



# Formation, characterization and modeling of emergent synthetic microbial communities <sup>☆</sup>



Jia Wang <sup>a</sup>, Dana L. Carper <sup>a</sup>, Leah H. Burdick <sup>a</sup>, Him K. Shrestha <sup>a,b</sup>, Manasa R. Appidi <sup>a,b</sup>, Paul E. Abraham <sup>a</sup>, Collin M. Timm <sup>a,c</sup>, Robert L. Hettich <sup>a</sup>, Dale A. Pelletier <sup>a,\*</sup>, Mitchel J. Doktycz <sup>a,\*</sup>

<sup>a</sup> Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

<sup>b</sup> Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, TN, USA

## ARTICLE INFO

### Article history:

Received 14 December 2020

Received in revised form 22 March 2021

Accepted 25 March 2021

Available online 9 April 2021

### Keywords:

Microbial community  
Rhizosphere bacteria  
Genome-scale model  
Flux balance analysis  
Metabolic interaction  
Metaproteomics

## ABSTRACT

Microbial communities colonize plant tissues and contribute to host function. How these communities form and how individual members contribute to shaping the microbial community are not well understood. Synthetic microbial communities, where defined individual isolates are combined, can serve as valuable model systems for uncovering the organizational principles of communities. Using genome-defined organisms, systematic analysis by computationally-based network reconstruction can lead to mechanistic insights and the metabolic interactions between species. In this study, 10 bacterial strains isolated from the *Populus deltoides* rhizosphere were combined and passaged in two different media environments to form stable microbial communities. The membership and relative abundances of the strains stabilized after around 5 growth cycles and resulted in just a few dominant strains that depended on the medium. To unravel the underlying metabolic interactions, flux balance analysis was used to model microbial growth and identify potential metabolic exchanges involved in shaping the microbial communities. These analyses were complemented by growth curves of the individual isolates, pairwise interaction screens, and metaproteomics of the community. A fast growth rate is identified as one factor that can provide an advantage for maintaining presence in the community. Final community selection can also depend on selective antagonistic relationships and metabolic exchanges. Revealing the mechanisms of interaction among plant-associated microorganisms provides insights into strategies for engineering microbial communities that can potentially increase plant growth and disease resistance. Further, deciphering the membership and metabolic potentials of a bacterial community will enable the design of synthetic communities with desired biological functions.

© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

<sup>☆</sup> This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

\* Corresponding authors.

E-mail addresses: [pelletierda@ornl.gov](mailto:pelletierda@ornl.gov) (D.A. Pelletier), [doktyczmj@ornl.gov](mailto:doktyczmj@ornl.gov) (M.J. Doktycz).

<sup>c</sup> Current address: Research and Exploratory Development Department, Johns Hopkins University Applied Physics Laboratory, Laurel, MD, USA.

## 1. Introduction

Bacterial communities exert significant influence over a wide range of biological processes, such as human disease, plant interactions, biogeochemical cycles, and food fermentation [1–3]. An important challenge exists in understanding the underlying mechanisms that contribute to how bacteria shape their community and how the resulting structure depends on distinct environmental niches [2,4,5]. The interactions between species are dynamic and community membership depends on possessing the metabolic capabilities needed to survive in a particular environment [6]. Determining these trophic exchanges and interdependent metabolic processes is difficult in natural microbial communities comprised of hundreds of members. For example, phylogenetic marker and metagenomics analyses have revealed the extreme diversity of rhizosphere bacterial communities and the complex

interplay among them [7–11]. It has been established that the composition and activity of root bacterial communities is spatially and temporally dynamic and can be influenced by both abiotic (e.g., soil nutrients, O<sub>2</sub>, pH, etc.) and biotic (e.g., host and microbe-microbe) factors [12–15]. This complexity prevents tracking of metabolic fluxes from specific donor to acceptor strains or identifying competitive and cooperative relationships that leads to community structure [16,17].

Simplified synthetic microbial communities are being considered as comprehensible systems for uncovering an in-depth view of community assembly principles. These systems are able to circumvent the complexity of natural ecosystems and allow the capture of community behaviors [17–21]. One approach is to reduce the complexity of natural communities by selection of microbial consortia under laboratory conditions from environmental samples [22]. This top-down approach can provide an overall co-occurrence correlation network but does not assess metabolic interactions in detail as individual genome and metabolic profiles are lacking [23]. A second approach is to construct synthetic bacterial communities from the bottom-up [24]. In the bottom-up approach, individual bacterial isolates are combined to give rise to a more complex microbial system where the original strains serve as sub-systems in an emergent community [25–27]. These easily manipulated bottom-up assemblies contribute to a promising approach for understanding interactions in natural communities [26,28,29]. Definition and characterization of each individual strain facilitates the study of potential synergistic effects in the synthetic community [25,30]. The metabolites driving interspecies interactions can be determined and modeling of the metabolite exchange is possible [16,19,29]. This bottom-up approach can be used to experimentally select and investigate stable microbial community assembly [22].

The interplay among bacterial members in a consortium can be reconstructed using community-wide genome-scale metabolic models [31]. Recent applications of computational biology and genome-scale modeling approaches to the analysis of bottom-up assembled communities is providing mechanistic insights into the dynamic interactions occurring in defined bacterial communities [17,23,25]. For example, modeling studies have been applied for understanding biodegradation and bioproductivity [32]. In these models, it is often assumed that species interact in a pairwise manner [33,34]. Two-species metabolic models assess cross-feeding networks and usually capture the positive interactions between the microorganisms [34–36]. Currently, metabolic exchanges among greater numbers of microbes are being modeled and found important for shaping community distribution [22,37]. Modeling these higher-order interactions will be helpful for addressing questions regarding how and why a stable bacterial community forms.

In this study, we describe the formation, characterization, and modeling of synthetic bacterial communities assembled from a highly characterized, phylogenetically diverse set of selected isolates in different media environments. The aim of these efforts is to define an approach to discovering simple, reproducible microbial communities, without predefined relationships, for detailed experimental studies that allow molecular and cellular level investigations into community structure. Using this discovery-based approach, we hypothesize that different community structures will result and depend on the media environment and the cooperative and competitive characteristics of the emergent community members. Ten bacterial strains (Table 1), isolated from *Populus deltoides* rhizosphere and with defined genome sequences, were co-cultured in either complex or minimal glucose media and serially transferred until a stable community structure formed. The resulting, reproducible system allows for understanding community assembly processes and investigation of causative molecular and cellular

level events. To this end, a combination of marker gene profiling and metaproteomics characterization was carried out for tracking community structure and for gaining mechanistic insights into interactions between isolates (Fig. 1). These data are complemented by growth curve analyses and pairwise interaction screens. Different stable communities assemble in these environments and the higher-order interactions among community members are investigated. To unravel potential metabolic interactions among the surviving community members, genome-scale, community-level metabolic models were constructed for predicting potential metabolic processes involved in shaping the bacterial communities. The approach of discovering new microbial community structures under laboratory-defined conditions will facilitate understanding of the formation and dynamics of natural communities and the rational design of synthetic consortia with desired biological functions.

