Keystone species determine the productivity of synthetic microbial biofilm communities

Xinli Sun[1, 2], Jiyu Xie[1], Daoyue Zheng[1], Wei Wang[1], Riyan Xia[1], Ruifu Zhang[1], Ákos T. Kovács[2*], Zhihui Xu[1*], Qirong Shen[1]

1 Jiangsu Provincial Key Lab of Solid Organic Waste Utilization, Jiangsu Collaborative Innovation Center of Solid Organic Wastes, Educational Ministry Engineering Center of Resource-Saving Fertilizers, The Key Laboratory of Plant Immunity, Nanjing Agricultural University, 210095 Nanjing, Jiangsu, Peoples R China

2 Bacterial Interactions and Evolution Group, DTU Bioengineering, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

Xinli Sun and Jiyu Xie contributed equally to this work.

*Corresponding authors: Zhihui Xu and Ákos T. Kovács

E-mail: xzh2068@njau.edu.cn, atkovacs@dtu.dk

Abstract

Microbes typically reside in multi-species communities, whose interactions have considerable impacts on the robustness and functionality of such communities. To manage microbial communities, it is essential to understand the factors driving their assemblage and maintenance. Even though the community composition could be easily assessed, interspecies interactions during community establishment remain poorly understood. Here, we combined co-occurrence network analysis with quantitative PCR to examine the importance of each species within synthetic communities (SynComs) of pellicle biofilms. Genome-scale metabolic models and *in vivo* experiments indicated that the biomass of SynComs was primarily affected by keystone species that are acting either as metabolic facilitators or as competitors. Our study suggests that a combination of co-occurrence network analysis and metabolic modeling could explain the importance of keystone species in SynComs.

Introduction

In natural microbiome systems, hundreds or thousands of species occupy and interact with each other, as well as the environment around them. To deal with this complexity, we can utilize synthetic communities (SynComs) to identify common principles governing and structuring microbiomes (Bengtsson-Palme, 2020). Several SynComs with moderate complexity and high controllability have been developed to represent different natural environments (Blasche et al., 2017; Chan et al., 2017;

Fu et al., 2020; Lebeis et al., 2015; Weiss et al., 2021). Such knowledge is essential to disentangle complex community interactions and to provide powerful tools for biotechnological, medical, or agricultural purposes (Abram, 2015; Cavaliere et al., 2017; Cho and Blaser, 2012; Fitzpatrick et al., 2020; Gómez-Godínez et al., 2021). To harness this potential, it is essential to preserve both the microenvironments in which these beneficial processes can take place, as well as the microbial populations that contribute to these processes. Based on the study of SynComs, researchers have found that the assemblage and robustness of communities can be affected by several factors including pH (Ortiz et al., 2021), spatial distribution (Liu et al., 2019), initial abundance (Gao et al., 2020, 2019), niche specificity (Estrela et al., 2021), nutrient availability (Ratzke et al., 2020), keystone species (Niu et al., 2017) and so on. Nevertheless, many questions remain about how diversity is maintained, and ultimately, how the community function can be preserved and controlled.

Multiple species are widely distributed in nature as surface-associated communities also referred to as biofilms (Sadiq et al., 2021). The biofilm formation process consists of the planktonic stage, initial attachment to a surface, the building of micro-aggregates, biofilm maturation, and biofilm dispersal (Lee and Yoon, 2017). These multi-species biofilm communities assemble as a result of interspecies interactions, which determine not only the composition and distribution but also their biological function. Compared with solitary bacteria, species residing in multispecies biofilms can gain fitness advantages, including enhanced biomass production (Ren et al., 2015), stress tolerance (Lee et al., 2014; Raghupathi et al., 2018), and access to complex nutrients (Breugelmans et al., 2008; Nielsen et al., 2000) through cooperative interactions. In contrast, ecological competition is more pervasive in multispecies biofilms (Foster and Bell, 2012; Oliveira et al., 2015). Bacteria actively inhibit competitors by producing toxins, biosurfactants, and antimicrobials (Rendueles and Ghigo, 2012), or indirectly impair the fitness of neighbors through nutrient scavenging. The spatial structuring property of biofilms supports the intermixing of cooperating species and the partitioning of niches by competing species. Understanding how these interactions affect the formation of multi-species biofilms is therefore important for their utilization or inhibition.

