The arginine dihydrolase pathway shapes human gut microbial community assembly, functions, and mammalian host colonization

Yiyi Liu[1,2], Yu-Yu Cheng[2], Zhichao Zhou[3], Eugenio I. Vivas[3,4], Karthik Anantharaman[3] and Ophelia S. Venturelli[1,2,3*]

[1]Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison WI 53706

[2]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706

[3]Department of Bacteriology, University of Wisconsin-Madison, WI 53706

[4]Gnotobiotic Animal Core Facility, University of Wisconsin-Madison, Madison, WI 53706, USA

*To whom correspondence should be addressed: venturelli@wisc.edu

Keywords: arginine dihydrolase pathway, pH, synthetic biology, human gut microbial community, computational modeling

ABSTRACT

The arginine dihydrolase metabolic pathway (*arc* operon) is present in the genomes of diverse human gut bacterial species and can modulate environmental pH. We lack an understanding of the role of this pathway on community-level interactions. We investigated the role of the *arc* operon in the probiotic *E. coli* Nissle on the assembly and production of the beneficial metabolite butyrate in a human gut community *in vitro* and in the murine gut. The presence of the *arc* operon shapes the dynamics of community assembly. With *arc* operon expression, we observed reduced variability in community structure from environmental pH perturbations. In particular, the abundance of butyrate producers was altered either transiently or longer-term. The *arc* operon can enhance or reduce butyrate production depending on the environmental pH. Dynamic computational modeling of community assembly reveals the extent of pH-mediated inter-species interactions on community assembly. Several species are identified with significant beneficial or detrimental non-pH-mediated interactions. Overall, we demonstrated that a single specialized metabolic pathway could shape community assembly, species colonization of the mammalian gut and butyrate production.

INTRODUCTION

The human gut microbiome substantially expands our genome's capabilities by encoding numerous specialized metabolic pathways[1–3]. Certain metabolic pathways are frequently found in constituent members of the gut microbiota, such as the enzymes to produce acetate[4]. However, other specialized metabolic pathways such as enzymes for modifying primary bile acids are present in a minority of resident gut species[5–7]. These specialized metabolic pathways can be harbored by multiple distinct taxa and contribute uniquely to the functions performed by gut microbiota. Addition or subtraction strategies of specialized microbial functions are also more accessible for engineering the microbiome because they are harbored by a minority of taxa[8,9].

Specialized metabolic pathways influence the interactions between constituent members of gut microbiota and their dynamic behaviors and functions. For example, *Bacteroides spp.* harbor polysaccharide utilization loci (PUL) that enable the breakdown of complex carbohydrates. These PULs have been shown to generate competition for limiting nutrient availability as well as enhancement of other constituent community members that can utilize the breakdown products and not the fiber for growth[3]. However, we lack an understanding of the role of other specialized metabolic pathways beyond polysaccharide utilization loci on the interactions in the gut microbiome. Specifically, butyrate production by the gut microbiome has links to several human diseases[10–12], yet predictable modulation of butyrate level via supplementation of butyrate-producing species is challenging and cannot be fully explained by PUL-related interactions[13,14], calling for characterization of specialized metabolic pathways beyond PULs.

The arginine dihydrolase pathway (*arc* operon) is a specialized metabolic pathway that converts arginine into several byproducts including ammonia, which could impact the gut community[15]. This pathway is harbored by many diverse gut species enabling utilization of arginine as an energy source under aerobic and anaerobic acidic conditions[16–18]. By releasing ammonia, the *arc* operon can increase environmental pH and provide additional nitrogen resource in nitrogen-limited environments such as the mammalian gut[19]. Specifically, modification of pH has been shown to impact *in vitro* gut community structure assembly and metabolite production in communities derived from human fecal samples[20,21]. The pH in mammalian gastrointestinal (GI) tract can also play a role in host health and gut ecosystem functions[22,23], suggesting the *in vivo* relevance of studying the effects of the *arc* operon expression.

Computational modeling of microbiomes can reveal the quantitative relationships between different species and their impact on community-level functions[24,25]. The generalized Lotka-Volterra (gLV) is a coupled ordinary differential equation model that captures species intrinsic growth and pairwise interactions. This model has been used to predict community assembly in different environments[26–29]. However, gLV models coarse grains the mechanisms of inter-species interactions by representing each interaction as a lumped constant parameter. Other computational models of microbial interactions have been proposed that focus on specific mechanism modalities. For example, a computational model that captures pH-mediated interactions captured pairwise interactions in a soil microbial community[30].

By combining bottom-up design of synthetic communities, computational modeling, and synthetic biology, we investigated the role of the *arc* operon in the human-associated probiotic strain *Escherichia coli* Nissle 1917 (EcN) in shaping human gut community assembly, butyrate production, and mammalian colonization. Exploiting genetic control, we engineered inducible control of the *arc* operon of EcN and investigated the effect of this pathway on the assembly of a defined human gut community both *in vitro* and in the murine gut. The variability in community structure due to changes in environmental pH significantly reduced under *arc* operon expression.

The presence of the *arc* operon caused transient or sustained changes in the abundance of butyrate producing species. In addition, the *arc* operon significantly enhanced or reduced butyrate production depending on the initial environmental pH. To quantify the role of pH modification of the *arc* operon on community assembly, a dynamic computational model was used to forecast community assembly as a function of monospecies growth and pH modification. The computational model revealed the quantitative contribution of pH-mediated interactions in a species-specific manner. In sum, we demonstrate that a single specialized metabolic pathway can shape community assembly, species colonization of the mammalian gut and butyrate production. A more quantitative and detailed understanding of how specialized metabolic pathways shape gut microbiota assembly and functions holds promise for engineering the gut microbiome to our benefit.

RESULTS

Activation of the arginine dihydrolase pathway increases environmental pH through arginine metabolism

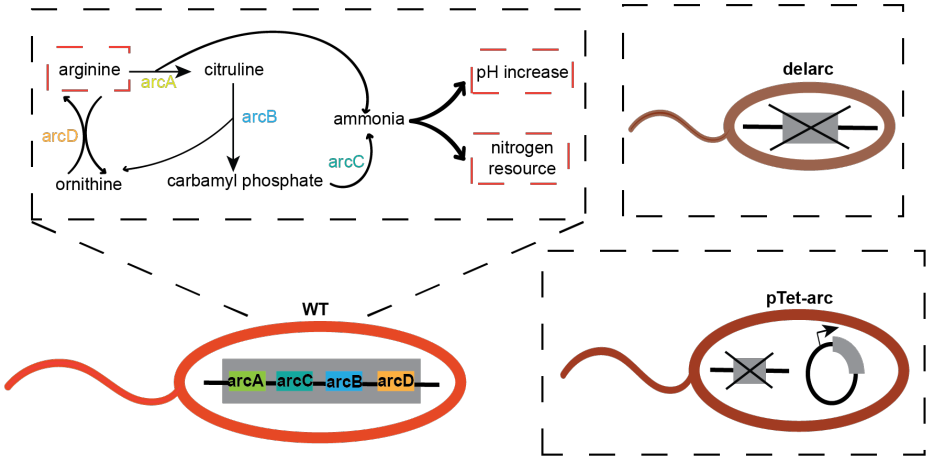
The arginine dihydrolase pathway is encoded by the *arc* operon in bacterial species, constituted by four genes, with mechanism illustrated in Figure 1A. Arginine deiminase, encoded by *arcA*, is responsible for arginine metabolism into citrulline, producing ammonia as a byproduct; Ornithine carbamoyltransferase, encoded by *arcB*, catalyze the conversion from citrulline to ornithine and carbamyl phosphate; Carbamate kinase, encoded by arcC, further metabolizes carbamyl phosphate into more ammonia; Finally, the arginine ornithine antiporter, encoded by *arcD*, catalyzes the electroneutral exchange between arginine and ornithine to enable efficient conversion in the arginine dihydrolase pathway. Altogether, the ammonia produced in the pathway increases environmental pH by conversion into ammonium, along with providing additional nitrogen resource to the environment.

To determine the prevalence of the arginine dihydrolase pathway in human gut microbiome, bioinformatics search was performed on a metagenome-assembled genome dataset containing 154,723 microbial genomes assembled from 9,428 samples of the human microbiome from diverse geographic location, body sites, diseases, and lifestyles. Out of the 533 total species identified from the database, 63 species were found to harbor the entire *arc* operon (Figure 1B), constituting 12% of all species. Therefore, the *arc* operon was determined to be a specialized metabolic operon for human gut microbes due to its low abundance among all microbes colonizing the human gut environment. Additionally, the operon was rarely found in Bacteroidetes, one of the most abundant phyla in human gut microbiome.

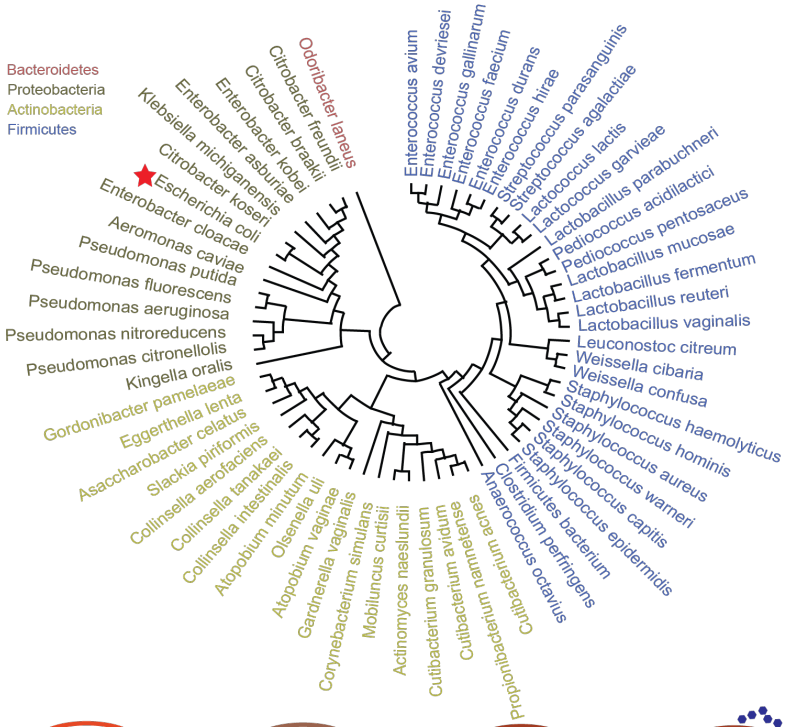
To investigate the impact of the *arc* operon on human gut microbiome, the gut probiotic strain *Escherichia coli* Nissle 1917 (EcN) which harbors the *arc* operon was selected for further engineering due to its extensive genetic modification toolbox and reported health relevance to human gut microbiome[31]. In order to assess impact of the *arc* operon on the fitness of EcN, the arginine dihydrolase pathway was first deleted from the genome of wildtype EcN (wt strain), constructing the delarc strain as negative control of *arc* operon expression (Figure 1A). To investigate the impact of inducible higher *arc* operon expression than the wt strain, the arginine dihydrolase pathway was re-introduced to the delarc strain under anhydrotetracycline chemical inducer control on a plasmid with kanamycin resistance, constructing the pTet-arc strain (Figure 1A). The higher gene expression in induced pTet-arc strain was verified by qPCR of RNA extracted at exponential phase of *in vitro* monoculture (Figure S5F).