## 2. Materials and methods

### 2.1. Community construction

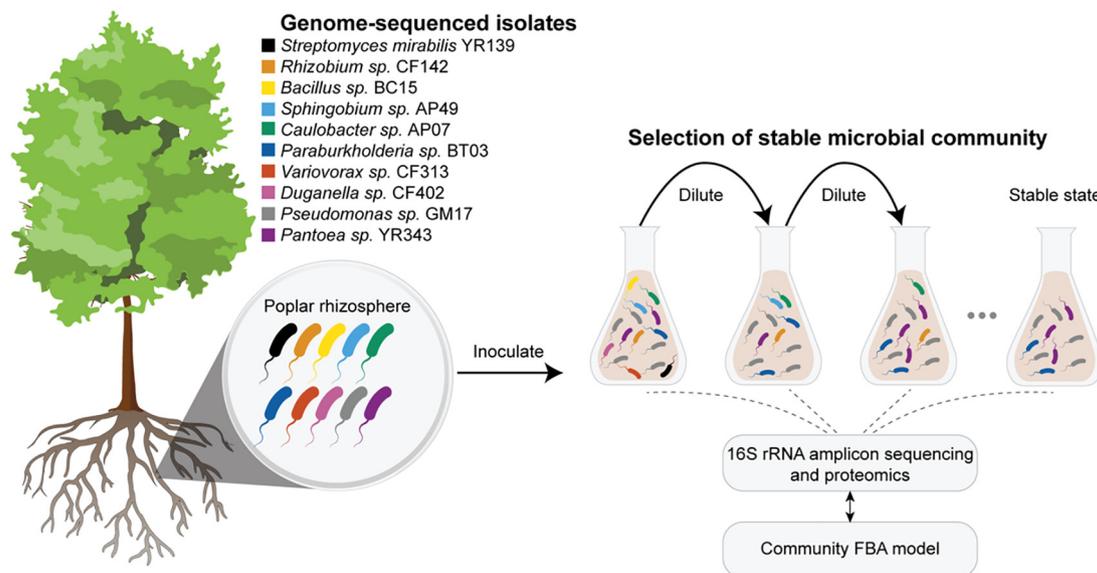
All 10 wild type bacterial strains (Table 1) were previously isolated from the rhizosphere of *Populus deltoides* [7,38–40]. Two media were utilized in the experiments. A complex medium and a minimal, chemically defined medium were used in order to ascertain the influence of the chemical environment on microbial community selection. R2A was selected as the complex medium as this medium was used in the original environmental isolations of the strains. The medium contains essential amino acids and both simple and complex carbon sources that can potentially be utilized for growth (e.g., pyruvate, starch, dextrose). In contrast, the second medium, MOPS-glucose, is a defined mineral salts medium with a single carbon source that also supports growth of each of the individual strains. The MOPS minimal medium was prepared as described in [41] and 0.2% glucose was added as the carbon source. The R2A complex medium was prepared as described in [42]. The bacterial community was cultured through a serial dilution protocol. Each strain was inoculated from an agar plate and grown individually in a test tube with 10 mL R2A or MOPS liquid medium and cultivated in a shaking incubator at 30 °C and 200 rpm for 48 h. The growth of each strain was evaluated by measuring optical density at 600 nm (OD<sub>600</sub>) and then the cultures were normalized to the culture which had the lowest OD<sub>600</sub> (0.228 in MOPS and 0.987 in R2A). Equal volumes of the normalized culture were mixed together, and the community was transferred into fresh medium using a 1:10 dilution. The starting inoculum for each medium contained a mixture of the ten bacterial strains at approximately equal concentrations as determined by OD<sub>600</sub>. Triplicate cultures, for both media, were incubated for 48 h at 30 °C with shaking at 200 rpm and passaged every 48 h at a 1:10 dilution for 15 passages. All bacterial strains in community are either obligate aerobes or facultative anaerobes, and shaking conditions ensured aeration required for growth. At the end of each passage, the bacterial cells were collected by centrifugation at 12,000 rpm for 15 min and stored at –80 °C.

### 2.2. DNA extraction and microbial relative abundance analysis

To determine the composition of the bacterial communities at each passage, genomic DNA was extracted from three replicate cell pellets collected at the end of each passage using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) according to the manufacturer's instructions. DNA concentrations were determined on an Invitrogen Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA). The 16S rRNA gene was amplified using 515

**Table 1**  
General features of the bacterial isolates utilized for community experiments.

Strain	Phylogeny	Genome size (bp)	Number of coding sequence (CDS)	G + C content (%)	Reference
<i>Pantoea</i> sp. YR343	$\gamma$ -Proteobacteria	5,391,843	4985	54.5	[7]
<i>Pseudomonas</i> sp. GM17	$\gamma$ -Proteobacteria	6,866,808	6199	62.8	[7]
<i>Sphingobium</i> sp. AP49	$\alpha$ -Proteobacteria	4,506,188	4280	64.1	[7]
<i>Rhizobium</i> sp. CF142	$\alpha$ -Proteobacteria	6,068,985	5714	66.8	[7]
<i>Variovorax</i> sp. CF313	$\beta$ -Proteobacteria	7,510,066	7608	60.1	[7]
<i>Bacillus</i> sp. BC15	Firmicutes	6,240,445	5413	62.9	[40]
<i>Caulobacter</i> sp. AP07	$\alpha$ -Proteobacteria	5,615,958	4915	68.9	[7]
<i>Duganella</i> sp. CF402	$\beta$ -Proteobacteria	11,048,459	9632	61.9	[39]
<i>Streptomyces mirabilis</i> YR139	Actinobacteria	5,742,731	5635	34.8	[39]
<i>Paraburkholderia</i> sp. BT03	$\beta$ -Proteobacteria	11,452,267	11,227	70.3	[38]



**Fig. 1.** Overview of experimental design for the bottom-up assembly of stable communities utilizing defined bacterial isolates.

forward and 806 reverse primers, barcoded, and prepared for sequencing using previously described methods [43]. The amplified products were then sequenced at Oak Ridge National Laboratory with a single  $2 \times 300$  paired-end sequencing kit on Illumina MiSeq (Illumina Inc., San Diego, CA). Raw sequences were trimmed using cutadapt (v.1.18) to remove primers. The sequences were imported into QIIME2 (v. 2019.1) for further processing. Sequence variants were assigned using DADA2 implemented in QIIME2 plugin. Taxonomy was assigned using the consensus vsearch option in QIIME2 against a database of 16S sequences from the 10 community members [44–47]. The resulting sequence variant table, mapping file, and taxonomy file were imported into Phyloseq (version 1.22.3) in R (version 3.4.4) for visualization. We corrected for 16S rRNA copy number using a custom R script (available at <https://github.com/dlcarper/CopyNumberCorrection>) and the number of copies obtained from the isolate genomes. Raw sequences were deposited in the NCBI SRA database under bioproject number PRJNA658537.

### 2.3. Maximum growth rate measurements

Individual bacterial strains were inoculated from R2A agar plates and grown in 10 mL of R2A or MOPS liquid media in test tubes for 48 h as seed cultures. These seed cultures were measured for  $OD_{600}$  and then normalized to the culture which had the lowest  $OD_{600}$  value. The seed culture of each strain was inoculated (5%, v/v) to the same liquid medium (10 mL) in a test tube and cultured at

30 °C with shaking at 200 rpm for 48 h. Optical density measurements at 600 nm were taken at 2-h intervals during the exponential phase of the growth curves. Bacterial cells were collected by centrifugation at 12,000 rpm and heated at 60 °C to measure the cell dry weight (CDW). A calibration curve for each strain was created by correlating the  $OD_{600}$  to the CDW and applied to transform the OD values to cell concentration. The maximum growth rates of bacterial strains were calculated as the slope of the plot between  $\ln(X/X_0)$  and time during the exponential growth phase. The X and  $X_0$  are the cell concentration (mg/mL) at time t and the time at the beginning of the exponential phase [48]. The length of lag phase was determined from a logarithmic plot of the growth curve and identified as the time point that results from extrapolating the slope of the exponential growth phase to the intersection with the initial inoculum OD value [49].

### 2.4. Pairwise interaction screens

Pairwise interactions were performed as described previously [11], briefly 5  $\mu$ L of overnight R2A liquid culture of a selected microbe was spotted on R2A agar plate containing a lawn of another selected microbe and incubated at 25 °C for 48 h. Results were recorded as antagonistic when a zone of inhibition was observed, commensal when there was no obvious phenotype observed, or mutualistic when enhanced growth was observed around the test strain.

## 2.5. Individual and community modeling and flux balance analysis

The Department of Energy Systems Biology Knowledgebase (KBase) platform ([www.kbase.us](http://www.kbase.us)) was applied for the automated reconstruction of individual metabolic models and flux balance analysis (FBA) for the 10 bacterial isolates [50]. The FBA was carried out following previously described methods [51]. In brief, the genome of each strain was annotated by Annotate Microbial Genome app, which is based on Rapid Annotations using Subsystems Technology (RAST) toolkit to annotate prokaryotic genomes [52–54]. The metabolic modeling (Build Metabolic Model app) was performed using the annotated genome of each strain, and then the model was gap-filled (Gapfill Metabolic Model app) on R2A or MOPS media growth conditions to ensure they are qualified to simulate growth. The statistics on the constructed metabolic models for each strain are shown in Tables S1 and S2. Finally, FBA of the gapfilled model was predicted using biomass production as the objective value.

For community modeling, the individual models of the major surviving community members determined by 16S amplicon sequencing were merged into community metabolic model using the Merge Metabolic Models into Community Model app. The relative ratio of the community members was adjusted by the 16S rRNA experimentally determined ratio of the final passage of co-culture of 10 bacterial isolates (Table S3) using the method in [51]. Then, the adjusted community models were gapfilled and analyzed by FBA using the same procedure as individual models to estimate the fluxes of biochemical reactions and biomass production in the bacterial community. For constraining microbial growth in the community models, a calibration curve between the carbon uptake limit and growth rate was built using the FBA model of each individual strain, and the carbon uptake limit was calculated based on the experimental growth rate of each strain using the linear regression equation of the corresponding calibration curve. The carbon uptake limit of the community model was the summation of individual growth rate  $\times$  relative percentage. The models used in this study are accessible online (<https://narrative.kbase.us/narrative/73218> and <https://narrative.kbase.us/narrative/73221>).

