growth on all of the available substrates that the bacterial group is capable of utilizing. *Bacteroides thetaiotaomicron*, for example, expresses enzymes for the utilization of different diet-derived polysaccharides simultaneously (Rogers *et al.*, 2013) and this flexibility has also been shown for numerous other examples (Kovarova-Kovar and Egli, 1998). However, it should be noted that it is also possible for individual bacteria to possess regulatory mechanisms that result in substrate preferences (Russell and Baldwin, 1978). Because substrate preferences are also likely to vary within functional groups, the group assumptions in the present model should be taken to represent an assumed average behaviour. As presented here, the model can be regarded as a learning tool to enhance our understanding of microbial communities in the gut, and because there is still much to be discovered about this system, the model has been coded to allow a reasonable amount of flexibility. For example, the metabolic pathways and stoichiometries for each group are simply input files (i.e. they are not hard coded) so that they can be updated as our knowledge of each bacterial group increases.

The model clearly has greater potential than we have illustrated here. In particular, there is potential to apply this model with appropriate refinements to the human colon. One refinement would be to greatly diversify the range of substrates to include endogenous secretions (mucins) and to subdivide the complex category of NSPs, with a



**Fig. 10.** The relationship between  $N_s$  (the number of strains per group) and variation in model realizations (here demonstrated with acetate concentrations for 50 realizations). The model was set up with biomass equally distributed between the 10 BFGs and kept at pH 6.5 for 400 h.  $N_s$  was set at (A) 1, (B) 10 and (C) 30 strains per group; all other conditions are as specified in Table S8. D shows how the variance of acetate (over 50 realizations) changes with  $N_s$ .

corresponding expansion of the number of BFGs. This could have a major impact on the maintenance of strain diversity, but would also be computationally demanding. It also appears necessary to represent both the proximal and distal colon – i.e. a two-box model rather than a one-box model. We also anticipate the future inclusion of feedback on pH from SCFA production, which will introduce a degree of self-regulation to the system (Kettle *et al.*, 2014). More than 90% of SCFA produced by the colonic microbiota are absorbed across the colonic mucosa, and it will therefore be important to consider the impact of changes in absorption that can result for example from changes in flow rate or intestinal physiology. Furthermore, the experiments considered here represent a pseudo steady-state system with a constant rate of substrate supply, while *in vivo* conditions resulting from meal feeding will show more variation with time. These factors can however be examined quite readily through modifications to the model. Finally, because the fermentor system was supplied with soluble substrates and represents a relatively homogeneous community, the model has not attempted to consider local spatial heterogeneity resulting from attachment to insoluble substrates, although this occurs *in vivo* (Swidsinski *et al.*, 2008; Walker *et al.*, 2008).

To conclude, the work presented here describes a first attempt at modelling microbial dynamics in the human colon incorporating both uncertainty and adaptation. Given the difficulty of accessing the colon itself, and the need to rely on measurements made on faecal samples, modelling can play an extremely important role in the interpretation of experimental data and the prediction of consequences from dietary modulation. It also provides a rigorous examination of our developing knowledge and understanding of the microbial community of the human large intestine.

## Experimental procedures

### Modelling procedures

The equations used in the model and details on the software are included in Appendix S1 and the model code is also available in Appendix S4. Furthermore, a sensitivity analysis of the model is given in Appendix S3. Here, however, we include some details on how we used the model to simulate the particular fermentor experiments described by Walker and colleagues (2005).

### Relating experimental data on microbiota composition to BFGs

The bacterial counts from the probes used in the laboratory experiments are apportioned to our 10 BFGs as shown in Table S6. Because microbial data are given in cell counts

per unit volume and our model variables are in mass per unit volume, we need to convert cell counts to mass for each BFG. This is problematic due to a lack of appropriate data and we therefore simply assume that all of the groups have the same mass per cell. The cell mass is estimated by assuming that the bacteria have the same density as water and have a tubular shape with radius  $0.3\ \mu\text{m}$  and length  $6\ \mu\text{m}$  (based on Duncan *et al.*, 2002), giving a cell mass of  $1.7\ 10^{-12}\ \text{g}$ . Assuming that the dry weight is 50% of wet weight (Bratbak and Dundas, 1984) gives the dry weight of one cell as  $0.85\ 10^{-12}\ \text{g}$ . This value was then used to convert the FISH probe counts ( $\text{count l}^{-1}$ ) into biomass ( $\text{g l}^{-1}$ ). However, it should be noted here that these values will vary between strains and species (e.g. Feijó Delgado *et al.*, 2013 found the water content of *E. coli* to be 72% of the mass) and therefore this conversion factor should be treated with caution.