The wt, delarc, and pTet-arc strains were grown anaerobically overnight in media supplemented with additional arginine with and without inducer at various starting pH ranging from 4 to 9, with timepoint measurements of pH and OD taken (Figure 1C). Overall, growth of EcN was enhanced at higher starting pH. The delarc strain and pTet-arc strain without inducer (pTet-arc (-)) exhibited similar growth behavior in OD and pH influence, demonstrating that the pTet promoter had low leakiness. While wt strain and pTet-arc strain with inducer (pTet-arc (+)) both resulted in higher environmental pH compared to delarc and pTet-arc (-) strains, pTet-arc (+) was found to increase pH earlier, resulting in considerable growth in lower starting pH (pH=5). Additionally, the end point OD after 24-hour incubation for pTet-arc (+) was the highest among all strains and conditions, exhibiting better growth independent of the initial media pH. The expression of the *arc* operon was confirmed to increase environmental pH, leading to better growth of EcN in lower starting pH.

A


arginine
arcA
citruline
arcB
ammonia
pH increase
nitrogen resource
arcD
ornithine
carbamyl phosphate
arcC
delarc
WT
arcA
arcC
arcB
arcD
pTet-arc

B


Bacteroidetes
Proteobacteria
Actinobacteria
Firmicutes
Odoribacter laneus
Citrobacter freundii
Citrobacter braakii
Citrobacter koseri
Enterobacter asburiae
Enterobacter cloacae
Klebsiella michiganensis
Escherichia coli
Aeromonas caviae
Pseudomonas putida
Pseudomonas fluorescens
Pseudomonas aeruginosa
Pseudomonas nitroreducens
Pseudomonas citronellolis
Kingella oralis
Gordonibacter pamelaeae
Eggerthella lenta
Asaccharobacter celatus
Slackia piriformis
Collinsella aerofaciens
Collinsella tanakaei
Collinsella intestinalis
Atopobium minutum
Olsenella uli
Atopobium vaginae
Gardnerella vaginalis
Corynebacterium simulans
Mobiluncus curtisii
Actinomyces naeslundii
Cutibacterium granulosum
Cutibacterium avidum
Cutibacterium namnetense
Propionibacterium acnes
Anaerococcus octavius
Clostridium perfringens
Firmicutes bacterium
Staphylococcus epidermidis
Staphylococcus capitis
Staphylococcus warneri
Staphylococcus aureus
Staphylococcus hominis
Staphylococcus haemolyticus
Weissella confusa
Weissella cibaria
Leuconostoc citreum
Lactobacillus vaginalis
Lactobacillus reuteri
Lactobacillus fermentum
Lactobacillus mucosae
Pediococcus pentosaceus
Pediococcus acidilactici
Lactobacillus parabuchneri
Lactococcus garvieae
Lactococcus lactis
Streptococcus agalactiae
Streptococcus parasanguinis
Streptococcus hirae
Enterococcus durans
Enterococcus faecium
Enterococcus gallinarum
Enterococcus devriesei
Enterococcus avium

C

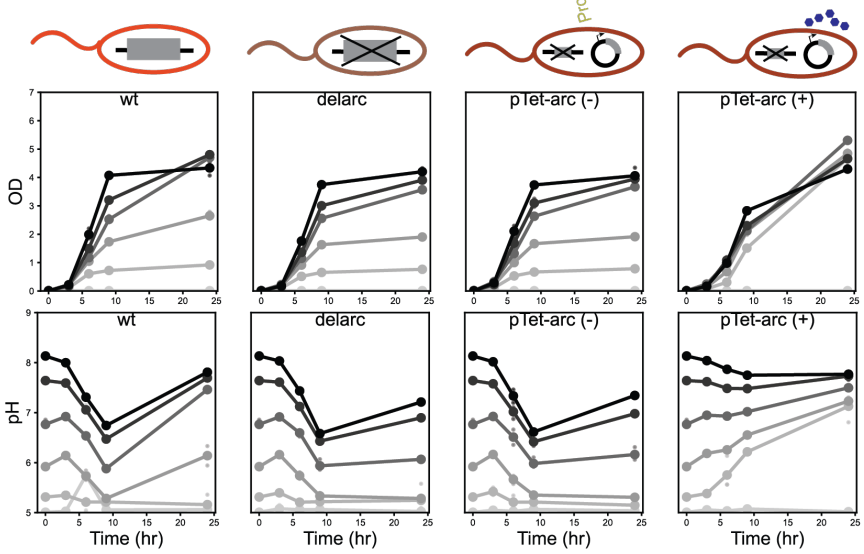

wt
delarc
pTet-arc (-)
pTet-arc (+)
OD
pH
Time (hr)

Figure 1 | (A) Mechanism of arginine dihydrolase pathway and strains engineered in this study. The *arc* operon encodes four color-coded genes, each responsible for various reactions in arginine metabolism. EcN strains with *arc* deletion and *arc* under inducible promoter on a plasmid were cloned for this study. (B) Phylogenetic tree of species harboring the arc operon in the human gut microbiome generated by metagenomic analysis. Species color indicates phylum. The species of interest in this study (*E. coli*) is marked by a red star. (C) Absolute abundance (calculated OD600) and supernatant pH of various EcN strains under different initial media pH ranging from 4 to 9 over time. Species were inoculated at an equal abundance based on OD600 measurements. Smaller sized points indicate experimental data replicates. Larger sized points and lines indicate average over all biological replicates. Different initial media pH is marked by different color of line and datapoints, with darker color corresponding to higher initial media pH.

Expression of the *arc* operon impacts community assembly in a synthetic gut microbial community *in vitro*

A previous study demonstrated that the *arc* operon provides an advantage for a pathogenic *E. coli* wildtype strain *in vitro* in the presence of arginine while conferring a disadvantage in the murine gut [16]. To verify this in Nissle, we performed *in vitro* anaerobic coculture of wt strain and delarc strain with arginine and initial pH titration in media as well as colonization competition of wt strain and delarc strain in a germ-free murine model (Figure S1). Our experiments confirmed *arc* operon confers advantage *in vitro* and colonization disadvantage *in vivo* for EcN. The pH of fecal and cecal samples had little variation across different days of the experiment in contrast to *in vitro* pH trends, indicating host homeostasis and other host-microbe interactions dominate pH regulation in gut.

We next sought to investigate the impact of the *arc* operon in *in vitro* community assembly. An 8-member synthetic gut microbial community (comm11) was assembled based on prevalence in human gut microbiome and lack of the *arc* operon. As a representative gut microbial community, the chosen consortium spanned the major human gut phyla *Bacteroidetes*, *Firmicutes*, and *Actinobacteria* (Figure 2A). The *in vitro* experimental media condition was optimized based on a previously identified chemically defined medium (DM38) that can support growth of all species in comm11[27]. Arginine concentration was selected for highest functional expression of the *arc* operon without EC takeover in communities (Figure S2).

Comm11 species were cocultured with either the pTet-arc (+) stain or the delarc strain for 24 hours anaerobically at various starting pH (Figure 2B). Timepoint pH data revealed distinguishing pH profiles between pTet-arc (+) condition and delarc condition, with higher pH in pTet-arc (+) condition observed in all timepoint and all initial pH than delarc condition, suggesting the *arc* operon can functionally express and increase environmental pH in a microbial community *in vitro* (Figure 2C). Relative abundances calculated from NGS data in either condition revealed distinguishing species relative abundance profiles as well (Figure 2D, S3A-D). Specifically, the species relative abundance comparison at 24-hour revealed similar trends to monoculture data (Figure S6, S7), where relative abundance for species that prefer lower pH (BL, CC) significantly increased in delarc condition while relative abundance for species that prefer higher pH (AC, BC, BT, EC) significantly increased in wt condition. This observation is verified by the integral species abundance over all experimental timepoints during coculture, which showed similar sensitivity to initial pH as integral OD of monoculture for all species except DF, which did not grow in monoculture at any initial pH (Figure S4E). The similarity between species sensitivity to pH in monoculture and in community suggest the influence of pH on a particular species is generally the same either in monoculture or in community. We also observed opposite relative abundance dynamic in earlier timepoint compared to the endpoint in AC (Figure S4 A-D), where, specifically in initial pH = 7 and 8 conditions, the relative abundance of AC is either higher or lower in pTet-arc (+) condition in earlier timepoints, but flipped in later timepoints, demonstrating the transient effect of the *arc* operon on species assembly.

In addition to community structure changes, we observed shift in production of butyrate, a beneficial gut microbe-derived metabolite (Figure 2E). At lower initial pH (pH = 5, 6), the butyrate concentration in 24-hour coculture supernatant is significantly higher in pTet-arc (+) condition compared to the delarc condition; however, at higher initial pH (pH=8), the reverse trend is observed where higher butyrate concentration is observed in delarc condition. This is potentially due to butyrate producing species AC significantly enriched at lower pH with *arc* operon expression while its growth is inhibited at higher pH due to EC outgrowth. Interestingly, the relative

abundance of another butyrate producing species CC was lower in pTet-arc (+) conditions, demonstrating the different effect of *arc* operon expression on different butyrate producers. The Euclidean distance between community structure at the endpoint of each initial pH condition was found to be lower in pTet-arc (+) condition than delarc condition (Figure 2F), suggesting the *arc* operon expression can stabilize in vitro community in response to pH changes. Overall, expression of the *arc* operon in EcN can modify supernatant pH, stabilize community assembly to pH changes, and enhance butyrate production at lower pH in a synthetic gut microbial community *in vitro*.