## 2.6. Community proteomics analysis

### 2.6.1. Cellular protein extraction

Cell pellets, collected at the end of the passage, were solubilized in 600  $\mu$ L of lysis buffer (4% sodium dodecyl sulfate (SDS) in 100 mM Tris, pH 8.0) supplemented with 1  $\times$  Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific, Waltham, MA). Samples were vortexed and then further disrupted with a Bullet Blender storm 24 (Next Advance) bead beater for 5 min using 0.15 mm Zirconium oxide beads at 3:1 sample to bead ratio. Samples were then placed in a heat-block for 10 min at 90  $^{\circ}$ C. Protein concentration was measured using a Nanodrop One spectrophotometer (Thermo Fisher Scientific, Waltham, MA). Protein disulfide bonds were reduced with 10 mM dithiothreitol (DTT) at 90  $^{\circ}$ C for 10 min and then alkylated with 30 mM iodoacetamide (IAA) for 15 min in the dark to prevent the reformation of disulfide bonds. As previously described [55], proteins were extracted by protein aggregation capture on Ser-Mag beads at 1:1 protein to beads ratio [55] and digested with sequencing grade Pierce trypsin (Thermo Fisher Scientific, Waltham, MA) at 1:75 (wt/wt) protein:trypsin ratio for overnight followed by a second 3 h digestion at 37  $^{\circ}$ C at constant shaking. Samples were vortexed and centrifuged at 12,000g for 15 min in 10 kDa molecular weight spin column filters (Vivaspin 500). Tryptic peptide flow-throughs were collected and then desalted using Pierce peptide desalting spin columns (Thermo Fisher Scientific, Waltham, MA) as per the manufacturer's instructions. Desalted peptides were vacuum dried with a SpeedVac Concentrator (Thermo Fisher Scientific, Waltham, MA) and then resolubilized in 0.1% formic acid. Peptide

concentrations were measured using a nanodrop (Thermo Fisher Scientific, Waltham, MA) and transferred to the auto-sampler vials for LC-MS/MS measurement.

### 2.6.2. Protein identification and quantification

Each sample was analyzed using two-dimensional (2D) liquid chromatography (LC) on an Ultimate 3000 RSLCnano system (Thermo Fisher Scientific, Waltham, MA) coupled with a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific, Waltham, MA). For each sample, an aliquot of digested peptide mixture was injected across an in-house built strong cation exchange (SCX) Luna trap column (5  $\mu$ m, 150  $\mu$ m  $\times$  50 mm; Phenomenex, Torrance, CA) followed by a nanoEase symmetry reversed-phase (RP) C18 trap column (5  $\mu$ m, 300  $\mu$ m  $\times$  50 mm; Waters, Milford, MA) and washed with an aqueous solvent. Cellular peptide mixtures were separated and analyzed across three successive SCX fractions of increasing concentrations of ammonium acetate (35 mM, 50 mM, and 500 mM), each followed by a 100-minute organic gradient (25 nL/min flow rate) to separate peptides across an in-house pulled nanospray emitter analytical column (75  $\mu$ m  $\times$  350 mm) packed with C18 Kinetex RP C18 resin (1.7  $\mu$ m; Phenomenex, Torrance, CA). All MS data were acquired with Thermo Xcalibur (version 4.2.47) using the topN method where N could be up to 10. Target values for the full scan MS spectra were  $3 \times 10^6$  charges in the 300–1500  $m/z$  range with a maximum injection time of 25 ms. Transient times corresponding to a resolution of 70,000 at  $m/z$  200 were chosen. A 1.6  $m/z$  isolation window and fragmentation of precursor ions was performed by higher-energy C-trap dissociation (HCD) with a normalized collision energy of 27 eV. MS/MS scans were performed at a resolution of 17,500 at  $m/z$  200 with an ion target value of  $1 \times 10^5$  and a maximum injection time of 50 ms. Dynamic exclusion was set to 20 s to avoid repeated sequencing of peptides.

All MS raw data files were analyzed using the Proteome Discoverer software (version 2.3, Thermo Fisher Scientific, Waltham, MA). Raw files were processed by the SEQUEST HT database search algorithm [56] and confidence in peptide-to-spectrum (PSM) matching was evaluated by Percolator [57]. Peptide and PSMs were considered identified at  $q < 0.01$  and proteins were required to have at least one unique peptide sequence. Protein relative abundance values were calculated by summing together peptide extracted ion chromatograms. Protein abundances were normalized by LOESS and mean central tendency procedures performed on log<sub>2</sub>-transformed values by InferoRDN [58]. From this normalized dataset, protein abundances subset for each microbe were extracted and further mean-centered by InferoRDN. All proteomics spectral data in this study were deposited at the ProteomeXchange Consortium via the MASSIVE repository (<https://massive.ucsd.edu/>). The data can be reviewed under the username “MSV000086551\_reviewer” and password “PMI”.

Organism relative abundance (i.e., population size) was assessed using two approaches [59]: 1) total protein count per organism divided by the total count of proteins per community and 2) summed total relative protein abundance per organism divided by the summed total protein abundance per community. Note, it is important to evaluate and compare both approaches to account for biases in protein expression. In this study, we found these two approaches provided similar results and report relative organism abundances using the second approach.

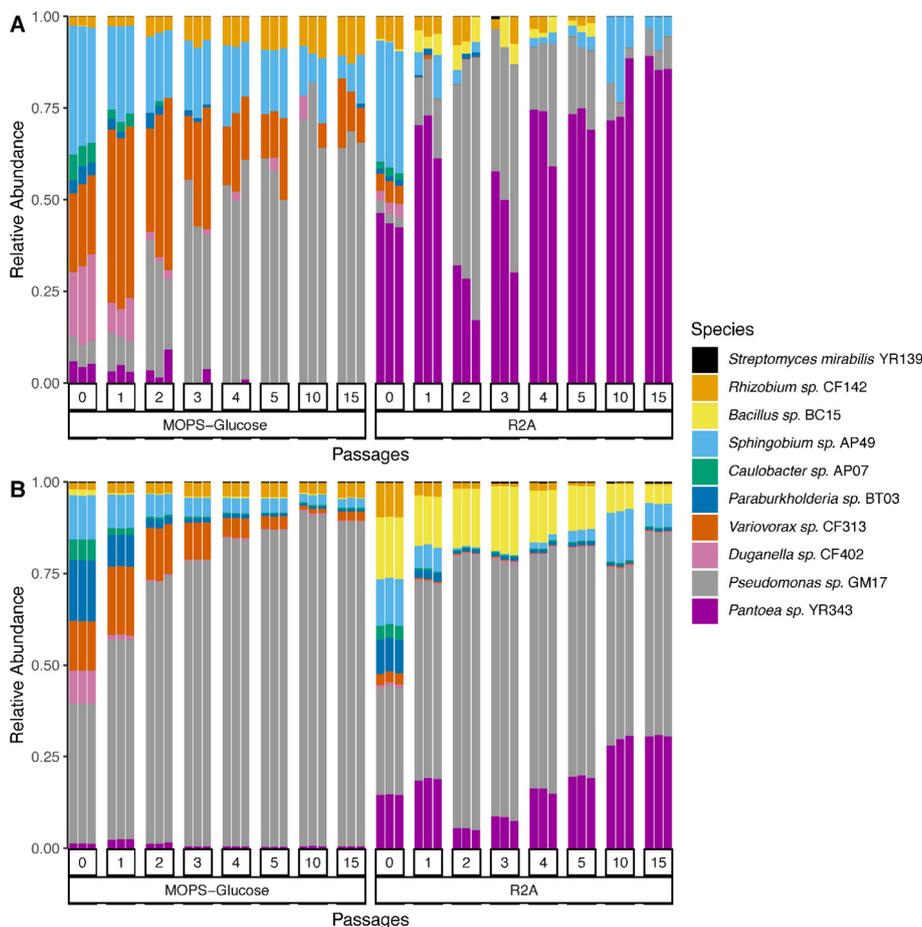
## 3. Results and discussion

### 3.1. Stable community structure in minimal and complex media

Stable microbial communities were formed by serial transfer of batch cultures containing a mixture of ten, phylogenetically

diverse bacterial strains derived from the *Populus* rhizosphere (Table 1). These strains represent phyla that are abundant in the rhizosphere of plants [43,60]. The use of batch cultures allows for effective exchange of metabolites and the preparation of samples for analytical measurements. The component strains' genomes are sequenced and comprise three  $\alpha$ -Proteobacteria, three  $\beta$ -Proteobacteria, two  $\gamma$ -Proteobacteria, one Firmicute and one Actinobacterium. These batch cultures were subsequently transferred to fresh medium every 48 h and the relative proportion of each member was analyzed by 16S rRNA gene amplicon Illumina sequencing and quantitative metaproteomics. They are in general agreement regarding the trends in membership of the stable microbial communities that are formed in the two different media environments (Fig. 2). In both environments, the bacterial diversity decreases from the initial inoculation, and both measurement approaches show changes in the relative proportions of the bacterial strains that settle into a similar distribution beginning with approximately the fifth dilution cycle (see Supplemental Fig. S1). This observation is consistent with the expectation that competition for local resources will reduce the genotypic diversity within a bacterial community [61]. In the MOPS minimal medium, four strains consistently dominate in abundance and stably coexist starting with passage No. 4. Organismal abundance trends in R2A show a fluctuation in organism relative abundances occurring between passages two and five that substantially alters the abundance of several members of the community until a community stably coexists starting with passage No. 10.