Each BFG contains 10 strains with the initial mass for that group evenly distributed between them. Although the initial conditions in terms of the biomass for each group are defined by the probe data for each donor (and therefore will change when simulating a different donor), this does not fully constrain the model because the model parameters (strain traits) required for every strain are not known (and there will also be some strains that have not been identified by the probes). In an attempt to find the community best suited to the particular donor, we therefore run the simulation many times, with each realization beginning with a slightly different bacterial community in terms of the strain traits [details in Appendix S1 (A5)].

### Modelling SCFA outputs

Concentrations of substrates, metabolites and each microbial strain are output from the model at 6 h intervals. For each experiment, we find the model realization that is the closest fit to the data and assume that this was started with a community most similar to the faecal donor who participated in the experiment. To compute the fit errors, we find the RMSE for the data from the three main SCFA (Ace, Pro and But) and the probe data for the 10 bacterial groups, and we then compute the total error as a weighted sum of these. Because the SCFA data are directly comparable with our model and subject to less uncertainty than the probe data, when minimizing the errors, after normalizing by the means, we weight the importance of the SCFA errors at 75% and the probe errors at 25%.

### Bacterial growth and SCFA changes of Roseburia strains at different initial pH values

Overnight cultures of two isolates deposited with Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), *Roseburia intestinalis* (L1-82; DSM 14610<sup>T</sup>) and *Roseburia hominis* (A2-183; DSM 16839<sup>T</sup>), were used to inoculate yeast casitone fatty acid glucose (YCFAG) medium (Lopez-Siles *et al.*, 2012; prepared at initial pH values of 5.5, 6.2 and 6.7) in triplicate. Following 24 h incubations at 37°C, the SCFA profiles were determined on culture supernatants (Richardson *et al.*, 1989).

## Acknowledgements

We would like to thank Thanasis Vogogias, David Nutter and Alec Mann for their assistance in developing the software for this model. We also acknowledge the Scottish Government's Rural and Environment Science and Analytical Services Division (RESAS) for their financial support. Furthermore, many thanks go to the two anonymous reviewers whose hard work has greatly improved this paper.