Despite a wealth of information on microbiota composition gathered from sequencing techniques, the knowledge of microbial interactions that occur in the microbiota is severely lacking. Building a co-occurrence network from high-throughput sequence data is a common method of deriving hypotheses from these massive data (Faust et al., 2012). For instance, it has been applied to link taxa to biological functions of interest (Wei et al., 2019), to identify potential biotic interactions (Durán et al., 2018), and to explore habitat differentiation (Barberán et al., 2012). Another emerging method for exploring microbial interactions is genome-scale metabolic modeling, which can provide insights into metabolic interaction potential and metabolic resource overlap in multi-species communities (Zelezniak et al., 2015; Zorrilla et al., 2021). Microbiome network analysis and metabolic modeling could guide the study of SynComs for animal and plant health (Heinken et al.,

2021; Poudel et al., 2016). However, the ecological relevance of predicted interactions remains poorly understood (Faust, 2021). The connectedness and strength of positive or negative interactions are not experimentally verified. The main challenge of studying microbial interactions in SynComs is fast and accurate quantification of the abundance of each species. This can be achieved by 16S amplicon sequencing, colony-forming unit counting, fluorescent labeling, and quantitative PCR. Both of the first two approaches are simple to employ, but they can be time-consuming. While the first two approaches are more convenient but time-consuming, the latter two approaches are more efficient but require extensive initial time and effort.

In this study, we constructed SynComs using isolates from a rhizosphere and predicted their interaction by analyzing co-occurrence networks. The network analysis was validated by quantitative assays. The potential metabolic interactions were further investigated through experiments and metabolic modeling. The results suggested that keystone species determine community productivity through metabolic exchanges or resource competition. We propose that our study could inform the rational design of synthetic communities.

Results

C. rhizoplanae and *P. eucrina* were predicted to be keystone species in the eleven-species SynCom

In a previous study, we isolated 267 bacteria from the cucumber rhizosphere (Sun et al., 2021). Eleven isolates that co-exist in the natural environment were chosen to create a synthetic biofilm community. These isolates originate from three different phyla: Firmicutes, Proteobacteria, and Bacteroidetes. We evaluated the stability of this SynCom by observing the compositional change of individual members using 16S rRNA gene amplicon sequencing. Several time points were selected to represent the development of a biofilm: initial aggregation (2d), establishment (4d), maturation (6d), and dispersal (8d). Chr (*Chryseobacterium rhizoplanae*) was identified to be the predominant species throughout the biofilm development (Fig. 1A). Three species Ach (*Achromobacter denitrificans*), Pxa (*Pseudoxanthomonas japonensis*), and Ste (*Stenotrophomonas maltophilia*) rapidly declined and could not establish themselves in the community. Network co-occurrence analysis was employed to examine the correlations among the species during succession. The correlations among the species varied from less to more (Fig. 1B). Pan (*Pantoea eucrina*) was consistently positively correlated with other species, while Chr was consistently negatively correlated with other species. These two species were predicted to act as hub taxa in this multi-species consortium. In the subsequent experiments, we examined the interactions of six of the most abundant species, including Pan, Chr, Com (*Comamonas odontotermitis*), Aci (*Acinetobacter baumannii*), Ent (*Enterobacter bugandensis*), and Pse (*Pseudomonas stutzeri*). Other species were

103 excluded either due to their low abundance at the biofilm maturation stage, or due to the lack of
104 correlation with the other species.