In general, 16S rRNA gene amplicon sequencing and metaproteomics results provide similar estimates of the relative microbial abundance distribution for each measured passage in the minimal MOPS medium but differ in the rich R2A medium. In MOPS medium, *Pseudomonas* sp. GM17 is the most abundant member along with three other strains, *Variovorax* sp. CF313, *Rhizobium* sp. CF142 and *Sphingobium* sp. AP49, that persist in consistent proportions (Fig. 2). Estimates of organism proportion in the R2A medium, on the other hand, are quite different between the two approaches. For 16S rRNA gene amplicon sequencing, three members dominate in abundance when grown in R2A complex medium, and *Pantoea* sp. YR343 is the member with highest content after 15 passages (Fig. 2A). The less abundant strains in R2A medium are *Pseudomonas* sp. GM17 and *Sphingobium* sp. AP49. Metaproteome analysis reveals the same dominant members, albeit the relative abundance of *Pantoea* sp. YR343 is strikingly different (Fig. 2B). Additionally, there is a notable abundance of *Bacillus* sp. BC15 not seen in 16S rRNA gene amplicon sequencing data. These differences can be attributed to measurement distinctions that result from using either a DNA- or protein-based approach to assess organism proportions in microbial communities. For example, examination of the proteins expressed by *Bacillus* sp. across the passages reveals an abundance of sporulation-related processes. Spores are notoriously challenging to measure by 16S rRNA gene amplicon sequencing because they are difficult to lyse, which negatively impacts DNA extraction efficiency [62]. Selective PCR amplification before amplicon sequencing may be another poten-



**Fig. 2.** Analysis of the relative abundances of the 10 bacterial strains after sequential passages in MOPS minimal and R2A complex media. The relative abundances of each bacterial strain in the community are based on A) 16S rRNA gene amplicon sequencing results and B) metaproteomic results. Passage 0 represents the end of the first growth cycle after the initial inoculation. Each bar is a replicate, with three replicates per passage. The numbers shown on the bottom represent the passage number for those samples.

tial source of bias, since the designation of a perfectly matching universal primer is not possible [63,64]. Nevertheless, there is clearly a benefit in using both approaches. For instance, the observed differences in the relative abundance of *Pantoea* sp. YR343 is likely because 16S rRNA gene amplicon sequencing measures DNA from viable cells as well as ‘relic’ DNA (i.e., DNA from dead cells), whereas metaproteomics is a more accurate estimate of biomass from viable, functioning cells [59,65,66]. Assimilating the results between these two measurements suggests that *Pantoea* sp. YR343 may experience a population dieback event prior to the timepoint of sampling.

### 3.2. Experimentally determined growth rates for the individual bacterial isolates

Individual microbial growth rates can impact interactions within a community [33], experimentally determined growth rates of the 10 individual strains were obtained for both media and show substantial differences. In the MOPS medium, strain *Pseudomonas* sp. GM17 has the highest growth rate ( $0.463 \text{ h}^{-1}$ ) among the group (Fig. 3A). It is also the dominant strain in the community growth experiment (Fig. 2). In the same medium, *Sphingobium* sp. AP49 ( $0.430 \text{ h}^{-1}$ ) has a maximum growth rate that is similar to that observed with the *Pseudomonas* sp. GM17 and also maintains a presence in the community. Compared to these two organisms, *Paraburkholderia* sp. BT03 ( $0.379 \text{ h}^{-1}$ ), *Variovorax* sp. CF313 ( $0.362 \text{ h}^{-1}$ ), *Pantoea* sp. YR343 ( $0.346 \text{ h}^{-1}$ ), *Duganella* sp. CF402 ( $0.333 \text{ h}^{-1}$ ) and *Caulobacter* sp. AP07 ( $0.296 \text{ h}^{-1}$ ) have slightly lower maximum growth rates, but only *Variovorax* sp. CF313 persists in the community. Three organisms, *Rhizobium* sp. CF142 ( $0.082 \text{ h}^{-1}$ ), *Bacillus* sp. BC15 ( $0.055 \text{ h}^{-1}$ ) and *Streptomyces mirabilis* YR139 ( $0.039 \text{ h}^{-1}$ ), have considerably lower growth rates when compared to the other organisms. Interestingly, despite having a relatively slow growth rate in the MOPS medium, *Rhizobium* sp. CF142 prevails as a dominant community member.

In R2A medium, there is a strong correlation between the growth rates of individual strains and community membership. Not surprisingly, the growth rates of all of the selected microbes are higher in this rich medium when compared to growth rates observed in MOPS medium. Overall, *Pantoea* sp. YR343 ( $0.656 \text{ h}^{-1}$ ), *Pseudomonas* sp. GM17 ( $0.563 \text{ h}^{-1}$ ) and *Sphingobium* sp. AP49 ( $0.479 \text{ h}^{-1}$ ) are relatively fast growers in R2A and are significant components of the emergent community. In contrast, *Paraburkholderia* sp. BT03 ( $0.472 \text{ h}^{-1}$ ), *Bacillus* sp. BC15 ( $0.442 \text{ h}^{-1}$ ), and *Duganella* sp. CF402 ( $0.411 \text{ h}^{-1}$ ), despite showing growth rates only slightly lower than *Sphingobium* sp. AP49, do not maintain a significant presence in successive community growth cycles (Fig. 3B).

Comparison of the individual member growth rates to the observed community composition indicates complex relationships among the community members. On the one hand, the presence of fast-growing microbes in the emerging community composition is expected and consistent with observations in other systems. The competitive exclusion principle predicts that when the bacterial members in a community compete for the same resources, the fitter strain will outcompete the other members [6] and has a greater opportunity for recurrent colonization that can allow for persistence in the community [67]. Yet, several fast-growing microbes are absent in the final passages of the community and slow growers are present. Often, slower growing microbes persist in communities and allow for maintenance of diversity. Here, the well-mixed conditions promote exchange of metabolites during bacterial growth and prevents spatial structuring that often allows for maintenance of slower growing organisms [22,68]. Therefore, bacterial strains must collectively adjust their behavior and selectively cooperate in order to emerge into a community with stable propor-

tions. The resulting supportive associations among community members likely proceeds through metabolic interactions such as the cross-feeding of essential nutrients [61]. One particularly interesting observation to this point is the persistence of *Rhizobium* sp. CF142 in the MOPS minimal medium community. When grown in monoculture in MOPS minimal medium, this strain has a slower growth rate and lower final OD when compared to the other strains. It could be speculated that metabolic cooperation emerges in these mixed microbial communities increasing the fitness of strain CF142, potentially by providing some missing nutrients made available by other community members [69,70]. This is consistent with the existence of both cooperative and competitive associations between the component members that lead to the formation of a community network structure [71].

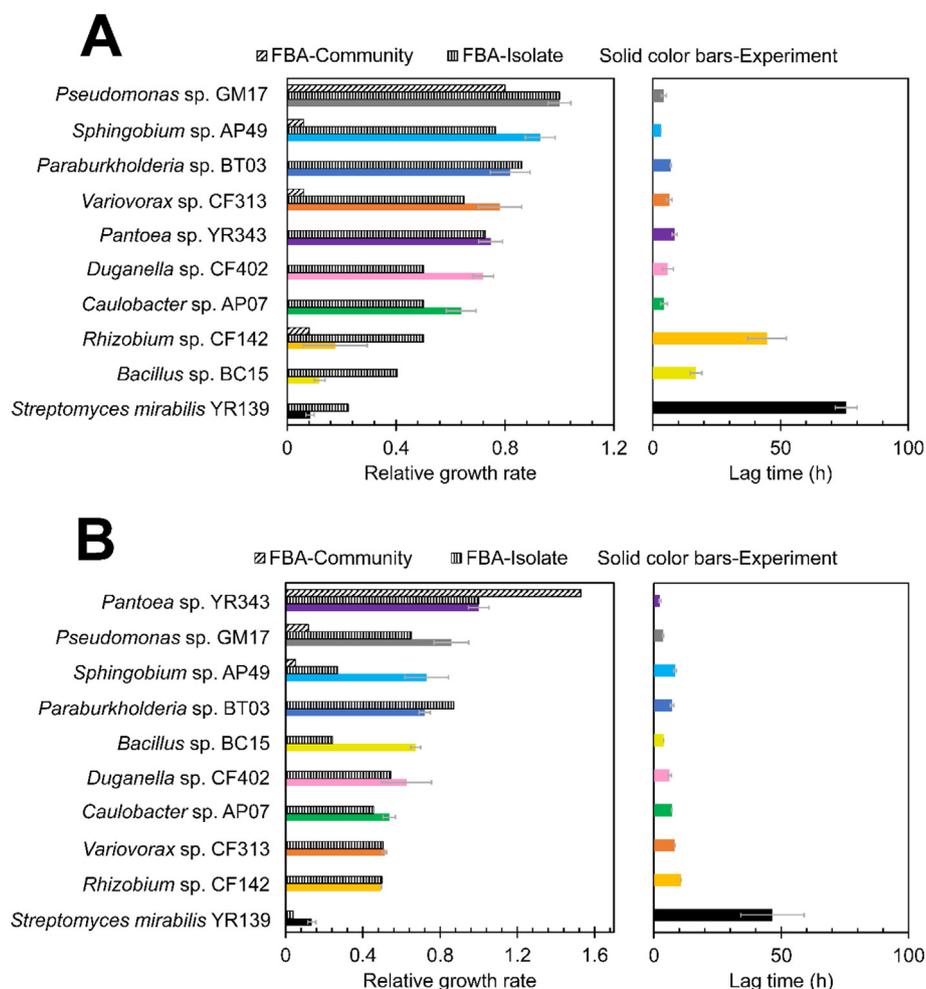
### 3.3. FBA-based growth prediction models for individual microbes