## References

- Aminov, R.I., Walker, A.W., Duncan, S.H., Harmsen, H.J.M., Welling, G.W., and Flint, H.J. (2006) Molecular detection, cultivation and improved FISH detection of a dominant group of human gut bacteria related to *Roseburia* and *Eubacterium rectale*. *Appl Environ Microbiol* **72**: 6371–6376.
- Bagramyan, K., and Trchounian, A. (2003) Structural and functional features of formate hydrogen lyase, an enzyme of mixed-acid fermentation from *Escherichia coli*. *Biochemistry (Mosc)* **68**: 1159–1170.
- Barcenilla, A., Pryde, S.E., Martin, J.C., Duncan, S.H., Stewart, C.S., Henderson, C., and Flint, H.J. (2000) Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Appl Environ Microbiol* **66**: 1654–1661.
- Bratbak, G., and Dundas, I. (1984) Bacterial dry matter content and biomass estimations. *Appl Environ Microbiol* **48**: 755–757.
- Cotillard, A., Kennedy, S.P., Kong, L.C., Prifti, E., Pons, N., Le Chatelier, E., et al. (2013) Dietary intervention impact on gut microbial gene richness. *Nature* **500**: 585–588.
- David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., et al. (2014) Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**: 559–563.
- Drake, H.L., Göbner, A.S., and Daniel, S.L. (2008) Old acetogens, new light. *Ann N Y Acad Sci* **1125**: 100–128.
- Duncan, S.H., and Flint, H.J. (2008) Proposal of a neotype strain (A1-86) for *Eubacterium rectale*. Request for an opinion. *Int J Syst Evol Microbiol* **58**: 1735–1736.
- Duncan, S.H., Hold, G.L., Harmsen, H.J.M., Stewart, C.S., and Flint, H.J. (2002) Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **52**: 2141–2146.
- Duncan, S.H., Louis, P., and Flint, H.J. (2004a) Lactate-utilising bacteria, isolated from human feces, that produce butyrate as a major fermentation product. *Appl Environ Microbiol* **70**: 5810–5817.
- Duncan, S.H., Holtrop, G., Lobley, G.E., Calder, G., Stewart, C.S., and Flint, H.J. (2004b) Contribution of acetate to butyrate formation by human faecal bacteria. *Br J Nutr* **91**: 915–923.
- Duncan, S.H., Aminov, R.I., Scott, K.P., Louis, P., Stanton, T.B., and Flint, H.J. (2006) Proposal of *Roseburia faecis* sp. nov., *Roseburia hominis* sp. nov. and *Roseburia inulinivorans* sp. nov., based on isolates from human faeces. *Int J Syst Evol Microbiol* **56**: 2437–2441.
- Duncan, S.H., Louis, P., Thomson, J.M., and Flint, H.J. (2009) The role of pH in determining the species composition of the human colonic microbiota. *Environ Microbiol* **11**: 2112–2122.
- Eckburg, P.B., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., et al. (2005) Diversity of the human intestinal microbial flora. *Science* **308**: 1638.
- Egert, M., de Graaf, A.A., Smidt, H., de Vos, W.M., and Venema, K. (2006) Beyond diversity: functional microbiomics of the human colon. *Trends Microbiol* **14**: 86–91.
- Feijó Delgado, F., Cermak, N., Hecht, V.C., Son, S., Li, Y., Knudsen, S.M., et al. (2013) Intracellular water exchange for measuring the dry mass, water mass and changes in chemical composition of living cells. *PLoS ONE* **8**: e67590. doi:10.1371/journal.pone.0067590.
- de Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J.B., Massart, S., et al. (2010) Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci USA* **107**: 14691–14696.
- Flint, H.J., Duncan, S.H., Scott, K.P., and Louis, P. (2007) Interactions and competition within the microbial community of the human colon: links between diet and health. *Environ Microbiol* **9**: 1101–1111.
- Flint, H.J., Scott, K.P., Louis, P., and Duncan, S.H. (2012a) The role of the gut microbiota in nutrition and health. *Nat Rev Gastroenterol Hepatol* **9**: 577–589.
- Flint, H.J., Scott, K.P., Duncan, S.H., Louis, P., and Forano, E. (2012b) Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* **3**: 289–306.
- Gottschalk, G. (1979) *Bacterial Metabolism*. New York, Heidelberg, Berlin: Springer Verlag.
- Kettle, H., Donnelly, R., Flint, H.J., and Marion, G. (2014) pH feedback and phenotypic diversity within bacterial functional groups of the human gut. *J Theor Biol* **342**: 62–69.
- Kovarova-Kovar, K., and Egli, T. (1998) Growth kinetics of suspended microbial cells: from single-substrate-controlled growth to mixed substrate kinetics. *Microbiol Mol Biol Rev* **62**: 646–666.
- Latham, M.J., and Wolin, M.J. (1977) Fermentation of cellulose by *Ruminococcus flavefaciens* in the presence and absence of *Methanobacterium ruminantium*. *Appl Environ Microbiol* **34**: 297–301.
- Liu, Y., and Whitman, W.B. (2008) Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Ann N Y Acad Sci* **1125**: 171–189.
- Lopez-Siles, M., Khan, T.M., Duncan, S.H., Harmsen, H.J.M., Garcia-Gil, L.J., and Flint, H.J. (2012) Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. *Appl Environ Microbiol* **78**: 420–428.
- Louis, P., Scott, K.P., Duncan, S.H., and Flint, H.J. (2007) Understanding the effects of diet on bacterial metabolism in the large intestine. *J Appl Microbiol* **102**: 1197–1208.
- Louis, P., Young, P., Holtrop, G., and Flint, H.J. (2010) Diversity of human colonic butyrate-producing bacteria revealed by analysis of the butyryl-CoA:acetate CoA-transferase gene. *Environ Microbiol* **12**: 304–314.