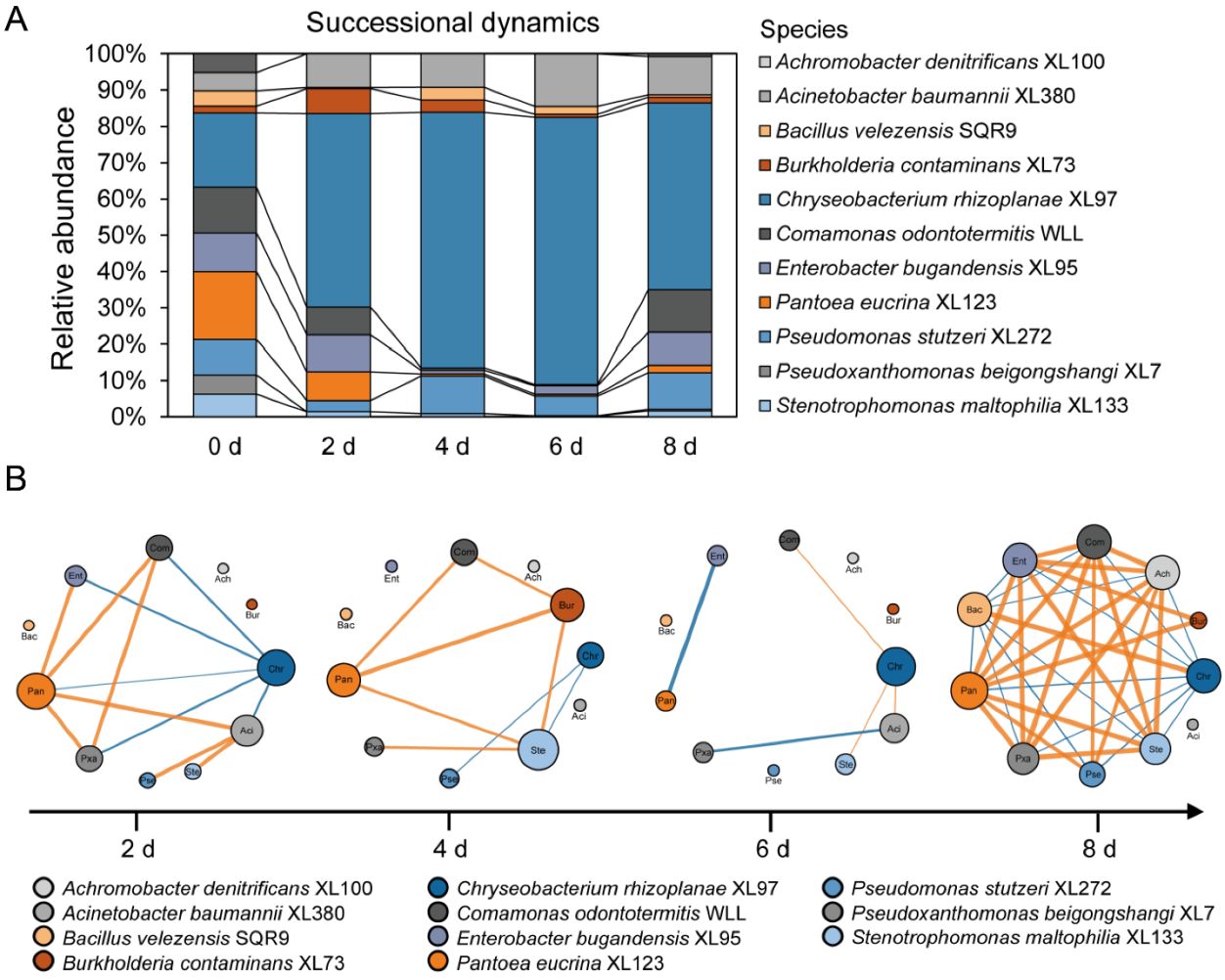
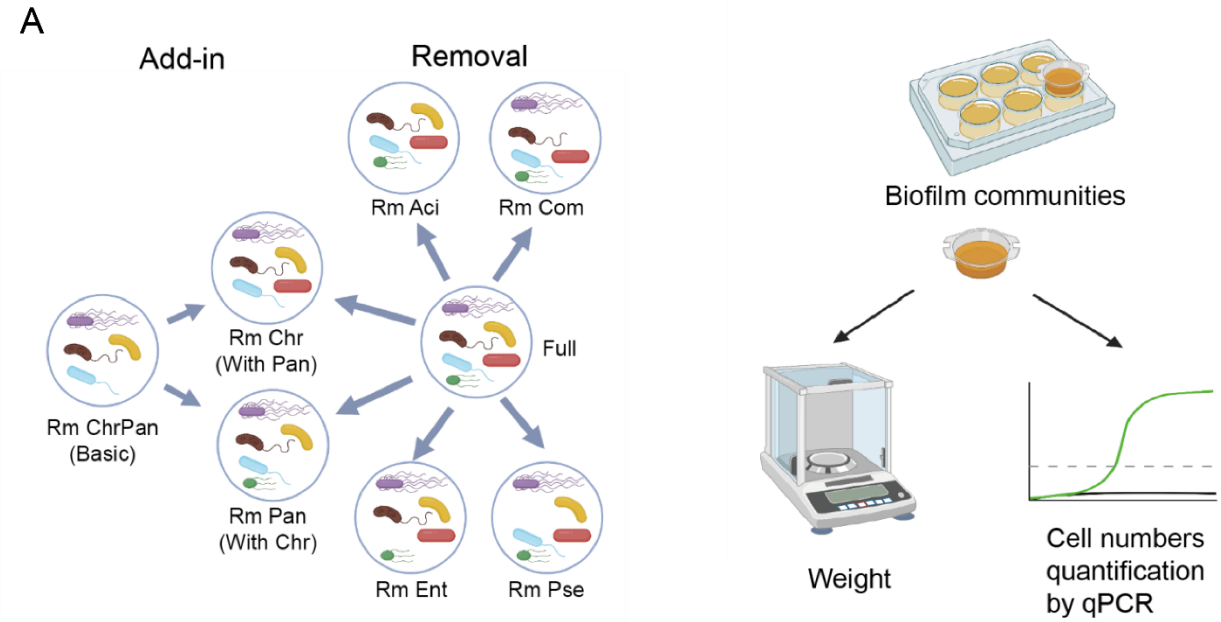