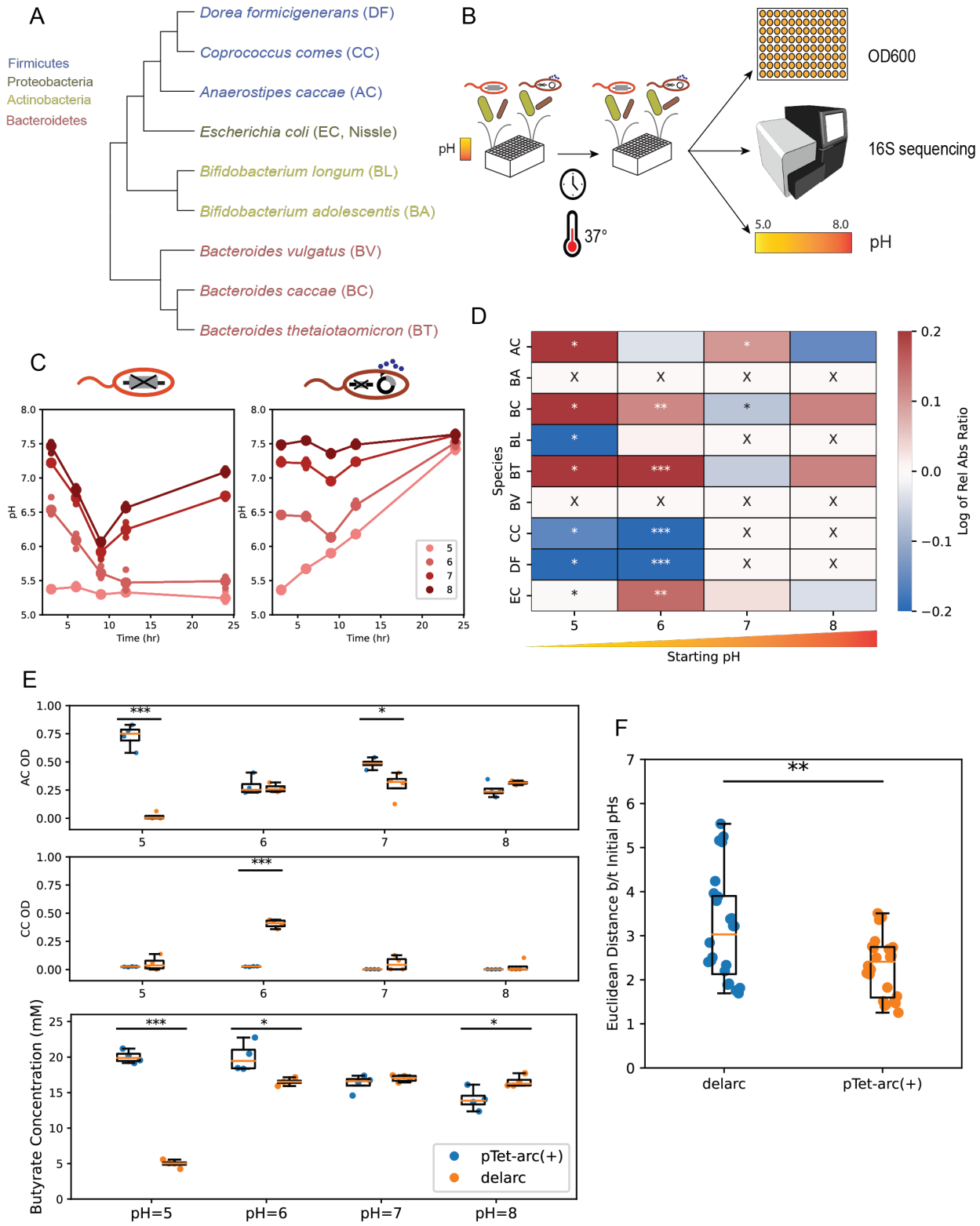


Figure 2 | (A) Phylogenetic tree of the synthetic human gut microbial community composed of 8 highly prevalent and diverse species. The species of interest harboring the *arc* operon (*E. coli*) is also included in the tree, with strain name noted. Species color indicates phylum. (B) Schematic

of the experimental workflow of *in vitro* comm11 coculture with various strains of EcN at various starting pH. Synthetic communities are cultured in microtiter plates in anaerobic conditions and incubated at 37°C. The absolute abundance of each species is determined by measuring cell density at 600 nm (OD600). The supernantent pH is taken by PhenolRed method. The Community composition is determined using multiplexed 16S rRNA sequencing. (C) Supernatant pH of comm11 coculture with delarc and pTet-arc (+) strains of EcN under different initial media pH ranging from 5 to 8. Species were inoculated at an equal abundance based on OD600 measurements. Datapoints indicate experimental data replicates. Lines indicate average over all biological replicates. Different initial media pH is marked by different color of line and datapoints, with darker color corresponding to higher initial media pH. (D) Heatmap of the log of the ratio between species relative abundance (mean value of biological replicates, n = 3) in pTet-arc (+) condition and delarc condition at various initial media pH. Species relative abundances are determined by multiplexed 16S rRNA sequencing. Red corresponds to higher relative abundance in wt conditions and blue corresponds to higher relative abundance in delarc conditions. Low growth conditions are marked by an 'X'. Asterisks represent statistical significance: *P < 0.05, **P < 0.01, ***P < 0.001 according to an unpaired t-test. (E) Box plot of butyrate concentration (mM) at 24-hour for comm11+EcN cocultures with either delarc strain or pTet-arc(+) strain at various starting pH. Datapoints indicate experimental data replicates. Datapoint colors indicate strain of EcN in coculture. The calculated species OD of butyrate producer species within comm11 were displayed in the same plot at the same initial pH and strain coculture conditions. The species OD was calculated by multiplying coculture OD with species relative abundance obtained from NGS. Independent statistical t-test is performed on each pair of conditions with the same starting pH. Asterisks represent statistical significance: *P < 0.05, **P < 0.01, ***P < 0.001 according to an unpaired t-test. (F) Community structure Euclidean distances between all pairs of initial pH conditions for delarc condition (blue) and pTet-arc(+) condition(orange). The Euclidean distances were calculated based on community relative abundance structure between each pair of initial pH condition for a particular strain condition. Asterisks represent statistical significance: *P < 0.05, **P < 0.01, ***P < 0.001 according to an unpaired t-test.

The *in vivo* impacts of the *arc* operon on colonization and community assembly exhibit opposite trends than *in vitro*

We next sought to investigate impact of the *arc* operon on community assembly *in vivo* through a murine model. Gnotobiotic male mice aged 6-8 weeks were place on a customized high arginine diet 3 days before the introduction of bacteria. Bacterial suspension of comm11 species at equal density was introduced to mice by oral gavage. The fecal samples were collected every 2-3 days, with the pH and strain relative abundance of fecal measured. After one week of comm11 colonization to reach stability in species compositions, mice were separated into three groups, with two groups introducing either wt or delarc strains of EcN by oral gavage and one group continuing comm11 colonization. After another two weeks, mice were sacrificed with cecal sample collected for pH and relative abundance analysis (Figure 3A). The expression of the *arc* operon in wt conditions and deletion of the *arc* operon in delarc conditions *in vivo* were confirmed by real-time quantitative reverse transcription PCR (qRT-PCR) (Figure S5F).

Contrary to *in vitro* results where EC made up 60% - 80% of the population in comm11 and EcN cocultures (Figure S3A-D), neither wt or delarc strains of EcN dominated the community *in vivo*, constituting at most 30% of the population (Figure 3B, S5A-B). The abundance for wt and delarc strains initially increased at the same rate until reaching a maximum at day 4 after introduction. Then, the delarc strain relative abundance decreased at a significantly slower rate than that of the wt strain, conferring colonization disadvantage to the wt strain. This result is in accordance with previously observed different relative abundance ranges among *in vitro* and *in vivo* experiments of delarc versus wt competition (Figure S1), suggesting the important role of host-microbe interaction in *in vivo* colonization leading to potentially reversed trends when scaling *in vitro* experiment into *in vivo*. The fecal pH was found to be statistically identical for different groups (Figure S5E), suggesting host homeostasis regulating pH regardless of presence of *arc*-harboring strains.

Species compositions of comm11 were compared between comm11+wt and comm11+delarc conditions (Figure 3D). After introduction of EcN at Day 7, certain species were significantly enriched in wt conditions at certain timepoints (CC, BC) while some significantly lower in coculture with wt (AC, DF) at certain timepoints. Additionally, the change in species relative abundance was found to be transient for BC and DF, only showing significance at timepoint 16. However, for CC and AC, the difference was found to be sustained at multiple timepoints. This suggested varying susceptibility of gut microbes to the *arc* operon, where some did not vary significantly (BA, BL, BT, BV), some were affected transiently (CC, BC), and some experienced sustained difference on colonization (AC, DF), which is in accordance with *in vitro* observations where species have different response to EcN strains (Figure S4). Interestingly, the CFU counts for wt and delarc conditions did not vary significantly (Figure S5D), suggesting the difference in EC relative abundance is likely due to the enrichment of other species in the presence of the *arc* operon. Along with the observation where pH did not significantly change across conditions and timepoints, our experimental result suggest the *arc* operon is affecting colonization of other species in community through non-pH mediated mechanisms. Altogether, the *arc* operon is shown to have the potential of mediating community assembly and colonization in murine model with varying temporal dynamics.