To gain a better understanding of individual strain metabolism and the potential variety of exchanged metabolites, FBA models of each of the component microbes were generated. By estimating the reaction fluxes to generate biomass constituents, the growth rate of the microorganism can be predicted [72]. The maximum relative growth rates of the ten bacterial isolates in R2A and MOPS media are displayed in Fig. 3. In these models, the objective value of growth is determined by setting a maximum glucose uptake flux of  $100 \text{ mmol/g DCW/h}$ . Among the 10 *Populus* bacterial isolates in MOPS medium, *Pseudomonas* sp. GM17 has the highest predicted and actual growth rate and was used for normalization. Similarly, among the 10 *Populus* bacterial isolates grown in R2A medium, *Pantoea* sp. YR343 has the maximal predicted and actual growth rate and was used as the reference standard.

In MOPS medium, the hierarchy of relative predicted growth rates generally matches the experimentally observed growth rates. In this medium, the maximal growth rate predictions for the slowest growing organisms, *Rhizobium* sp. CF142, *Bacillus* sp. BC15, and *Streptomyces mirabilis* YR139 are significantly overestimated and likely reflect imperfect understanding of metabolism in these organisms. When compared to the others, these organisms all have long lag phases (Fig. 3A) which likely reflects unknown adaptations to environmental conditions [73]. Further, these organisms may adopt different growth strategies that do not prioritize the conversion of glucose to biomass. In the case of *Streptomyces mirabilis* YR139, the unusual growth and morphological characteristics of this genus can contribute to experimental and predictive errors.

In R2A media, growth rate predictions show a different trend. In general, the predicted relative growth rates of the slower growing organisms match the experimentally observed maximal growth rates. In contrast, growth rate predictions are poor for several of the faster growing microbes (Fig. 3B). In particular, growth rate predictions are significantly underestimated for *Sphingobium* sp. AP49 and *Bacillus* sp. BC15 in this complex medium. Again, this may reflect unknown limits on metabolism for these species. The FBA predicted growth rates assume ideal conditions; all nutrients in the medium are made available at the maximum uptake flux. In the growth experiments of those individual strains, the growth rate could not be ideally as high as the FBA models.

In the FBA model using R2A medium, the dominant strain *Pantoea* sp. YR343 has the highest number of exchange reactions of nutrients (Table S4), and it has more transporters based on genome annotation compared with the other community members. For *Pantoea* sp. YR343, the higher number of transporter genes may be related to its stronger metabolic interaction potential encoded in the genome. In the nutrient rich complex medium, *Pantoea* sp. YR343 is capable of utilizing a variety of nutrients which is consistent with the observation that this organism has the highest simulated growth rate. In contrast, the number of exchange reactions



**Fig. 3.** Comparison of experimentally determined growth parameters with FBA model predicted growth. A) Relative growth rates and lag times for the individual strains in MOPS medium; B) relative growth rates and lag times for the strains in R2A medium. Each data column of experimental results represents the mean and error bars are the standard deviation over three parallel experiments.

identified in individual FBA models carried out in MOPS medium is similar (Table S5). This likely results from the minimal medium environment where nutrition is relatively limited compared with the complex R2A medium. In MOPS medium models, the medium-specific metabolic processes may override the strain-specific metabolisms, leading to much less difference in simulated growth rates among the 10 isolates when compared to models employing R2A medium [74]. The discrepancy of simulated growth rates of the same strain between MOPS and R2A media also corresponds to the experimental observation that bacteria utilize their metabolism for survival in the minimal medium, in contrast they tend to have more active growth in the complex medium [75].

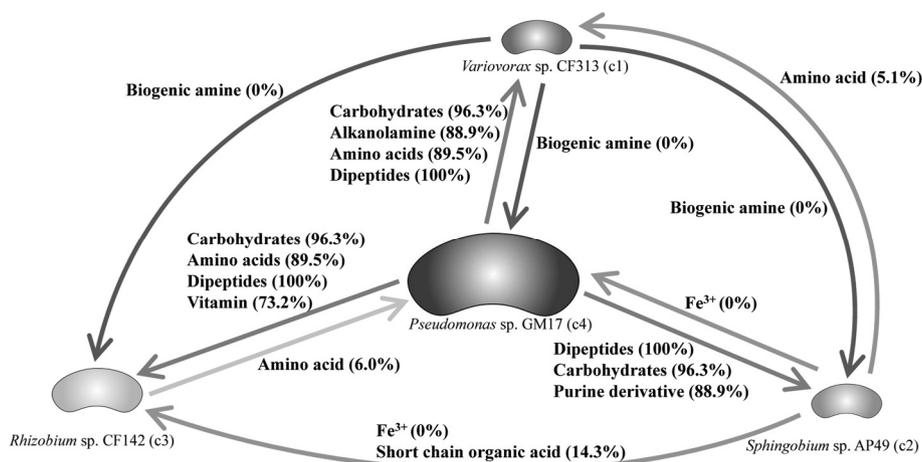
### 3.4. Evaluation of microbial community models

#### 3.4.1. FBA-based growth prediction for community models

FBA-based community models were assembled in order to identify microbial features that account for the observed stability of the community and for assessing the suitability of these models and their use in understanding the molecular genetic bases for the resulting community structure. Flux changes of the metabolites involved in exchange may explain the interaction mechanisms between the component members. Compartmentalized models were created using the KBase platform to allow for the community members to secrete and take up metabolites from a

shared environment [17,51]. To create these models, the genome-scale metabolic models of the primary constituents of the final community were combined using the relative ratio of the community members determined from the 16S rRNA gene marker data for the final passage. These final, persistent strains are considered the best-performing species and as community drivers that affect dependent species and community organization [76,77].

The FBA-based community models predict altered growth rates for the constituent members when compared to their individual growth rates. In MOPS medium, the dominant strain *Pseudomonas* sp. GM17 is predicted to have the fastest growth among the consortium members. However, this growth rate is lower when compared with its individual FBA model (Fig. 3A). This is most likely due to competition with other community members in this limited nutrient environment and thus a lower growth rate is not surprising. In comparison, the three other major strains in this consortium have significantly lower, but similar, predicted growth rates when compared to *Pseudomonas* sp. GM17 as well as to their individual models (Fig. 3A). Altered growth rates can result from competitive interactions and have been observed in other studies using minimal media for microbial community assembly [78–80]. These prior studies indicate that the metabolic capability associated with each strain can influence community composition and lead to the survival of the strongest competitors [22,81].



**Fig. 4.** Predicted metabolite exchange among the four dominant members of the microbial community formed in MOPS medium as proposed by the community FBA model. The percentage of detected enzymes by metaproteomics analyses is shown in parentheses.

When using R2A medium in the community metabolic model, the dominant strain *Pantoea* sp. YR343 has a much higher theoretical growth rate than that of its individual model. This suggests that it is beneficial for *Pantoea* sp. YR343 to grow in the presence of the other two members in the community. The higher predicted growth rate for *Pantoea* sp. YR343 in the community model may relate to the organism's broad spectrum of transport reactions [82]. In a community, the bacterial members that are metabolic generalists have a better chance of survival compared to those that are adapted to specific substrates [22,83]. For both media, the community-FBA models support the dominant presence of a fast-growing member and suggest that members of the community influence each other's growth rate.