FIG 1 Dynamic changes of the 11-species consortium composition in biofilm. (**A**) The relative abundance of species determined by 16S rRNA gene amplicon sequencing method (n=8). (**B**) Co-occurrence of bacterial clades within biofilm communities. Each node represents a bacterial species, node size is proportional to betweenness centrality. Line width indicates the interactive strength of interaction, and color indicates the sign of association (orange indicates positive, blue indicates negative).
105

106 Productivity of multi-species biofilm was affected by key members

107 To test the central role that Pan and Chr play in biofilm development, we evaluated the relative
108 productivity of different combinations of the six strains. “Add-in” and “Removal” strategies were
109 applied (Fig 2A). According to the “Add-in” strategy, the “Basic community” consisted of four
110 members (Com, Aci, Ent, and Pse), upon which we added the predicted positive or negative
111 member to the community and compared the total biomass. Based on our hypothesis, we
112 anticipated that supplementation of the positive member would increase total biomass, while we
113 anticipated the opposite for the negative member. On the other hand, the “Removal” strategy
114 included all the six species as a full community, then one isolate was dropped out to obtain the five-
115 species reduced communities. Biofilm productivity was assessed in two ways: fresh weight and total
116 cell numbers (Fig 2A).

Compared to the "Full community", “Rm Chr” and “Rm Com” resulted in a higher biofilm weight, whereas “Rm Ent” and “Rm Pan” resulted in a lower biofilm weight (Fig 2B). This indicated that these four species are important to maintain the productivity of this synthetic community. However, the contribution of each species to the population productivity could not be determined using this approach. The previously used amplicon sequencing method was relatively slow and could only provide information on the relative abundance. The main challenge in determining the composition is to establish a fast and accurate method for quantifying the absolute abundance of each species. To achieve this goal, we performed a genome-scale comparison and identified strain-specific genes in each species. We designed multiple primers to amplify these genes and tested their specificity against other species. Consequently, strain-specific primers were developed to selectively amplify only one isolate from the six-members community (Table 1). Using these primers, we were able to determine the contribution of each isolate to the community productivity. “Rm Chr” reached the highest number of total cell numbers, while “Rm Pan” greatly decreased the total cell numbers (Fig 2C). Taken together, we identified Pan and Chr as the keystone positive and negative species in terms of community productivity, respectively.