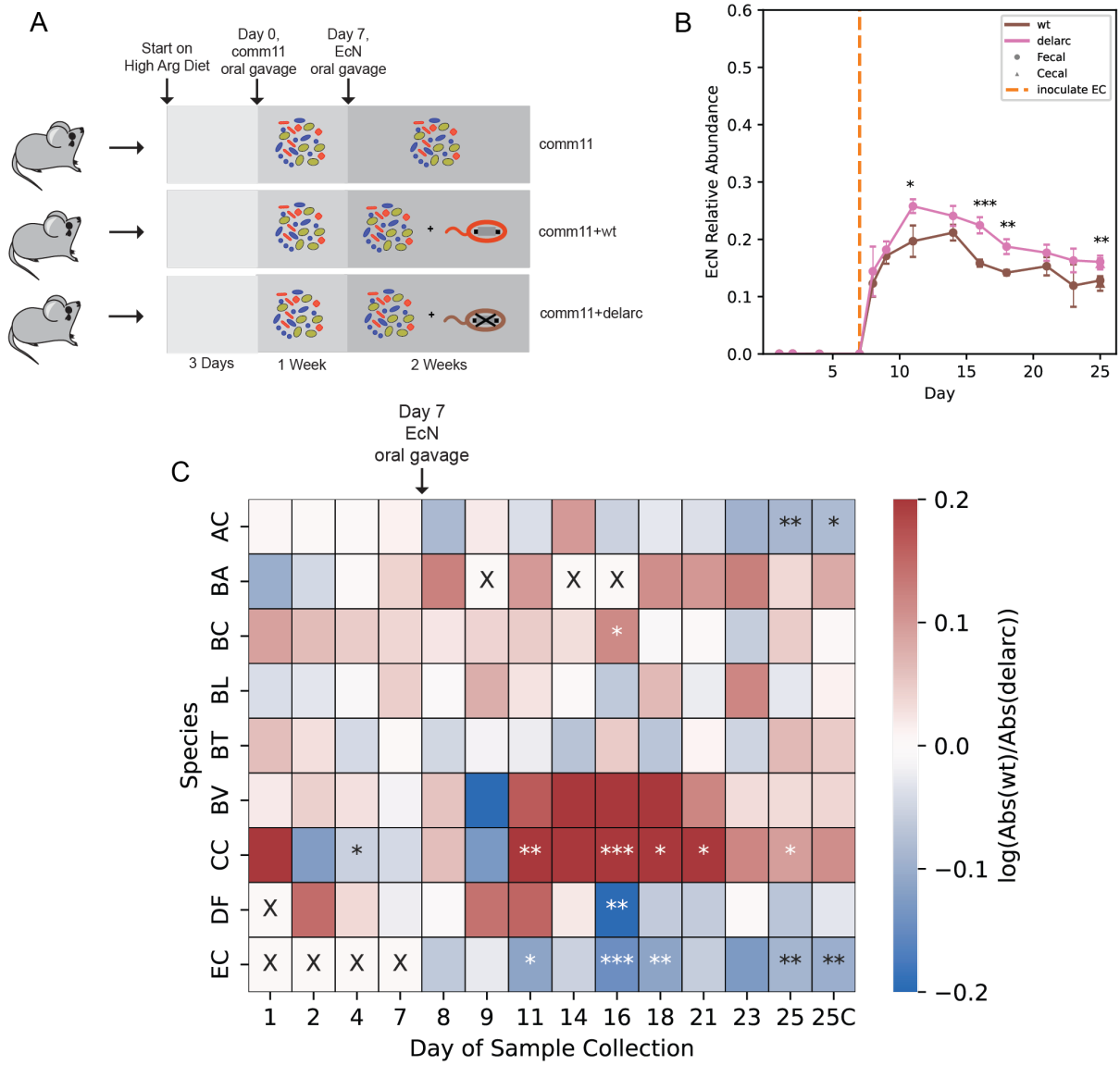


Figure 3 | (A) Experimental workflow of *in vivo* comm11 coculture with various strains of EcN. Mice were tagged and placed on customized high arginine diet 3 days before oral gavage of equal part comm11 bacterial suspension. Fecal samples were collected every two to three days for determination of species relative abundances by multiplexed 16S rRNA sequencing and fecal sample pH by measurement of diluted sample with pH meter. After 1 week of colonization, wt and delarc strains of EcN were introduced by oral gavage into mice in designated groups. After 2 weeks, mice were sacrificed with cecum samples extracted for analysis of species composition and cecal sample pH. (B) EcN relative abundance from fecal and cecal sample in wt and delarc conditions. EcN relative abundances are determined by multiplexed 16S rRNA sequencing on fecal samples and cecal samples. Datapoints indicate experimental data replicates. Lines indicate average over all biological replicates. EcN inoculation is marked by orange dashed line on Day 7. Asterisks represent statistical significance: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ according to an unpaired t-test. (C) Heatmap of the difference of species relative abundance (mean value of biological replicates, $n = 3$) in wt condition and delarc condition at various timepoints during the mouse experiment. Species relative abundances are determined by multiplexed 16S rRNA sequencing. Red corresponds to higher relative abundance in wt conditions and blue corresponds

to higher relative abundance in delarc conditions. EcN inoculation is marked on Day 7. Asterisks represent statistical significance: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ according to an unpaired t-test.

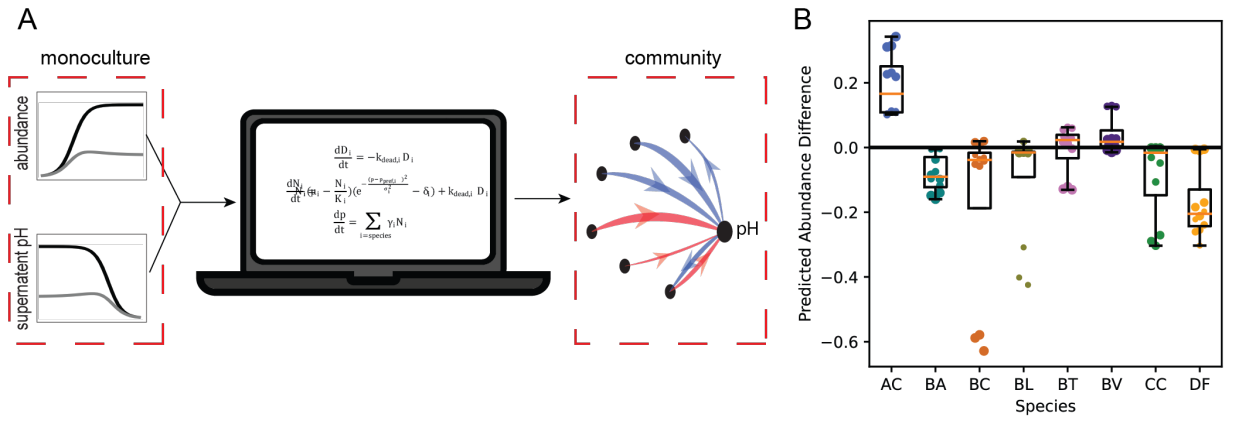
A dynamic computational model provides insight into community assembly due to pH-mediated inter-species interactions

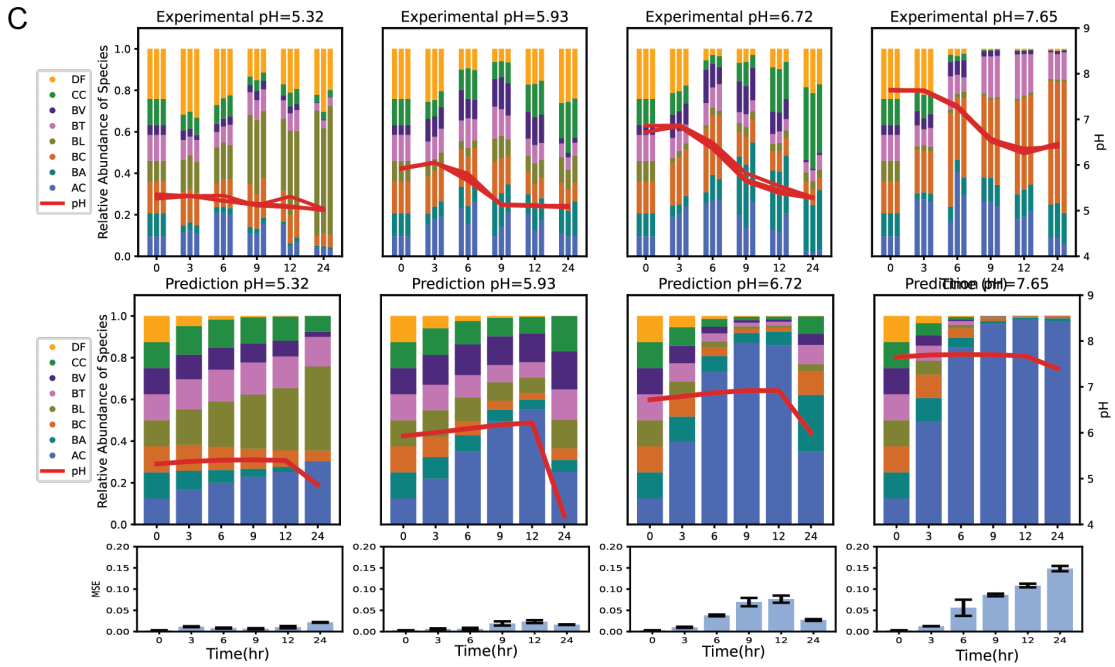
To quantitatively understand the impact of the *arc* operon on *in vitro* community assembly through pH, a dynamic computational model adapted from a previously identified microbe-pH model[30] was employed to unravel the extent of community assembly attributed to pH-mediated inter-species interactions. Monoculture data at various starting pH for species in comm11 and all EcN strains surveyed in this study were fitted to the model to obtain optimized parameters (Figure 4A, S6, S7). The monoculture fitted parameters yield strong linear correlation between predicted and experimental OD600 at all timepoints with a Pearson correlation coefficient of 0.99 and p-value of 0.000 (Figure S10A). Compared to the generic logistic model which assumes no interspecies interaction, our microbe-pH model predicted comm11 species assembly with on-par accuracy at lower pH and better accuracy at high pH (Figure S9A). Additionally, the pH model performed better than the logistic model at predicting comm11+EcN experiments in most initial pH conditions except pH=7, with lower predicted community Euclidean distance to experimental community (Figure S9B, S9C), suggesting the pH model can capture interspecies interactions not captured in logistic model to a certain degree.