### 3.4.2. FBA-based predictions of metabolic exchanges

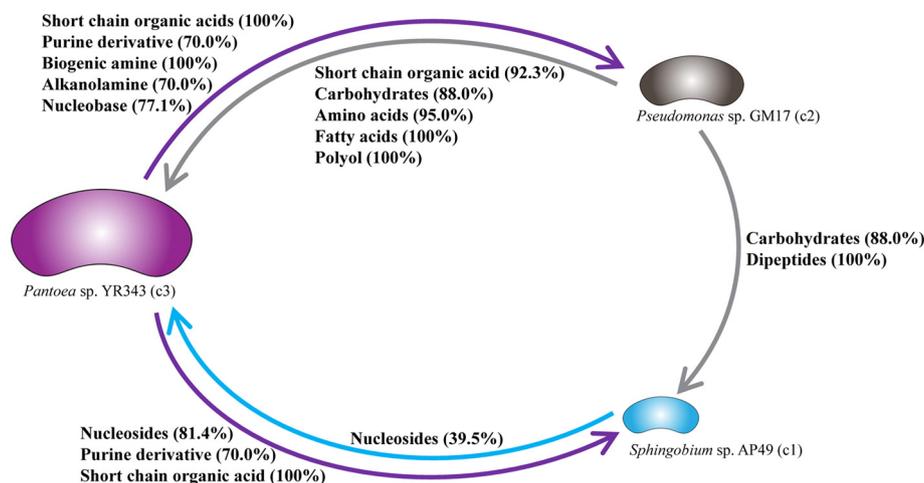
Limiting the uptake flux of carbon is the foundation of a constraint-based FBA model and allows prediction of the distribution of metabolic fluxes that depend on the medium [84]. Without a carbon source uptake limitation, metabolite exchanges between community members will be overestimated. Therefore, a limitation of carbon uptake flux was added to the FBA models for the communities modeled in either MOPS or R2A medium. Calibration curves relating carbon source uptake flux and growth rate objective value were established for each community member, and the individual carbon uptake flux limitation of each strain was calculated based on the experimental growth rate. The carbon uptake limit for a community FBA model was calculated as the sum of the relative ratios of the individual microbial components multiplied by their individual carbon uptake rates.

The extracellular metabolites involved in interspecies exchanges can be predicted by the community FBA models [51]. Fig. 4 illustrates the predicted metabolite exchanges among the four component members in MOPS medium. The dominant strain, *Pseudomonas* sp. GM17 is predicted to supply more metabolites to the other community members than it receives. Key among these predicted metabolites are amino acids, sugars and purine derivatives. In the minimal medium environment, cellular building materials must be wholly prepared from the glucose carbon source, or, in a community environment, scavenged from the excretions of other microbes. Accessing excreted metabolites may be vital for the maintenance of those members with minor content in the community. In turn, *Pseudomonas* sp. GM17 may rely on the production of metabolites from these minor community members as evidenced by its lower growth rate in the community FBA model when compared with its individual FBA model. The metapro-

teomics results provide an opportunity to assess whether the appropriate pathways are represented. The percentage of the detected enzymes in the metabolite supplier is related to the total protein involved in the KEGG pathways for the exchanged metabolites. Considering the dominant microbial component *Pseudomonas* sp. GM17, there is high representation of the enzymes related to the metabolites supplied by this organism (Fig. 4). In contrast, enzymes involved in the preparation of metabolites shared from the other organisms are low. This lower representation is a consequence of the relatively lower number of proteins detected from these organisms relative to *Pseudomonas* sp. GM17.

In the community model using R2A medium, metabolite exchange can also be predicted. Here, the dominant strain *Pantoea* sp. YR343, based on the 16S rRNA marker data, is expected to export a greater range of metabolites than it receives from the other two members of the community (Fig. 5). Organic acids, purine derivatives and biogenic amines are predicted to be excreted and support the growth of the other community members. The community FBA model predicts a faster growth rate for *Pantoea* sp. YR343 in the community model when compared to the individual model and this may result from access to metabolites excreted by the other community members. Again, the community metaproteomics results can be used to assess the presence of relevant pathways. The high representation of both *Pantoea* sp. YR343 and *Pseudomonas* sp. GM17 in the metaproteomics data allows detection of most of the enzymes expected to participate in the FBA-based predictions. As in the case of the minimal medium environment, representation of the minor microbial components is poor, and confirmation of relevant components of metabolism is much lower when compared to the major microbial components.

The combination of omics data with FBA modeling can aid with understanding community function and stability. Here, the metaproteomics results help support predictions of the FBA model by providing evidence for the metabolic capabilities of each community member and making them more faithful representations of the biological system being interrogated [23]. While the coverage of enzymes implicated in the exchanged metabolites is relatively high for prominent members of the community, as communities grow in membership diversity, metaproteomic sequencing depth will need to improve for the more minor members, which may disclose their survival mechanisms. Additionally, the integration of metabolomics data will be valuable for constraining FBA models and for confirming predictions of interdependencies between community members.



**Fig. 5.** Predicted metabolites exchange among the three dominant members of the microbial community formed in R2A medium as proposed by the community FBA model. The percentage of detected enzymes by metaproteomics analyses is shown in parentheses.

**Table 2**

Pairwise interaction screen results. Strain designations across top of table indicate lawn of microbes spread on R2A agar plate and designations on left indicate cells spotted on center of lawn. + indicates a positive interaction while - indicates an antagonistic interaction. Empty cells indicate no obvious colony phenotype change.

Genus	Strain	YR343	GM17	AP49	CF142	CF313	BC15	AP07	CF402	YR139*	BT03
<i>Pantoea</i>	YR343				+	+				ND	
<i>Pseudomonas</i>	GM17			-	-	-	-	-	-	ND	-
<i>Sphingobium</i>	AP49					-	+			ND	
<i>Rhizobium</i>	CF142						+			ND	
<i>Variovorax</i>	CF313				-					ND	-
<i>Bacillus</i>	BC15								-	ND	
<i>Caulobacter</i>	AP07			-						ND	
<i>Duganella</i>	CF402					+	+			ND	
<i>Streptomyces</i>	YR139			-	-			-		ND	
<i>Paraburkholderia</i>	BT03				-					ND	

\* Due to the growth characteristics of *Streptomyces* sp. YR139, a lawn of bacteria could not be prepared and resulted in no data (ND).

The inherent complexity of metabolic interactions is a challenge in the modeling of microbial communities [85]. The genome-scale, mechanistic modeling provided by the FBA approach is presently insufficient to account for a large fraction of intracellular networks or assess dynamic, population level changes that likely lead to community structuring [72]. Integration with population-level dynamic models, such Lotka-Volterra (LV) [86–88], or r/K selection strategies [89] may help to describe the temporal progress of species abundances and community formation processes. Dynamic FBA, which simulates the dynamics of community growth and substrate consumption in time-dependent processes, can also extend current FBA approaches to temporal changes [90–92]. Effective application of these tools will require new, time dependent global sampling and measurement strategies for verifying the efficacy of dynamic models.

### 3.5. Pairwise interaction screens

While the present metabolic models help to understand interactions that support growth of a stable community, they provide only partial insight into the selection process that leads to community formation. In both tested environments, a fast-growing microbe emerges as a dominant component that is supportive of other community members. However, other relatively fast-growing microbes are out competed in early growth cycles, and their abundance fades from the composition of the community. Antagonistic interactions may be present that facilitate the com-

munity selection process. To assess this possibility, community members were screened for mutualistic, commensal, or antagonistic colony phenotypes in a pairwise interaction screen. While a majority of the interactions appear to be commensal, with no obvious phenotypes, the screen did identify both positive and negative interactions (Table 2). *Pseudomonas* sp. GM17 cells were antagonistic to the growth of the majority of the other community members, while *Pantoea* sp. YR343 demonstrated positive interactions with several strains indicating an obvious role for competition, antimicrobial production and or beneficial interactions in microbial community selection, structure and stability.

Interestingly, although the growth of strain AP49, CF313 and CF142 were inhibited by strain GM17 in pairwise interaction screens, these strains still co-existed with strain GM17 during the 10-member community cultivation. This may be attributed to intertwined metabolic interactions among these four strains, in which the beneficial effect from the metabolites in a shared extracellular environment for growth overwhelms antagonistic effects by strain GM17. Alternatively, different experimental conditions may account for the unexpected result. The multi-member community was grown in a well-mixed liquid medium condition which is different from the static agar plate conditions of the pairwise interaction screen and may account for the observed discrepancy [93]. The integration of temporal modeling of microbial communities with time course of community composition within the growth period in a passage can be promising to accommodate these antagonistic interactions and the dynamic processes that shape community structure.

## 4. Conclusions

Distinct bacterial communities can be formed from a more complex mixture of microbial isolates. Diverse bacterial species, without previously known or expected obligate relationships, were combined and put through dilution cycles. After approximately five cycles, select members prevail and form fairly stable community structures that persist through successive cycles. Further, the resulting community structures depend on the media environment used for the dilution cycles. The described approach to discovering stable, media-dependent emergent communities takes advantage of genome-defined isolates to allow for effective implementation of systems biology tools. Growth curve analyses, metaproteomics and FBA analyses were employed to identify key factors that contribute to the resulting community structure. Growth rate was identified as providing an advantage for a microbial member to maintain presence in the community. Interestingly, some of the fastest growing organisms remain in the final community structure but other rapid growing organisms do not. Further, under minimal medium conditions, a relatively slow growing organism was found to persist. Pairwise interaction measurements highlight that selective antagonistic relationships may contribute to the final structuring of the community. In order to gain a molecular-level understanding of the resulting microbial organization, FBA analyses were performed. Metabolic exchanges between organisms can be identified and likely underpin the shaping of community membership. Metaproteomic results support general findings of the FBA models. However, presently accessible FBA tools primarily account for central metabolite fluxes with cell growth as the final objective. Understanding dynamic processes at the molecular and cell population levels will be required to understand community formation, dynamics, and structure. Improved modeling capabilities, coupled with time dependent measurements and the described, scalable approach to identifying stable communities will facilitate definition of the molecular events that result in microbial community structure and dynamics.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This research was sponsored by the Genomic Science Program, U.S. Department of Energy, Office of Science, Biological and Environmental Research, as part of the Plant Microbe Interfaces Scientific Focus Area (<http://pmi.ornl.gov>). Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2021.03.034>.