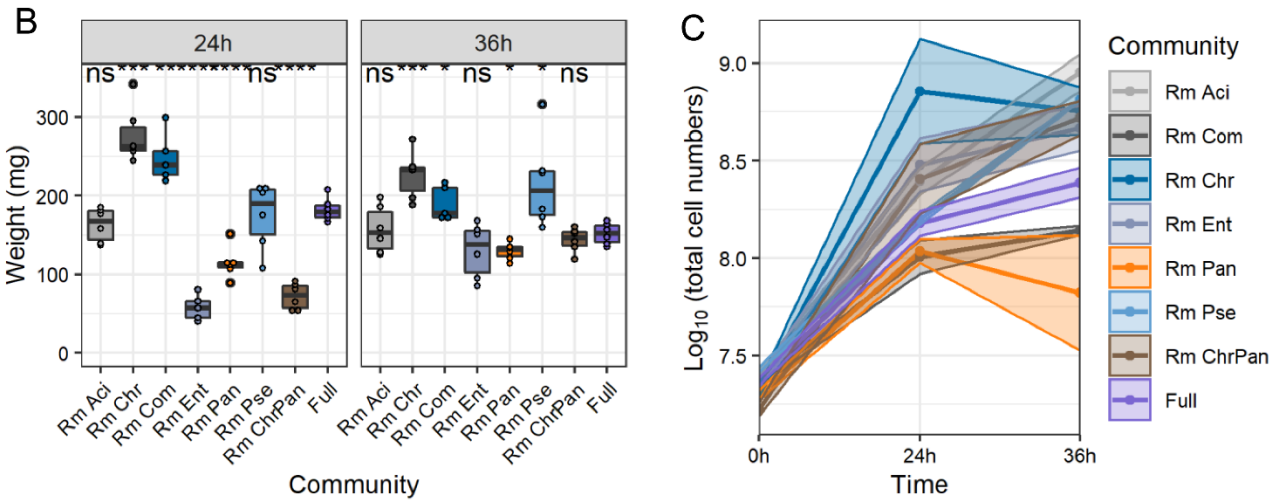


FIG 2 Productivity of the reduced SynComs. (**A**) Schematic diagram of experimental setup. Add in: positive or negative species were added in the four-species community (Rm ChrPan) to create five-species communities. Removal: each isolate was dropped out from the six-species community (Full) to create five-species communities. Biofilm biomass were quantified by fresh weight and cell numbers were quantified by qPCR. Diagrams were generated in Biorender. (**B**) Biofilm biomass quantified by fresh weight (n=6). The "Full community" was set as the control group. Asterisks indicate statistically significant ($p < 0.01$) according to unpaired student's *t* test via R. (**C**) Total cell numbers in the communities (n=6). Color strip indicates standard deviation. All represents the full six-species communities, Rm Aci, Rm Chry, Rm Com, Rm Ent, Rm Pan, Rm Pse represent the five-species communities resulting from the removal of *Acinetobacter baumannii* XL380, *Chryseobacterium rhizoplanae* XL97, *Comamonas odontotermitis* WLL, *Enterobacter ludwigii* XL95, *Pantoea eucrina* XL123 and *Pseudomonas stutzeri* XL272, respectively. Rm ChrPan represents the four-species community resulting from the removal of *C. rhizoplanae* and *P. eucrina*.

C. rhizoplanae and *P. eucrina* were validated to be keystone species in the reduced community

The compositions of different communities were clustered based on the Bray-Curtis dissimilarity index (Fig 3A-C). After 24 hours, the communities clustered into two main groups: "Full community", "Basic community" (Rm ChrPan), "Rm Pan", and "Rm Chr" communities clustered together, whereas the other communities were grouped in one branch. The upper branch was dominated by Ent, the lower branch by Pan. Based on the hierarchical clustering, Pan or Chr played deterministic roles in species composition by influencing the relative abundance of Ent. Additionally, the random forest analysis revealed that Pan played a major role in determining the differences of community compositions (Fig 3D). After 36 hours, the communities were divided into two branches without any apparent disciplinarian. We found that the proportion of Pan cells was highest in the "Rm Ent" community, while the proportion of Ent cells was highest in the "Rm Pan" community (Fig 3C). Furthermore, random forest analysis revealed that Chr was the keystone species determining the compositional differences (Fig 3E).