The fitted parameters were then used to predict community assembly in comm11 and compared to experimental results. The predictions were able to capture the general acidifying trend of comm11 growth at all starting pH (Figure 4C). Most species present at a given pH are predicted by the model with difference in relative abundance, indicating the robustness of the assumption of pH as the only factor of inter-species interactions. Overall, most species are either consistently overpredicted (AC, BV) or underpredicted (DF, CC, BL, BC, BA) (Figure 4B). Since the computational model assumes only pH-mediated interaction among species, the discrepancy between prediction and experimental result point to the direction and magnitude of non-pH-mediated interactions for that species. Overpredicted species likely have negative non-pH-mediated interactions and underpredicted species potentially have positive non-pH-mediated interactions. Larger difference between prediction and experimental result indicates greater non-pH-mediated interactions. Specifically, DF in monoculture did not grow in any initial pH but grew in community at some of the same starting pH, producing the poor fit for DF in community and pointing to potential beneficial non-pH-mediated interactions with DF.

The fitted parameters from monoculture growth dynamics were then used to predict community assembly in comm11 when cocultured with various strains of EcN to understand the extent of pH-mediated interaction by the *arc* operon. Like the comm11 community structure prediction, species presence and absence and general pH trends are captured by the prediction, while differing in relative abundances, confirming pH-mediated interaction can explain most community assembly dynamics due to the *arc* operon (Figure D, E). Specifically, the model prediction can capture the overall higher final pH in pTet-arc (+) conditions versus delarc conditions for all initial pH conditions. The model prediction also picks up the EC dominance in higher initial pH conditions, with an EC percentage over 50% in both experimental data and model prediction. For all conditions, the model underpredicts the relative abundance of EC (Figure S10B), suggesting potential non-pH-mediated beneficial interaction for growth of EC in a community. The model prediction has better accuracy at lower starting pH than higher starting pH due to AC being the only species growing at higher pH in monoculture, suggesting the null hypothesis in model of dominating pH-mediated interaction fails at higher initial pH. The better prediction at lower pH can also attribute to the positive correlation between community structure Euclidean distance and pH difference in lower pH, specifically in the initial pH=6 condition (Figure S8), potentially conferring larger impact of the *arc* operon on community structure at lower pH. Overall, our modeling result verifies pH as the driving force behind community assembly due to *arc* operon *in vitro* while also

pointing to the direction and strength of potential non-pH-mediated interactions for all species in a synthetic microbial community.


A
monoculture
abundance
supernatent pH
community
pH
B
Predicted Abundance Difference
0.2
0.0
−0.2
−0.4
−0.6
AC
BA
BC
BL
BT
BV
CC
DF
Species


C
Experimental pH=5.32
Experimental pH=5.93
Experimental pH=6.72
Experimental pH=7.65
Prediction pH=5.32
Prediction pH=5.93
Prediction pH=6.72
Prediction pH=7.65
Relative Abundance of Species
pH
MSE
Time(hr)
DF
CC
BV
BT
BL
BC
BA
AC

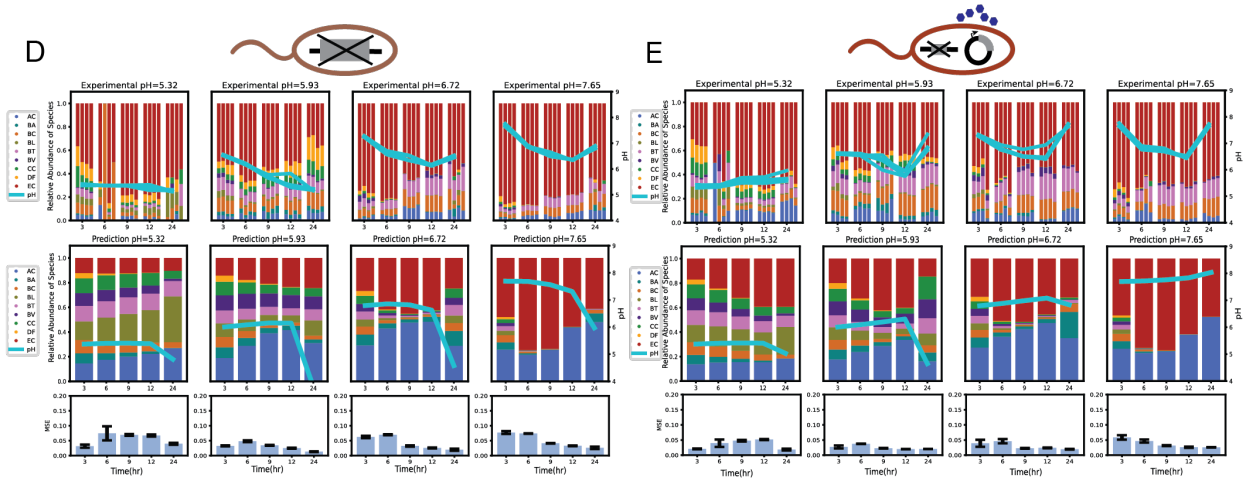

D
E
Experimental pH=5.32
Experimental pH=5.93
Experimental pH=6.72
Experimental pH=7.65
Prediction pH=5.32
Prediction pH=5.93
Prediction pH=6.72
Prediction pH=7.65
Relative Abundance of Species
pH
MSE
Time(hr)

Figure 4 | (A) Modeling framework for predicting community assembly. The modified microbe-pH interaction model represents community dynamics and environmental pH dynamics. Model parameters were optimized for monoculture data of comm11 species and strains of EcN by minimizing differences between predicted and experimental species abundance and supernatant pH. The community assembly was predicted using monoculture parameters to infer extent of pH-mediated interactions for species in a synthetic gut microbial community. (B) Box plot of the difference between predicted relative abundance and experimental relative abundance at 24-hour for comm11 coculture. Datapoints indicate experimental data replicates. Datapoint colors indicate species. Datapoint sizes indicate initial pH condition, with larger initial pH represented by larger datapoints. The y=0 line is marked by a black line to indicate perfect fitting. (C) Barplot of experimental and predicted composition of comm11 communities. Color indicates species identity. The red line indicates pH. The mean squared error between experimental and predicted species relative abundances are calculated for initial pH at each timepoint, with error bars indicating standard deviation of the mean squared errors over three biological replicates. (D) Barplot of experimental and predicted composition of comm11 coculture with delarc strain. Color indicates species identity. The red line indicates pH. The mean squared error between experimental and predicted species relative abundances are calculated for initial pH at each timepoint, with error bars indicating standard deviation of the mean squared errors over three biological replicates. (E) Barplot of experimental and predicted composition of comm11 coculture with pTet-arc(+) strain. Color indicates species identity. The red line indicates pH. The mean squared error between experimental and predicted species relative abundances are calculated for initial pH at each timepoint, with error bars indicating standard deviation of the mean squared errors over three biological replicates.

DISCUSSION

Our results showed that deletion of the *arc* operon from wildtype EcN introduces competition disadvantage *in vitro* while promoting colonization *in vivo* when in coculture with a synthetic gut microbial community. The comm11 and EcN coculture experiments also demonstrated the effectiveness in reshaping microbial community structure and metabolite production *in vitro* through the change of environmental pH introduced by tuning the expression of the *arc* operon. Co-colonization experiment in murine model further testified the capacity of the *arc* operon in affecting community assembly *in vivo*. Finally, we applied dynamic modeling to develop mechanistic understanding of community assembly *in vitro* and validated the expression of the *arc* operon affects in vitro community assembly predominantly through pH-mediated interspecies interactions.

Previous study on the arginine deiminase operon had similar results where the *arc* operon enhanced fitness of pathogenic EC species *in vitro* but was a burden in a mouse model of intestinal colonization[16]. While previous study concluded the arc operon was deleterious for EC, our result from co-colonization with a synthetic gut microbial community revealed unvaried EC CFU counts between wt and delarc coditions, suggesting presence of the *arc* operon yield the same fitness in EC while promoting growth of other species, presenting another interpretation for the apparent *arc*-operon-induced colonization disadvantage. The hypothesis that expression of the *arc* operon aid in growth of other species in community was also supported by the observation where the *arc* operon stabilize community against pH disturbance in our *in vitro* data. The promotion of other species in the synthetic gut community is potentially caused by the excess nitrogen resource because of the *arc* operon expression. The mammalian gut was previously shown to be limited in total nitrogen supply, and Bacteroidetes were shown to consume nitrogen in large intestine more readily than other commensals[32], potentially explaining the enrichment of BC in wildtype co-colonization conditions in our results. The interplay of pH modification and nitrogen supply further demonstrates the complexity of reshaping microbial communities *in vivo* compared to *in vitro*, calling for future studies involving species nitrogen uptake analysis when studying effect of pathways that produce nitrogen resource.

Our results also demonstrated the effectiveness of tuning the expression of the *arc* operon in both transient and sustained reshaping of microbial community structure *in vitro* and *in vivo*. In our *in vitro* experiment coculturing comm11 with different strains of EcN in various starting media pH, we observed enrichment of certain species in pTet-arc (+) coculture conditions (BC, BT, EC) while decline in some other species (BL, CC, DF) compared to delarc coculture conditions. Additionally, we observed butyrate production increase at lower initial pH correlating to AC relative abundance increases in coculture with *arc* operon expression, which is in accordance with previous observation that AC relative abundance correlates well with butyrate production in synthetic microbial communities[33] and provides theoretical basis for tuning beneficial metabolite production by *arc* operon expression. While we did not measure the butyrate level in mouse intestine, we predict a difference on butyrate production since we observed significant transient or sustained relative abundance changes in butyrate producers AC and CC. Additionally, in our *in vivo* experiment co-colonizing comm11 with wt and delarc strains of EcN, we observed enrichment of a different set of species in wildtype conditions (BC, CC) and decline in some others (AC, DF, EC). Specifically, both butyrate producing species (AC and CC) demonstrated opposite trend of relative abundance *in vitro* versus *in vivo*, suggesting a different mechanism governing *in vivo* colonization versus *in vitro* abundance, potentially related to nitrogen cross feeding as discuss above. The different community assembly structures with *arc* operon expression *in vitro*

and *in vivo* allude to the difficulty in scaling batch culture experiments to mouse models for microbial community studies.