- Macfarlane, G.T., and Gibson, G.R. (1997) Carbohydrate fermentation, energy transduction and gas metabolism in the human large intestine. In *Gastrointestinal Microbiology*, Vol. 1. Mackie, R.I., and White, B.A. (eds). New York, NY, USA: Chapman and Hall, pp. 269–318.
- Marquet, P., Duncan, S.H., Chassard, C., Bernalier-Donadille, A., and Flint, H.J. (2009) Lactate has the potential to promote hydrogen sulphide formation in the human colon. *FEMS Microbiol Lett* **299**: 128–134.
- Moore, W.E.C., Cato, E.P., and Holdeman, L.V. (1972) *Ruminococcus bromii* sp. n. and emendation of the description of *Ruminococcus* Sijpestein. *Int J Syst Bacteriol* **22**: 78–80.
- Munoz-Tamayo, R., Laroche, B., Walert, E., Dore, J., and Leclerc, M. (2010) Mathematical modelling of carbohydrate degradation by human colonic microbiota. *J Theor Biol* **266**: 189–201.
- Nava, G.M., Carbonero, F., Croix, J.A., Greenberg, E., and Gaskins, H.R. (2012) Abundance and diversity of mucosa-associated hydrogenotrophic microbes in the healthy human colon. *ISME J* **6**: 57–70.
- Norberg, J., Swaney, D.P., Dushoff, J., Lin, J., Casagrandi, R., and Levin, S.A. (2001) Phenotypic diversity and ecosystem functioning in changing environments: a theoretical framework. *Proc Natl Acad Sci USA* **98**: 11376–11381.
- Prins, R.A. (1977) Biochemical activities of gut microorganisms. In *Microbial Ecology of the Gut*. Clarke, R.T.J., and Bauchop, T. (eds). London, UK: Academic Press, pp. 73–183.
- Qin, J., Ruiqiang, L., Raes, J., Arumugam, M., Solvsten Burgdorf, K., Manichanh, C., *et al.* (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**: 59–65.
- Ragsdale, S.W., and Pierce, E. (2008) Acetogenesis and the Wood-Ljungdahl pathway of CO<sub>2</sub> fixation. *Biochim Biophys Acta* **1784**: 1873–1898.
- Reichardt, N., Duncan, S.H., Young, P., Belenguer, A., McWilliam Leitch, C., Scott, K.P., *et al.* (2014) Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISMEJ* **8**: 1323–1335.
- Richardson, A.J., Calder, A.G., Stewart, C.S., and Smith, A. (1989) Simultaneous determination of volatile and non-volatile acidic fermentation products of anaerobes by capillary gas chromatography. *Lett Appl Microbiol* **9**: 5–8.
- Robert, C., and Bernalier-Donadille, A. (2003) The cellulolytic microflora of the human colon: evidence of microcrystalline cellulose-degrading bacteria in methane-excreting subjects. *FEMS Microbiol Ecol* **46**: 81–89.
- Rogers, T.E., Pudlo, N.A., Koropatkin, N.M., Bell, J.S.K., Balasch, M.M., Jasker, K., and Martens, E.C. (2013) Dynamic responses of *Bacteroides thetaiotaomicron* during growth on glycan mixtures. *Mol Microbiol* **88**: 876–890.
- Russell, J.B., and Baldwin, R.L. (1978) Substrate preferences in rumen bacteria: evidence of catabolite regulatory mechanisms. *Appl Environ Microbiol* **36**: 319–329.
- Salyers, A.A., Vercellotti, J.R., West, S.E.H., and Wilkins, T.D. (1977) Fermentation of mucins and plant polysaccharides by strains of *Bacteroides* from the human colon. *Appl Environ Microbiol* **33**: 319–322.
- Seeliger, S., Janssen, P.H., and Schink, B. (2002) Energetics and kinetics of lactate fermentation to acetate and propionate via methylmalonyl-CoA or acrylyl-CoA. *FEMS Microbiol Lett* **211**: 65–70.
- Stephen, A.M., Wiggins, H.S., and Cummings, J.H. (1987) Effect of changing transit time on colonic microbial metabolism in man. *Gut* **28**: 601–609.
- Swidsinski, A., Loening-Baucke, V., Verstraelen, H., Osowska, S., and Doerffel, Y. (2008) Biostructure of fecal microbiota in healthy subjects and patients with chronic idiopathic diarrhea. *Gastroenterology* **135**: 568–579.
- Tap, J., Mondot, S., Levenez, F., Pelletier, E., Caron, C., Furet, J.-P., *et al.* (2009) Towards the human intestinal microbiota phylogenetic core. *Environ Microbiol* **11**: 2574–2584.
- Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature* **457**: 480–484.
- Walker, A.W., Duncan, S.H., McWilliam Leitch, E.C., Child, M.W., and Flint, H.J. (2005) pH and peptide can radically alter bacterial populations and short-chain fatty acid ratios within microbial communities from the human colon. *Appl Environ Microbiol* **71**: 3692–3700.
- Walker, A.W., Duncan, S.H., Harmsen, H.J.M., Holtrop, G., Welling, G.W., and Flint, H.J. (2008) The species composition of the human intestinal microbiota differs between particle-associated and liquid phase communities. *Environ Microbiol* **10**: 3275–3283.
- Walker, A.W., Ince, J., Duncan, S.H., Webster, L.M., Holtrop, G., Ze, X., *et al.* (2011) Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J* **5**: 220–230.
- Whitman, W.B., Coleman, D.C., and Wiebe, W.J. (1998) Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* **95**: 6578–6583.
- Wolin, M.J., Miller, T.J., Collins, M.D., and Lawson, P.A. (2003) Formate-dependent growth and homoacetogenic fermentation by a bacterium from human feces: description of *Bryantella formatexigenes* gen. nov., sp. nov. *Appl Environ Microbiol* **69**: 6321–6326.
- Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S.A., *et al.* (2011) Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**: 105–108.
- Yachi, S., and Loreau, M. (1999) Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc Natl Acad Sci USA* **96**: 1463–1468.
- Ze, X., Duncan, S.H., Louis, P., and Flint, H.J. (2012) *Ruminococcus bromii* is a keystone species for the degradation of resistant starch in the human colon. *ISME J* **6**: 1535–1543.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Table S1.** Metabolic stoichiometries (relative number of moles of resource that are consumed or produced by each BFG).