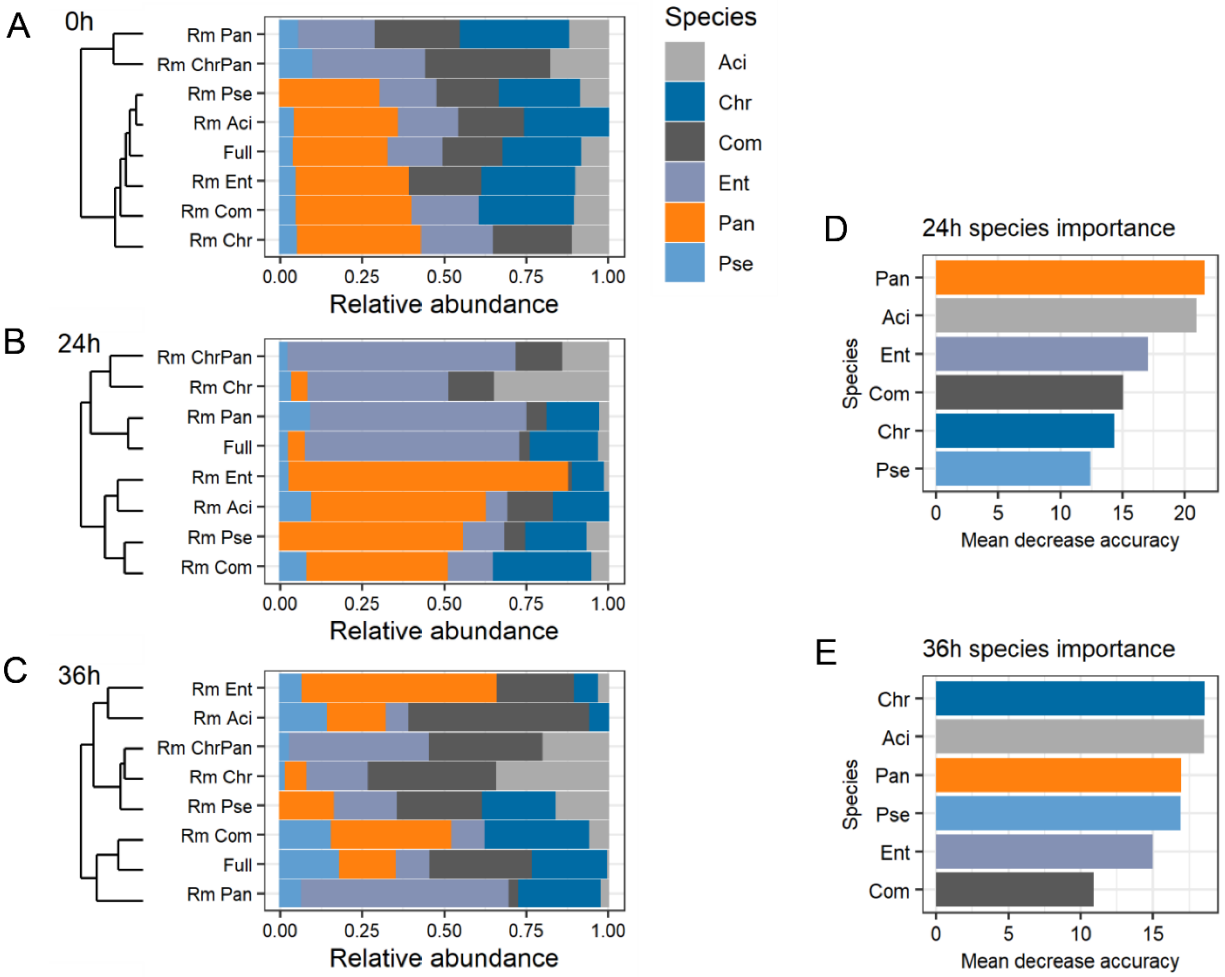


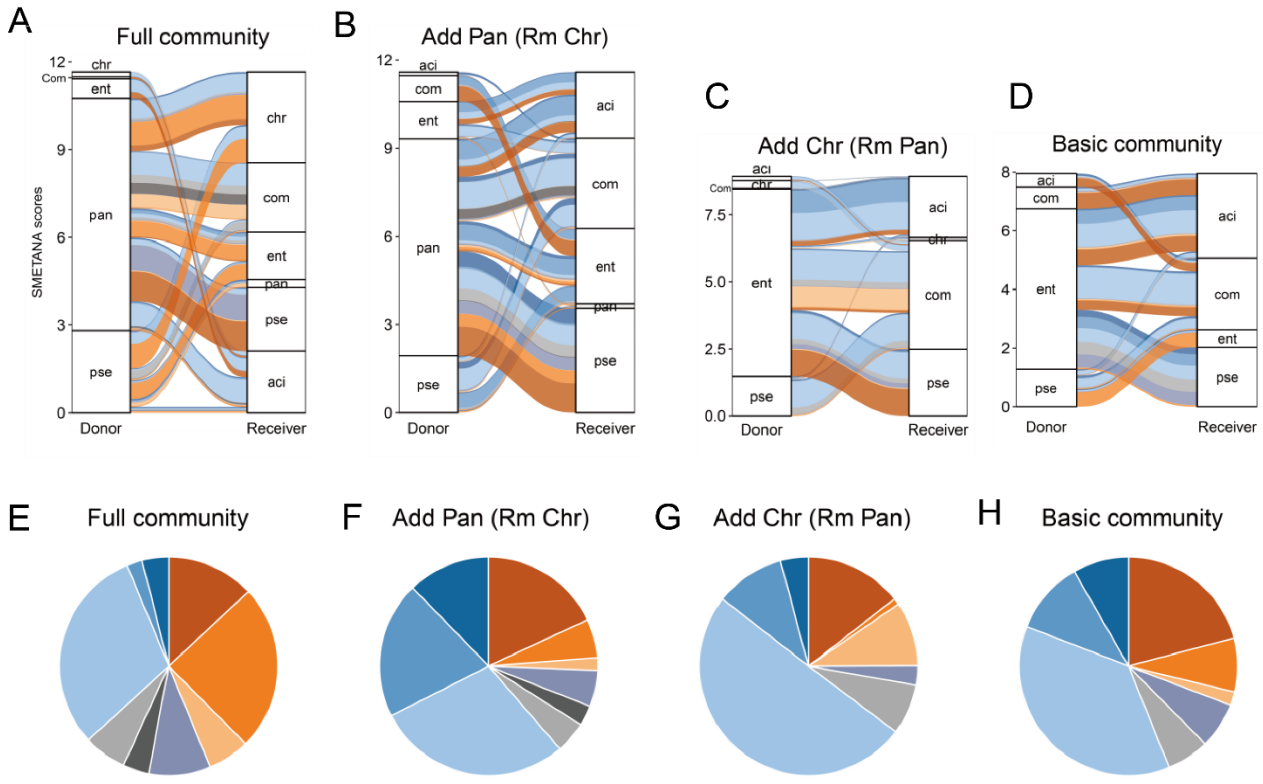
FIG 3 Species composition and importance. (**A-C**) Hierarchical clustering analysis of the synthetic community composition (n=6). Samples include a six-species communities (Full), a four-species community (Rm ChrPan) and six five-species communities (Rm communities). Rm Aci, Rm Chr, Rm Com, Rm Ent, Rm Pan, Rm Pse represent the five-species communities resulting from the removal of *Acinetobacter baumannii* XL380, *Chryseobacterium rhizoplanae* XL97, *Comamonas odontotermitis* WLL, *Enterobacter ludwigii* XL95, *Pantoea eucrina* XL123 and *Pseudomonas stutzeri* XL272, respectively. The samples were collected at the initial inoculation (A), 24h (B) and 36h (C). (**D-E**) Species importance in the SynComs as predicted by the random forest analysis.

Metabolic facilitation explained the positive interaction

To gain insights into the metabolic interaction potential of the SynCom members, we reconstructed genome-scale metabolic models, using the M9 glucose minimal medium as the input medium. We derived the likely exchanged metabolites across communities and the strength of metabolic coupling - SMETANA score (Zelezniak et al., 2015). The “Full community” and the “Add Pan” community displayed higher cross-feeding potential than the “Add Chr community” and the “Basic community” (Fig 4A-D). Pan was the principal donor in the communities with Pan, whereas Ent was the principal donor in the communities without Pan. This result explained the dominance of these two species in the SynComs. In contrast, adding Chr to the SynComs had little effect on the global metabolic exchanges. Chr acted more as a receiver rather than a donor in the “Full community” (Fig 4A). In addition, the add-in or removal of keystone species also affected the proportion of exchanged compounds (Fig 4E-H). Amino acids and phosphates were the major categories of

metabolites exchanged across communities. When the positive member Pan was removed, we observed higher reliance on amino acid cross-feeding and lower dependence on organoheterocyclic compounds exchange. Moreover, we highlighted the interactions involving Pan as the donor in the "Full community" (Fig 4I) and in the "Add Pan" (Rm Chr) community (Fig 4J). We observed that the metabolic contribution of Pan benefited all community members.