## References

- [1] Fisher CK, Mehta P. Identifying keystone species in the human gut microbiome from metagenomic timeseries using sparse linear regression. *PLoS ONE* 2014;9:e102451.
- [2] Harcombe William R, Riehl William J, Dukovski I, Granger Brian R, Betts A, et al. Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Rep* 2014;7:1104–15.
- [3] Hanemaaijer M, Röhling WFM, Olivier BG, Khandelwal RA, Teusink B, et al. Systems modeling approaches for microbial community studies: from metagenomics to inference of the community structure. *Front Microbiol* 2015;6.
- [4] Wright ES, Vetsigian KH. Inhibitory interactions promote frequent bistability among competing bacteria. *Nat Commun* 2016;7:11274.
- [5] Khandelwal RA, Olivier BG, Röhling WFM, Teusink B, Bruggeman FJ. Community flux balance analysis for microbial consortia at balanced growth. *PLoS ONE* 2013;8:e64567.
- [6] Gilmore SP, Lankiewicz TS, Wilken SE, Brown JL, Sexton JA, et al. Top-down enrichment guides in formation of synthetic microbial consortia for biomass degradation. *ACS Synth Biol* 2019;8:2174–85.
- [7] Brown SD, Utturkar SM, Klingeman DM, Johnson CM, Martin SL, et al. Twenty-one genome sequences from *Pseudomonas* species and 19 genome sequences from diverse bacteria isolated from the rhizosphere and endosphere of *Populus deltoides*. *J Bacteriol* 2012;194:5991.
- [8] Kotoky R, Pandey P. Difference in the rhizosphere microbiome of melia azedarach during removal of benzo(a)pyrene from cadmium co-contaminated soil. *Chemosphere* 2020;258:127175.
- [9] Yin L, Liu B, Wang H, Zhang Y, Wang S, et al. The rhizosphere microbiome of *Mikania micrantha* provides insight into adaptation and invasion. *Front Microbiol* 2020;11:1462.
- [10] de la Torre-Hernández ME, Salinas-Virgen LI, Aguirre-Garrido JF, Fernández-González AJ, Martínez-Abarca F, et al. Composition, structure, and GPRT traits of the rhizospheric bacterial communities associated with wild and cultivated *Echinocactus platyacanthus* and *Neobuxbaumia polylopha*. *Front Microbiol* 2020;11:1424.
- [11] Timm CM, Campbell AG, Utturkar SM, Jun S-R, Parales RE, et al. Metabolic functions of *Pseudomonas fluorescens* strains from *Populus deltoides* depend on rhizosphere or endosphere isolation compartment. *Front Microbiol* 2015;6:1118.
- [12] Wagner MR, Lundberg DS, del Rio TG, Tringe SG, Dangl JL, et al. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat Commun* 2016;7:12151.
- [13] Pent M, Pöldmaa K, Bahram M. Bacterial communities in boreal forest mushrooms are shaped both by soil parameters and host identity. *Front Microbiol* 2017;8:836.
- [14] Pii Y, Mimmo T, Tomasi N, Terzano R, Cesco S, et al. Microbial interactions in the rhizosphere: Beneficial influences of plant growth-promoting rhizobacteria on nutrient acquisition process. A review. *Biol Fertil Soils* 2015;51:403–15.
- [15] Timm CM, Carter KR, Carrell AA, Jun S-R, Jawdy SS, et al. Abiotic stresses shift belowground *Populus*-associated bacteria toward a core stress microbiome. *mSystems* 2018;3:e00070–17.
- [16] Zeleznik A, Andrejev S, Ponomarova O, Mende DR, Bork P, et al. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *PNAS* 2015;112:6449.
- [17] Tan J, Zuniga C, Zengler K. Unraveling interactions in microbial communities-from co-cultures to microbiomes. *J Microbiol* 2015;53:295–305.
- [18] De Roy K, Marzorati M, Van den Abbeele P, Van de Wiele T, Boon N. Synthetic microbial ecosystems: an exciting tool to understand and apply microbial communities. *Environ Microbiol* 2014;16:1472–81.
- [19] Großkopf T, Soyer OS. Synthetic microbial communities. *Curr Opin Microbiol* 2014;18:72–7.
- [20] Qu Q, Zhang Z, Peijnenburg WJGM, Liu W, Lu T, et al. Rhizosphere microbiome assembly and its impact on plant growth. *J Agric Food Chem* 2020;68:5024–38.
- [21] Chodkowski JL, Shade A. A synthetic community system for probing microbial interactions driven by exometabolites. *mSystems* 2017;2:e00129–17.
- [22] Goldford JE, Lu N, Bajić D, Estrela S, Tikhonov M, et al. Emergent simplicity in microbial community assembly. *Science* 2018;361:469.
- [23] Bauer E, Thiele I. From network analysis to functional metabolic modeling of the human gut microbiota. *mSystems* 2018;3:e00209–17.
- [24] Dahabieh MS, Thevelein JM, Gibson B. Multimodal microorganism development: Integrating top-down biological engineering with bottom-up rational design. *Trends Biotechnol* 2020;38:241–53.
- [25] Peng XN, Gilmore SP, O'Malley MA. Microbial communities for bioprocessing: lessons learned from nature. *Curr Opin Chem Eng* 2016;14:103–9.
- [26] Elzinga J, van der Oost J, de Vos WM, Smidt H. The use of defined microbial communities to model host-microbe interactions in the human gut. *Microbiol Mol Biol Rev* 2019;83:e00054–18.
- [27] Liu F, Mao J, Kong W, Hua Q, Feng Y, et al. Interaction variability shapes succession of synthetic microbial ecosystems. *Nat Commun* 2020;11:309.
- [28] Guo X, Boedicker JQ. The contribution of high-order metabolic interactions to the global activity of a four-species microbial community. *PLoS Comput Biol* 2016;12:e1005079.
- [29] Venturilli OS, Carr AV, Fisher G, Hsu RH, Lau R, et al. Deciphering microbial interactions in synthetic human gut microbiome communities. *Mol Syst Biol* 2018;14:e8157.
- [30] Rodríguez Amor D, Dal Bello M. Bottom-up approaches to synthetic cooperation in microbial communities. *Life* 2019;9:22.
- [31] Faust K, Raes J. Microbial interactions: From networks to models. *Nat Rev Microbiol* 2012;10:538–50.
- [32] Xu X, Zarecki R, Medina S, Ofaim S, Liu X, et al. Modeling microbial communities from atrazine contaminated soils promotes the development of biostimulation solutions. *ISME J* 2019;13:494–508.
- [33] Friedman J, Higgins LM, Gore J. Community structure follows simple assembly rules in microbial microcosms. *Nat Ecol Evol* 2017;1:0109.