**Table S2.** Estimated relative maximum growth rates ( $\mu_{fj}^M$ , dimensionless – for absolute values multiply by  $\mu_{max}$ , which is  $24\text{ d}^{-1}$ ) for each BFG.

**Table S3.** Properties of resources: molar mass ( $m$ ,  $\text{g mol}^{-1}$ ), and yield ( $Y\text{ g g}^{-1}$ ).

**Table S4.** The centre of the pH preference range  $\text{pH}_m$  for use in the pH limitation function  $H_i$  [see Appendix S1 (A2.4 and A5) for details on the pH limitation function].

**Table S5.** Initial (or fixed) conditions for the fermentor experiments (Walker *et al.*, 2005).

**Table S6.** Assigning the probes used in the fermentor experiment by Walker and colleagues (2005) to the bacterial functional groups in the model.

**Table S7.** Percentage covered by individual probes used for deriving data for the ten bacterial groups, expressed as percentage of Eub probe.

**Table S8.** Model variables and parameters (see Appendix S1 for details and usage).

**Table S9.** The metabolic pathways for each BFG. Consumption and production of dietary substrates (left) and metabolites (right) by each BFG.

**Appendix S1.** Describing the governing equations and assigning strain traits.

**Appendix S2.** Fermentation pathways for each BFG.

**Appendix S3.** Describing a sensitivity analysis of model.

**Appendix S4.** AMES\_SourceCode.tar.gz Model code (Fortran 90).