The computational workflow trained on monoculture data also offered insight into how the *arc* operon influences community structure through pH-mediated interactions. The species consistently overpredicted by the null model (AC) point to negative non-pH-mediated interactions while species consistently underpredicted by the null model (BC, DF, EC) suggest positive non-pH-mediated interactions. BT and BV were found to be well predicted by the pH model across all conditions, suggesting the interspecies interactions for these species are predominantly pH mediated. Other species in comm11 (BA, BL, CC) had varying prediction direction across conditions, suggesting a change of pH-mediated interaction in different coculture structures. Specifically, DF had minimal growth in all starting pH surveyed, but showed considerable growth in community, which is in accordance with previous studies suggesting potential byproduct cross feedings between Firmicutes and Bacteroidetes and within Firmicutes species[34,35]. Additionaly, species are either overpredicted or underpredicted to the similar degree in the prediction for community structure of comm11 coculture with delarc condition or pTet-arc (+) condition (Figure S10C, S10D). The similar difference between prediction and experimental data for all species in both conditions suggest similar direction and magnitude of non-pH-mediated interactions for species in either condition. Therefore, most differences in community assembly between conditions with and without *arc* expression can be captured by our model assuming only pH-mediated interspecies interaction. Overall, our computational workflow identified pH as the main driving force behind community structure changes due to expression of the *arc* operon *in vitro*.

Since our comm11 and EcN coculture experiments in mouse models yield limited difference in community structures in wt and delarc conditions, the expression level of the *arc* operon in the wt strain can be tuned for more stable and higher expression of the pathway through integration of the inducible *arc* operon into the genome of EcN. Additionally, even though we focused on variation in media arginine concentration and initial media pH for this study, future studies could investigate effect of *arc* operon on colonization and community assembly under other contexts, including other metabolites in diet, inflamed gut environment, and other perturbations such as antibiotic supplementation. The predictions from our platform support the further investigation of non-pH-mediated interactions such as metabolite cross feedings within gut microbial communities for better mechanistic understandings of the extent of such interactions.

In sum, our study demonstrated modulation of host colonization, microbial community assembly, and metabolite production by tuning the expression of a unique metabolic pathway in a probiotic species both *in vitro* and *in vivo*. Additionally, we illustrated the potential of computational modeling in aiding mechanistic understanding of community assembly under specialized metabolic pathways. While we employed this framework to examine pH-mediated interspecies interactions, such strategy could also be applied to non-pH-mediated effects conferred by other specialized metabolic pathways, providing theoretical basis for model-guided control of *in vitro* community composition and metabolite production.

METHOD

Strain maintenance and culturing

All aerobic culturing was carried out in 37°C incubator with shaking. All anaerobic culturing was carried out in an anaerobic chamber with an atmosphere of 2 ± 0.5% H2, 15 ± 1% CO2 and balance N2. All prepared media and materials for anaerobic experiments were placed in the chamber at least overnight before use to equilibrate with the chamber atmosphere. The permanent stocks of each strain used in this paper were stored in 25% glycerol at −80 °C. Batches of single-use glycerol stocks were produced for each strain by first growing a culture from the permanent stock in YBHI media (Acumedia, Bacto, and Sigma-Aldrich) to stationary phase, mixing the culture in an equal volume of 50% glycerol, and aliquoting 400 µL into Matrix Tubes (ThermoFisher) for storage at −80 °C. Quality control for each batch of single-use glycerol stocks included Illumina sequencing of 16S rDNA isolated from pellets of the aliquoted mixture to verify the identity of the organism. For each experiment, precultures of each species were prepared by thawing a single-use glycerol stock (SUGS) and adding 100 µL of SUGS to 5 mL fresh YBHI media for stationary incubation at 37 °C for 16 hours. All experiments were performed in a chemically defined medium (DM38). This medium was developed by a previously studied medium in lab that enabled the growth of all our species in this experiment.

Monoculture and community culturing dynamic growth quantification

Each species' preculture was diluted to an OD600 of 0.01 (ThermoFisher Nanodrop, 1 mL in cuvette) in DM38 and aliquoted into three or four replicates of 1.5 mL each in a 96 Deep Well (96DW) plate and covered with a semi-permeable membrane (Diversified Biotech) and incubated at 37 °C without shaking. At each time point, samples were mixed and OD600 and pH were measured by the plate reader. Cell pellet was saved at each timepoint for species abundance quantification at -80°C. For pH measurement, phenol red solution was diluted to 0.005% weight per volume in MilliQ water. Bacterial supernatant (20 µL) was added to 180 µL of phenol red solution, and absorbance was measured at 560 nm (Tecan Spark Plate Reader). A standard curve was produced by fitting the Henderson–Hasselbach equation to fresh media with a pH ranging between 3 and 11 measured using a standard electro-chemical pH probe (Mettler-Toledo). We used the following equation to map the pH values to the absorbance measurements.

$$\mathrm{pH} = a + b * \ln A_{560}$$

The parameters a and b were determined using a linear regression between pH and the log term for the standards in the linear range of absorbance with A representing the absorbance of each condition.

Genomic DNA extraction from cell pellets

Genomic DNA was extracted from cell pellets using a modified version of the Qiagen DNeasy Blood and Tissue Kit protocol. First, pellets in 96DW plates were removed from −80 °C and thawed in a room temperature water bath. Each pellet was resuspended in 180 µL of enzymatic lysis buffer (20 mM Tris–HCl (Invitrogen), 2 mM sodium EDTA (Sigma-Aldrich), 1.2% Triton X-100 (Sigma-Aldrich), 20 mg/mL lysozyme from chicken egg white (Sigma-Aldrich)). Plates were then covered with a foil seal and incubated at 37 °C for 30 min with orbital shaking at 600 RPM. Then, 25 µL of 20 mg mL−1 Proteinase K (VWR) and 200 µL of Buffer AL (QIAGEN) were added to each sample before mixing with a pipette. Plates were then covered by a foil seal and incubated at 56 °C for 30 min with orbital shaking at 600 RPM. Next, 200 µL of 100% ethanol (Koptec) was added to each sample before mixing and samples were transferred to a nucleic acid binding (NAB) plate (Pall) on a vacuum manifold with a 96DW collection plate. Each well in the NAB plate was then washed once with 500 µL buffer AW1 (QIAGEN) and once with 500 µL of buffer AW2 (QIAGEN). A vacuum was applied to the Pall NAB plate for an additional 10 min to remove any

excess ethanol. Samples were then eluted into a clean 96DW plate from each well using 110 µL of buffer AE (QIAGEN) preheated to 56 °C. Genomic DNA samples were stored at −20 °C until further processing.

Sequencing library preparation

Genomic DNA concentrations were measured using a SYBR Green fluorescence assay and then normalized to a concentration of 1 ng µL−1 by diluting in molecular grade water using a Tecan Evo Liquid Handling Robot. First, genomic DNA samples were removed from −20 °C and thawed in a room temperature water bath. Then, 1 µL of each sample was combined with 95 µL of SYBR Green (Invitrogen) diluted by a factor of 100 in TE buffer (Integrated DNA Technologies) in a black 384-well microplate. This process was repeated with two replicates of each DNA standard with concentrations of 0, 0.5, 1, 2, 4, and 6 ng µL−1. Each sample was then measured for fluorescence with an excitation/emission of 485/535 nm using a Tecan Spark plate reader. Concentrations of each sample were calculated using the standard curve and a custom Python script was used to compute the dilution factors and write a worklist for the Tecan Evo Liquid Handling Robot to normalize each sample to 1 ng µL−1 in molecular grade water. Samples with DNA concentration <1 ng µL−1 were not diluted. Diluted genomic DNA samples were stored at −20 °C until further processing.

Amplicon libraries were generated from diluted genomic DNA samples by PCR amplification of the V3–V4 of the 16S rRNA gene using custom dual-indexed primers (Supplementary Data 5) for multiplexed next-generation amplicon sequencing on Illumina platforms (Method adapted from Venturelli et al. Mol. Syst. Biol., 2018). Primers were arrayed in skirted 96-well PCR plates (VWR) using an acoustic liquid handling robot (Labcyte Echo 550) such that each well received a different combination of one forward and one reverse primer (0.1 µL of each). After liquid evaporated, dry primers were stored at −20 °C. Primers were resuspended in 15 µL PCR master mix (0.2 µL Phusion High Fidelity DNA Polymerase (Thermo Scientific), 0.4 µL 10 mM dNTP solution (New England Biolabs), 4 µL 5× phusion HF buffer (Thermo Scientific), 4 µL 5 M Betaine (Sigma-Aldrich), 6.4 µL Water) and 5 µL of normalized genomic DNA to give a final concentration of 0.05 µM of each primer. Primer plates were sealed with Microplate B seals (Bio-Rad) and PCR was performed using a Bio-Rad C1000 Thermal Cycler with the following program: initial denaturation at 98 °C (30 s); 25 cycles of denaturation at 98 °C (10 s), annealing at 60 °C (30 s), extension at 72 °C (60 s); and final extension at 72 °C (10 min). 2 µL of PCR products from each well were pooled and purified using the DNA Clean & Concentrator (Zymo) and eluted in water. The resulting libraries were sequenced on an Illumina MiSeq using a MiSeq Reagent Kit v3 (600-cycle) to generate 2 × 300 paired-end reads.

Bioinformatic analysis for quantification of species abundance

Sequencing data were demultiplexed using Basespace Sequencing Hub's FastQ Generation program. Custom python scripts were used for further data processing (method adapted from Venturelli et al. Mol. Syst. Biol., 2018)16. Paired end reads were merged using PEAR (v0.9.10)83 after which reads without forward and reverse annealing regions were filtered out. A reference database of the V3–V5 16S rRNA gene sequences was created using consensus sequences from next-generation sequencing data or Sanger sequencing data of monospecies cultures. Sequences were mapped to the reference database using the mothur (v1.40.5)84 command classify.seqs (Wang method with a bootstrap cutoff value of 60). Relative abundance was calculated as the read count mapped to each species divided by the total number of reads for each condition. Absolute abundance of each species was calculated by multiplying the relative abundance by the OD600 measurement for each sample. Samples were excluded from further analysis if >1% of the reads were assigned to a species not expected to be in the community

(indicating contamination) or if they had <1000 total reads and OD600 > 0.1 (indicating that there were insufficient reads for analysis and this was not due to lack of community growth).