- [34] Stolyar S, Van Dien S, Hillesland KL, Pintel N, Lie TJ, et al. Metabolic modeling of a mutualistic microbial community. *Mol Syst Biol* 2007;3:92.
- [35] Ye C, Zou W, Xu N, Liu L. Metabolic model reconstruction and analysis of an artificial microbial ecosystem for vitamin C production. *J Biotechnol* 2014;182–183:61–7.
- [36] Ravikrishnan A, Blank LM, Srivastava S, Raman K. Investigating metabolic interactions in a microbial co-culture through integrated modelling and experiments. *Comput Struct Biotechnol J* 2020;18:1249–58.
- [37] Chang C-Y, Osborne ML, Bajic D, Sanchez A. Artificially selecting bacterial communities using propagule strategies. *Evolution* 2020;74:2392–403.
- [38] Schaefer AL, Lappala CR, Morlen RP, Pelletier DA, Lu T-Y-S, et al. LuxR- and LuxI-type quorum-sensing circuits are prevalent in members of the *Populus deltoides* microbiome. *Appl Environ Microbiol* 2013;79:5745.
- [39] Levy A, Salas Gonzalez I, Mittelviehhaus M, Clingenpeel S, Herrera Paredes S, et al. Genomic features of bacterial adaptation to plants. *Nat Genet* 2018;50:138–50.
- [40] Hasim S, Allison DP, Mendez B, Farmer AT, Pelletier DA, et al. Elucidating duramycin's bacterial selectivity and mode of action on the bacterial cell envelope. *Front Microbiol* 2018;9:219.
- [41] Neidhardt FC, Bloch PL, Smith DF. Culture medium for enterobacteria. *J Bacteriol* 1974;119:736.
- [42] Reasoner DJ, Geldreich EE. A new medium for the enumeration and subculture of bacteria from potable water. *Appl Environ Microbiol* 1985;49:1.
- [43] Cregger MA, Veach AM, Yang ZK, Crouch MJ, Vilgalys R, et al. The *Populus* holobiont: Dissecting the effects of plant niches and genotype on the microbiome. *Microbiome* 2018;6:31.
- [44] Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 2016;4:e2584.
- [45] Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 2011;17:10–2.
- [46] Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37:852–7.
- [47] Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;13:581–3.
- [48] Jagannath S, Ramachandran KB. Influence of competing metabolic processes on the molecular weight of hyaluronic acid synthesized by *Streptococcus zooepidemicus*. *Biochem Eng J* 2010;48:148–58.
- [49] Shiroda M, Pratt JR, Döpfer D, Wong ACL, Kaspar CW. RpoS impacts the lag phase of *Salmonella enterica* during osmotic stress. *FEMS Microbiol Lett* 2014;357:195–200.
- [50] Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, et al. KBase: The United States department of energy systems biology knowledgebase. *Nat Biotechnol* 2018;36:566–9.
- [51] Henry CS, Bernstein HC, Weisenhorn P, Taylor RC, Lee J-Y, et al. Microbial community metabolic modeling: a community data-driven network reconstruction. *J Cell Physiol* 2016;231:2339–45.
- [52] Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, et al. The RAST server: rapid annotations using subsystems technology. *BMC Genom* 2008;9:75.
- [53] Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, et al. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res* 2014;42:D206–14.
- [54] Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 2015;5:8365.
- [55] Batth TS, Tollenaere MAX, Rütther PL, Gonzalez-Franquesa A, Prabhakar BS, et al. Protein aggregation capture on microparticles enables multipurpose proteomics sample preparation. *Mol Cell Proteom* 2019;18:1027–35.
- [56] Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 1994;5:976–89.
- [57] Käll L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat Methods* 2007;4:923–5.
- [58] Polpitiya AD, Qian W-J, Jaitly N, Petyuk VA, Adkins JN, et al. DANTE: a statistical tool for quantitative analysis of -omics data. *Bioinformatics* 2008;24:1556–8.
- [59] Kleiner M, Thorson E, Sharp CE, Dong X, Liu D, et al. Assessing species biomass contributions in microbial communities via metaproteomics. *Nat Commun* 2017;8:1558.
- [60] Gittel NR, Castro HF, Kerley M, Yang Z, Pelletier DA, et al. Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types. *Appl Environ Microbiol* 2011;77:5934.
- [61] Germerodt S, Bohl K, Lück A, Pande S, Schröter A, et al. Pervasive selection for cooperative cross-feeding in bacterial communities. *PLOS Comput Biol* 2016;12:e1004986.
- [62] Pollock J, Glendinning L, Wisedchanwet T, Watson M. The madness of microbiome: attempting to find consensus “best practice” for 16S microbiome studies. *Appl Environ Microbiol* 2018;84:e02627–17.
- [63] Sipos R, Székely AJ, Palatinszky M, Révész S, Márialigeti K, et al. Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targeting bacterial community analysis. *FEMS Microbiol Ecol* 2007;60:341–50.
- [64] Campanaro S, Treu L, Kougias PG, Zhu X, Angelidaki I. Taxonomy of anaerobic digestion microbiome reveals biases associated with the applied high throughput sequencing strategies. *Sci Rep* 2018;8:1926.
- [65] Neurohr GE, Amon A. Relevance and regulation of cell density. *Trends Cell Biol* 2020;30:213–25.
- [66] Milo R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* 2013;35:1050–5.
- [67] Smith CCR, Snowberg LK, Gregory Caporaso J, Knight R, Bolnick DI. Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *ISME J* 2015;9:2515–26.
- [68] Saxer G, Doebeli M, Travisano M. Spatial structure leads to ecological breakdown and loss of diversity. *Proc R Soc B Biol Sci* 2009;276:2065–70.
- [69] Benomar S, Ranava D, Cárdenas ML, Trably E, Raftafi Y, et al. Nutritional stress induces exchange of cell material and energetic coupling between bacterial species. *Nat Commun* 2015;6:6283.
- [70] Zengler K, Zaramela LS. The social network of microorganisms — how autotrophies shape complex communities. *Nat Rev Microbiol* 2018;16:383–90.
- [71] Calatayud J, Andivia E, Escudero A, Melián CJ, Bernardo-Madrid R, et al. Positive associations among rare species and their persistence in ecological assemblages. *Nat Ecol Evol* 2020;4:40–5.
- [72] Orth JD, Thiele I, Palsson BØ. What is flux balance analysis?. *Nat Biotechnol* 2010;28:245–8.
- [73] Bertrand RL. Lag phase is a dynamic, organized, adaptive, and evolvable period that prepares bacteria for cell division. *J Bacteriol* 2019;201:e00697–18.
- [74] Lee SA, Park J, Chu B, Kim JM, Joa J-H, et al. Comparative analysis of bacterial diversity in the rhizosphere of tomato by culture-dependent and-independent approaches. *J Microbiol* 2016;54:823–31.
- [75] Kim J, Kim KH. Effects of minimal media vs. complex media on the metabolite profiles of *Escherichia coli* and *Saccharomyces cerevisiae*. *Process Biochem* 2017;57:64–71.
- [76] Whitham TG, Young WP, Martinsen GD, Gehring CA, Schweitzer JA, et al. Community and ecosystem genetics: a consequence of the extended phenotype. *Ecology* 2003;84:559–73.
- [77] Langenheder S, Bulling MT, Prosser JL, Solan M. Role of functionally dominant species in varying environmental regimes: evidence for the performance-enhancing effect of biodiversity. *BMC Ecol* 2012;12:14.
- [78] Abreu CI, Friedman J, Andersen Woltz VL, Gore J. Mortality causes universal changes in microbial community composition. *Nat Commun* 2019;10:2120.
- [79] Concepción-Acevedo J, Weiss HN, Chaudhry WN, Levin BR. Malthusian parameters as estimators of the fitness of microbes: a cautionary tale about the low side of high throughput. *PLoS ONE* 2015;10:e0126915.
- [80] Pekkonen M, Ketola T, Laakso JT. Resource availability and competition shape the evolution of survival and growth ability in a bacterial community. *PLoS ONE* 2013;8:e76471.
- [81] Bittleston LS, Gralka M, Leventhal GE, Mizrahi I, Cordero OX. Context-dependent dynamics lead to the assembly of functionally distinct microbial communities. *Nat Commun* 2020;11:1440.
- [82] Lipson DA. The complex relationship between microbial growth rate and yield and its implications for ecosystem processes. *Front Microbiol* 2015;6:615.
- [83] Zarecki R, Oberhardt MA, Reshef L, Gophna U, Ruppin E. A novel nutritional predictor links microbial fastidiousness with lowered ubiquity, growth rate, and cooperativeness. *PLoS Comput Biol* 2014;10:e1003726.
- [84] Goelzer A, Fromion V. Bacterial growth rate reflects a bottleneck in resource allocation. *Biochim Biophys Acta Gen Subj* 2011;1810:978–88.
- [85] Franzosa EA, Hsu T, Sirota-Madi A, Shafquat A, Abu-Ali G, et al. Sequencing and beyond: integrating molecular ‘omics’ for microbial community profiling. *Nat Rev Microbiol* 2015;13:360–72.
- [86] Zomorrodi AR, Segrè D. Synthetic ecology of microbes: mathematical models and applications. *J Mol Biol* 2016;428:837–61.
- [87] Sousa A, Frazão N, Ramiro RS, Gordo I. Evolution of commensal bacteria in the intestinal tract of mice. *Curr Opin Microbiol* 2017;38:114–21.
- [88] Joseph TA, Shenhav L, Xavier JB, Halperin E, Pe'er I. Compositional lotka-volterra describes microbial dynamics in the simplex. *PLoS Comput Biol* 2020;16:e1007917.
- [89] Song H-K, Song W, Kim M, Tripathi BM, Kim H, et al. Bacterial strategies along nutrient and time gradients, revealed by metagenomic analysis of laboratory microcosms. *FEMS Microbiol Ecol* 2017;93.
- [90] Widder S, Allen RJ, Pfeiffer T, Curtis TP, Wiuf C, et al. Challenges in microbial ecology: building predictive understanding of community function and dynamics. *ISME J* 2016;10:2557–68.
- [91] Louca S, Doebeli M. Calibration and analysis of genome-based models for microbial ecology. *eLife* 2015;4:e08208.
- [92] Cardona C, Weisenhorn P, Henry C, Gilbert JA. Network-based metabolic analysis and microbial community modeling. *Curr Opin Microbiol* 2016;31:124–31.
- [93] Bae S, Fleet GH, Heard GM. Occurrence and significance of *Bacillus thuringiensis* on wine grapes. *Int J Food Microbiol* 2004;94:301–12.