HPLC quantification of organic acids

Supernatant samples were thawed in a room temperature water bath before addition of 2 µL of H_2SO_4 to precipitate any components that might be incompatible with the running buffer. The samples were then centrifuged at 3500 × rpm for 10 min and then 150 µL of each sample was filtered through a 0.2 µm filter using a vacuum manifold before transferring 70 µL of each sample to an HPLC vial. HPLC analysis was performed using either a Shimadzu HPLC system equipped with a SPD-20AV UV detector (210 nm). Compounds were separated on a 250 × 4.6 mm Rezex© ROA-Organic acid LC column (Phenomenex Torrance, CA) run with a flow rate of 0.2 mL min^{-1} and at a column temperature of 50 °C. The samples were held at 4 °C prior to injection. Separation was isocratic with a mobile phase of HPLC grade water acidified with 0.015 N H_2SO_4 (415 µL L−1). At least two standard sets were run along with each sample set. Standards were 100, 20, and 4 mM concentrations of butyrate, succinate, lactate, and acetate, respectively. The injection volume for both sample and standard was 25 µL. The resultant data was analyzed using Shimadzu LabSolutions software package with peaks manually reintegrated if necessary.

Gnotobiotic mouse experiments

All germ-free mouse experiments were performed following protocols approved by the University of Wisconsin-Madison Animal Care and Use Committee. Two diets were used in this experiment: regular diet with lower arginine concentration (Chow diet, Purina, LabDiet 5021) and high arginine diet (Envigo, TD.210715). All strains were grown at 37 ºC anaerobically in YBHI media (Acumedia, Bacto, and Sigma-Aldrich) to stationary phase. All strains for a oral gavage were mixed in equal proportions based on OD600 and transferred on ice prior to oral gavage. We used 8-week old C57BL/6 gnotobiotic male mice (wild-type) fed the specific diets 3 days prior to oral gavage (Figure 2E, 3E). The mice were housed in biocontainment cages (Allentown Inc.) for the duration of the experiment. Mice were maintained on the same experimental diets with autoclaved water. Groups of mice (4 mice) fed a given condition of consortium were co-housed in a single cage. Fecal samples were collected every 2-3 days after oral gavage for NGS sequencing and pH measurement. At the end of the experiment, mice were euthanized, and the cecal contents were collected for NGS sequencing and pH measurement.

Genomic DNA extraction from fecal and cecal samples

The DNA extraction for fecal and cecal samples was performed as described previously with some modifications[36]. Fecal samples (~50 mg) were transferred into solvent-resistant screw-cap tubes (Sarstedt Inc) with 500 µL 0.1 mm zirconia/silica beads (BioSpec Products) and one 3.2 mm stainless steel bead (BioSpec Products). The samples were resuspended in 500 µL of Buffer A (200 mM NaCl (DOT Scientific), 20 mM EDTA (Sigma) and 200 mM Tris·HCl pH 8.0 (Research Products International)), 210 µL 20% SDS (Alfa Aesar) and 500 µL phenol/chloroform/isoamyl alcohol (Invitrogen). Cells were lysed by mechanical disruption with a bead-beater (BioSpec Products) for 3 min twice to prevent overheating. Next, cells were centrifuged for 5 min at 8,000 x g at 4°C, and the supernatant was transferred to a Eppendof tube. We added 60 µL 3M sodium acetate (Sigma) and 600 µL isopropanol (LabChem) to the supernatant and incubated on ice for 1 hr. Next, samples were centrifuged for 20 min at 18,000 x g at 4°C. The harvested DNA pellets were washed once with 500 µL of 100% ethanol (Koptec). The remaining trace ethanol was removed by air drying the samples. Finally, the DNA pellets were then resuspended into 200 µL of AE buffer (Qiagen). The crude DNA extracts were purified by a Zymo DNA Clean & Concentrator™-5 kit (Zymo Research) prior to PCR amplification and then NGS sequencing was performed.

Gene expression measurements of the *arc* operon in germ-free mice

The RNA extraction of the contents of each cecum was performed as described previously with some modifications[37]. Specifically, cecal samples (100~150 mg) were transferred into solvent-resistant screw-cap tubes (Sarstedt Inc) with 500 µL of acid washed beads (212-300 µm, Sigma). The samples were resuspended in 500 µL of Acid-Phenol:Chloroform (with IAA, 125:24:1, pH 4.5, Thermo Fisher Scientific), 500 µL of Buffer B (200mM NaCl (DOT Scientific), 20 mM EDTA (Sigma), pH 8.0), 210 µL 20% SDS (Alfa Aesar). Cells were lysed by mechanical disruption with a bead-beater (BioSpec Products) for 2 min twice to prevent overheating. Next, cells were centrifuged for 10 min at 8,000 x g at 4°C, and the supernatant was transferred into an Eppendof tube. We added 60 µL 3M sodium acetate (pH 5.5, Sigma) and 600 µL isopropanol (LabChem) to the supernatant and incubated at −80°C for 6 min. Next, samples were centrifuged for 15 min at 18,200 x g at 4°C. The harvested RNA pellets were washed once with 900 µL of 100% ethanol (Koptec). Finally, the RNA pellets were resuspended into 100 µL of RNase-free water. The crude RNA extracts were purified by a RNeasy Mini kit (Qiagen) with a DNase I (Qiagen) treatment step to eliminate DNA contamination in the sample. To remove any remaining DNA, the RNA products were treated with Baseline-ZERO DNase I (Epicentre) and then purified again using the RNeasy Mini kit.

We performed cDNA synthesis with 0.5-1 µg of total purified RNA using the iScript Select cDNA Synthesis Kit (Bio-Rad Laboratories). We performed quantitative reverse transcription PCR (qRT-PCR) on the Bio-Rad CFX connect Real-Time PCR instrument with SYBR™ Green PCR Master Mix (Thermo Fisher Scientific). We computed the fold changes of the target genes by normalizing to the two reference genes 16S rRNA gene and cysG[38] using the geometric mean[39]. We used three technical replicates per gene for qRT-PCR.

pH model and training

The pH model shown was adapted from the growth model with pH mediation developed by the Gore lab[30].

$$\frac{\mathrm{d}D_i}{\mathrm{d}t} = -k_{dead,i} D_i$$

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = N_i \left(\mu_i - \frac{N_i}{K_i}\right)\left(e^{-\frac{(p-p_{pref,i})^2}{\sigma_i^2}} - \delta_i\right) + k_{dead,i} D_i$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \sum_{i=species} \gamma_i N_i$$

Bacterial species density N_i for species i follows a logistic growth defined by maximum growth rate μ_i and carrying capacity K_i. A Gaussian pH modification is applied to the growth rate which is defined around preferred pH $p_{pref,i}$ with standard deviation σ_i and death rate δ_i. The species growth rate reaches maximum at a preferred pH and decreases outside a growth-promoting pH range. The environmental pH p is affected by the summation of each species by a pH-change-factor γ_i, where the factor can take values of both positive and negative depending on the species being an acidifier or basifier. To capture the initial lag phase of batch culture, D_i was introduced as the species density unavailable for replication. The unavailable density decreases with an exponential curve defined by $k_{dead,i}$, granting cells to replicate at the same rate in N_i.

REFERENCES

1. Schirmer, M., Smeekens, S.P., Vlamakis, H., Jaeger, M., Oosting, M., Franzosa, E.A., Jansen, T., Jacobs, L., Bonder, M.J., Kurilshikov, A., et al. (2016). Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. Cell *167*, 1125-1136.e8. 10.1016/J.CELL.2016.10.020.
2. Jones, B. v., Begley, M., Hill, C., Gahan, C.G.M., and Marchesi, J.R. (2008). Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. Proc Natl Acad Sci U S A *105*, 13580–13585. 10.1073/PNAS.0804437105/SUPPL_FILE/0804437105SI.PDF.
3. Feng, J., Qian, Y., Zhou, Z., Ertmer, S., Vivas, E.I., Lan, F., Hamilton, J.J., Rey, F.E., Anantharaman, K., and Venturelli, O.S. (2022). Polysaccharide utilization loci in Bacteroides determine population fitness and community-level interactions. Cell Host Microbe *30*, 200-215.e12. 10.1016/J.CHOM.2021.12.006/ATTACHMENT/6689EF7A-6A3E-4DFB-A8A5-3910C3E8AE19/MMC6.XLSX.
4. Wolfe, A.J. (2015). Glycolysis for Microbiome Generation. Microbiol Spectr *3*. 10.1128/MICROBIOLSPEC.MBP-0014-2014/ASSET/BE7B794E-2D1A-4E87-95C8-80AC07A129EF/ASSETS/GRAPHIC/MBP-0014-2014-FIG7.GIF.
5. Lucas, L.N., Barrett, K., Kerby, R.L., Zhang, Q., Cattaneo, L.E., Stevenson, D., Rey, F.E., and Amador-Noguez, D. (2021). Dominant Bacterial Phyla from the Human Gut Show Widespread Ability To Transform and Conjugate Bile Acids. mSystems *6*. 10.1128/MSYSTEMS.00805-21/SUPPL_FILE/MSYSTEMS.00805-21-ST006.XLSX.
6. Rowland, I., Gibson, G., Heinken, A., Scott, K., Swann, J., Thiele, I., and Tuohy, K. (2017). Gut microbiota functions: metabolism of nutrients and other food components. European Journal of Nutrition 2017 57:1 *57*, 1–24. 10.1007/S00394-017-1445-8.
7. Visconti, A., le Roy, C.I., Rosa, F., Rossi, N., Martin, T.C., Mohney, R.P., Li, W., de Rinaldis, E., Bell, J.T., Venter, J.C., et al. (2019). Interplay between the human gut microbiome and host metabolism. Nature Communications 2019 10:1 *10*, 1–10. 10.1038/s41467-019-12476-z.
8. Shepherd, E.S., Deloache, W.C., Pruss, K.M., Whitaker, W.R., and Sonnenburg, J.L. (2018). An exclusive metabolic niche enables strain engraftment in the gut microbiota. Nature *557*, 434–438. 10.1038/s41586-018-0092-4.
9. Sheth, R.U., Cabral, V., Chen, S.P., Wang, H.H., and Wang, H.H. (2016). Manipulating Bacterial Communities by in situ Microbiome Engineering. Trends in Genetics *32*. 10.1016/j.tig.2016.01.005.
10. Furusawa, Y., Obata, Y., Fukuda, S., Endo, T.A., Nakato, G., Takahashi, D., Nakanishi, Y., Uetake, C., Kato, K., Kato, T., et al. (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature 2013 504:7480 *504*, 446–450. 10.1038/nature12721.
11. Rivera-Chávez, F., Zhang, L.F., Faber, F., Lopez, C.A., Byndloss, M.X., Olsan, E.E., Xu, G., Velazquez, E.M., Lebrilla, C.B., Winter, S.E., et al. (2016). Depletion of Butyrate-Producing Clostridia from the Gut Microbiota Drives an Aerobic Luminal Expansion of Salmonella. Cell Host Microbe *19*, 443–454. 10.1016/J.CHOM.2016.03.004.
12. Fung, K.Y.C., Cosgrove, L., Lockett, T., Head, R., and Topping, D.L. (2012). A review of the potential mechanisms for the lowering of colorectal oncogenesis by butyrate. British Journal of Nutrition *108*, 820–831. 10.1017/S0007114512001948.
13. Gilijamse, P.W., Hartstra, A. v., Levin, E., Wortelboer, K., Serlie, M.J., Ackermans, M.T., Herrema, H., Nederveen, A.J., Imangaliyev, S., Aalvink, S., et al. (2020). Treatment with Anaerobutyricum soehngenii: a pilot study of safety and dose–response effects on glucose metabolism in human subjects with metabolic syndrome. npj Biofilms and Microbiomes 2020 6:1 *6*, 1–10. 10.1038/s41522-020-0127-0.

14. Geirnaert, A., Calatayud, M., Grootaert, C., Laukens, D., Devriese, S., Smagghe, G., de Vos, M., Boon, N., and van de Wiele, T. (2017). Butyrate-producing bacteria supplemented in vitro to Crohn's disease patient microbiota increased butyrate production and enhanced intestinal epithelial barrier integrity. Sci Rep *7*. 10.1038/s41598-017-11734-8.
15. Zhang, H., Liu, Y., Nie, X., Liu, L., Hua, Q., Zhao, G.P., and Yang, C. (2018). The cyanobacterial ornithine–ammonia cycle involves an arginine dihydrolase. Nature Chemical Biology 2018 14:6 *14*, 575–581. 10.1038/s41589-018-0038-z.
16. Billard-Pomares, T., Clermont, O., Castellanos, M., Magdoud, F., Royer, G., Condamine, B., Fouteau, S., Barbe, V., Roche, D., Cruveiller, S., et al. (2019). The arginine deiminase operon is responsible for a fitness trade-off in extended-spectrum-β-lactamase-producing strains of Escherichia coli. Antimicrob Agents Chemother *63*. 10.1128/AAC.00635-19/SUPPL_FILE/AAC.00635-19-S0001.PDF.
17. Marquis, R.E., Bender, G.R., Murray, D.R., and Wong, A. (1987). Arginine deiminase system and bacterial adaptation to acid environments. Appl Environ Microbiol *53*, 198–200. 10.1128/AEM.53.1.198-200.1987.
18. Wauven, C. v., Pierard, A., Kley-Raymann, M., and Haas, D. (1984). Pseudomonas aeruginosa mutants affected in anaerobic growth on arginine: evidence for a four-gene cluster encoding the arginine deiminase pathway. J Bacteriol *160*, 918–934. 10.1128/JB.160.3.928-934.1984.
19. Reese, A.T., Pereira, F.C., Schintlmeister, A., Berry, D., Wagner, M., Hale, L.P., Wu, A., Jiang, S., Durand, H.K., Zhou, X., et al. (2018). Microbial nitrogen limitation in the mammalian large intestine. Nature Microbiology 2018 3:12 *3*, 1441–1450. 10.1038/s41564-018-0267-7.
20. Jenni Firrman, LinShu Liu, Karley Mahalak, Ceylan Tanes, Kyle Bittinger, Vincent Tu, Jamshed Bobokalonov, Lisa Mattei, Huanjia Zhang, P.V. den A. (2016). The impact of environmental pH on the gut microbiota community structure and short chain fatty acid production. FEMS Microbiol Ecol, 1–11. 10.1093/FEMSEC/FIAC038.
21. Wang, S.P., Rubio, L.A., Duncan, S.H., Donachie, G.E., Holtrop, G., Lo, G., Farquharson, F.M., Wagner, J., Parkhill, J., Louis, P., et al. (2020). Pivotal Roles for pH, Lactate, and Lactate-Utilizing Bacteria in the Stability of a Human Colonic Microbial Ecosystem. mSystems *5*. 10.1128/MSYSTEMS.00645-20/SUPPL_FILE/MSYSTEMS.00645-20-SF007.TIF.
22. Nugent, S.G., Kumar, D., Rampton, D.S., and Evans, D.F. (2001). Intestinal luminal pH in inflammatory bowel disease: possible determinants and implications for therapy with aminosalicylates and other drugs. Gut *48*, 571–577. 10.1136/GUT.48.4.571.
23. Sofi, M.H., Gudi, R., Karumuthil-Melethil, S., Perez, N., Johnson, B.M., and Vasu, C. (2014). PH of drinking water influences the composition of gut microbiome and type 1 diabetes incidence. Diabetes *63*, 632–644. 10.2337/db13-0981.
24. Leggieri, P.A., Liu, Y., Hayes, M., Connors, B., Seppälä, S., O'Malley, M.A., and Venturelli, O.S. (2021). Integrating Systems and Synthetic Biology to Understand and Engineer Microbiomes. https://doi.org/10.1146/annurev-bioeng-082120-022836 *23*, 169–201. 10.1146/ANNUREV-BIOENG-082120-022836.
25. Qian, Y., Lan, F., and Venturelli, O.S. (2021). Towards a deeper understanding of microbial communities: integrating experimental data with dynamic models. Curr Opin Microbiol *62*, 84–92. 10.1016/J.MIB.2021.05.003.
26. Venturelli, O.S., Carr, A. v, Fisher, G., Hsu, R.H., Lau, R., Bowen, B.P., Hromada, S., Northen, T., and Arkin, A.P. (2018). Deciphering microbial interactions in synthetic human gut microbiome communities. Mol Syst Biol *14*. 10.15252/msb.20178157.
27. Clark, R.L., Connors, B.M., Stevenson, D.M., Hromada, S.E., Hamilton, J.J., Amador-Noguez, D., and Venturelli, O.S. (2021). Design of synthetic human gut microbiome

assembly and butyrate production. Nature Communications 2021 12:1 *12*, 1–16. 10.1038/s41467-021-22938-y.
28. Hromada, S., Qian, Y., Jacobson, T.B., Clark, R.L., Watson, L., Safdar, N., Amador-Noguez, D., and Venturelli, O.S. (2021). Negative interactions determine Clostridioides difficile growth in synthetic human gut communities. Mol Syst Biol *17*, e10355. 10.15252/MSB.202110355.
29. Ostrem Loss, E., Thompson, J., Cheung, P.L.K., Qian, Y., and Venturelli, O.S. (2023). Carbohydrate complexity limits microbial growth and reduces the sensitivity of human gut communities to perturbations. Nature Ecology & Evolution 2023 7:1 *7*, 127–142. 10.1038/s41559-022-01930-9.
30. Ratzke, C., and Gore, J. (2018). Modifying and reacting to the environmental pH can drive bacterial interactions. PLoS Biol *16*, e2004248. 10.1371/JOURNAL.PBIO.2004248.
31. Altenhoefer, A., Oswald, S., Sonnenborn, U., Enders, C., Schulze, J., Hacker, J., and Oelschlaeger, T.A. (2004). The probiotic Escherichia coli strain Nissle 1917 interferes with invasion of human intestinal epithelial cells by different enteroinvasive bacterial pathogens. FEMS Immunol Med Microbiol *40*, 223–229. 10.1016/S0928-8244(03)00368-7.
32. Reese, A.T., Pereira, F.C., Schintlmeister, A., Berry, D., Wagner, M., Hale, L.P., Wu, A., Jiang, S., Durand, H.K., Zhou, X., et al. (2018). Microbial nitrogen limitation in the mammalian large intestine. Nature Microbiology 2018 3:12 *3*, 1441–1450. 10.1038/s41564-018-0267-7.
33. Clark, R.L., Connors, B.M., Stevenson, D.M., Hromada, S.E., Hamilton, J.J., Amador-Noguez, D., and Venturelli, O.S. (2021). Design of synthetic human gut microbiome assembly and butyrate production. Nature Communications 2021 12:1 *12*, 1–16. 10.1038/s41467-021-22938-y.
34. Soto-Martin, E.C., Warnke, I., Farquharson, F.M., Christodoulou, M., Horgan, G., Derrien, M., Faurie, J.M., Flint, H.J., Duncan, S.H., and Louis, P. (2020). Vitamin biosynthesis by human gut butyrate-producing bacteria and cross-feeding in synthetic microbial communities. mBio *11*, 1–18. 10.1128/MBIO.00886-20/SUPPL_FILE/MBIO.00886-20-ST006.PDF.
35. Henson, M., and Phalak, P. (2017). Byproduct Cross Feeding and Community Stability in an In Silico Biofilm Model of the Gut Microbiome. Processes *5*, 13. 10.3390/pr5010013.
36. Goodman, A.L., Kallstrom, G., Faith, J.J., Reyes, A., Moore, A., Dantas, G., and Gordon, J.I. (2011). Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. Proc Natl Acad Sci U S A *108*, 6252–6257. 10.1073/PNAS.1102938108/SUPPL_FILE/ST01.XLS.
37. Rey, F.E., Faith, J.J., Bain, J., Muehlbauer, M.J., Stevens, R.D., Newgard, C.B., and Gordon, J.I. (2010). Dissecting the in Vivo Metabolic Potential of Two Human Gut Acetogens. J Biol Chem *285*, 22082. 10.1074/JBC.M110.117713.
38. Zhou, K., Zhou, L., Lim, Q., Zou, R., Stephanopoulos, G., and Too, H.P. (2011). Novel reference genes for quantifying transcriptional responses of Escherichia coli to protein overexpression by quantitative PCR. BMC Mol Biol *12*, 1–9. 10.1186/1471-2199-12-18/FIGURES/5.
39. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol *3*, 1–12. 10.1186/GB-2002-3-7-RESEARCH0034/COMMENTS.