Growth assays were used to determine the accuracy of the metabolic models. Firstly, all species except Chr were able to grow individually in the M9 glucose minimal medium (data not shown), thus five of them have the potential to serve as donors. Secondly, growth in the spent culture medium was utilized to determine metabolic interaction potential. Each isolate was cultured in the M9 glucose medium until the glucose was undetectable. We obtained sterile spent medium and used it to cultivate every other member of the community and themselves. The maximum growing capacity was determined and defined as growth facilitation (Fig 4K). Results showed that the spent medium from Pan and Ent could support the growth of all other isolates. Pse's metabolic by-products were also capable of supporting the growth of three isolates. Aci, Com, and Chr acted only as receivers but not donors. These results fully supported the metabolic models, suggesting Pan and Ent contributed to the overall productivity by cross-feeding, whereas Chr acted as a cheater.



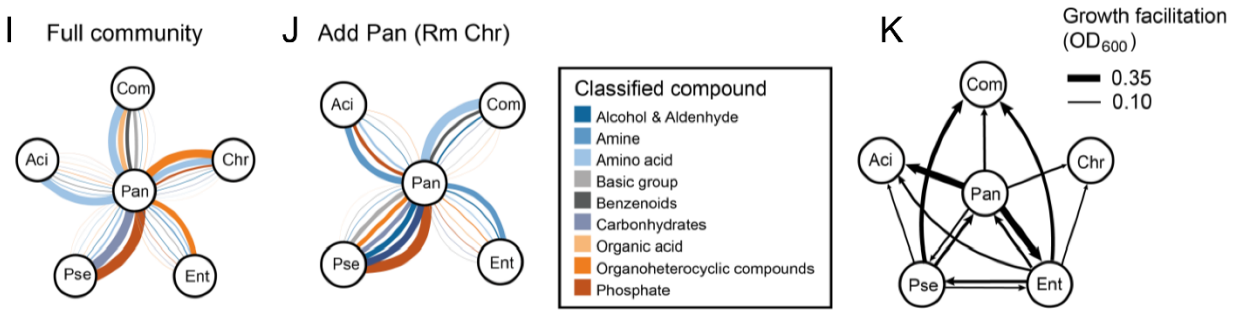


FIG 4 Metabolic facilitation. (**A-J**) Metabolic interaction potential simulated by genome-scale metabolic modelling. (**A-D**) Alluvial diagrams showing compounds exchanged within communities. Thickness of lines are SMETANA scores. (**E-H**) Pie charts showing the proportion of compounds exchanged within communities. (**I-J**) Flower plots showing interactions involving positive species *P. eucrina* XL123 (centered in each subgraph) as a donor across communities. Thickness of lines are proportional to magnitude of SMETANA score. Colors of all plots are the classification of compounds. (**K**) Growth facilitation assessed by growth in spent medium. The spent medium was prepared by growing the corresponding species in M9 medium with 1% glucose till the glucose was under detection then filter sterilized. The species were grown in the filtered spent medium for 24 h and the growth were measured as the OD_{600}. Line width indicates the OD_{600}.

Resource competition explained the negative interaction

Higher-order interactions may further be explained by metabolic competition. Using metabolic modeling, the metabolic resource overlap was simulated (Fig 5A). We observed high metabolic resource overlap across all communities, indicating intense resource competition among SynCom members. The addition of the negative member Chr resulted in the highest metabolic resource overlap while adding the positive member Pan resulted in a reduction of resource competition. To test the possibility of direct competition, we performed spot-on-lawn and pair-wise spot assays on TSB agar plates (Fig 5B-C). Both Aci and Ent completely inhibited the growth of Com. Similarly, Chr and Ent inhibited Pse growth. However, these inhibitions were not caused by direct antagonism as no inhibition zones were observed in the respective spot-on-lawn assay plates (Fig 5B, red square). In light of these results, no clear boundary was observed between two strains when they were spotted adjacent to one another (Fig 5C). Interestingly, Pan and Ent could change the colony color of Chr from light yellow to white (Fig 5B, yellow square). We also assessed the carbon source metabolic ability by high-throughput phenotypic microarrays (Fig 5D). Ent and Pan were generalists that can utilize a wide variety of carbon sources. Other community members were specialized in using amino acids and organic acids as carbon sources while having a limited ability to utilize nucleotides, sugar alcohols, and sugars. Collectively, resource competition was widespread among the SynCom members.

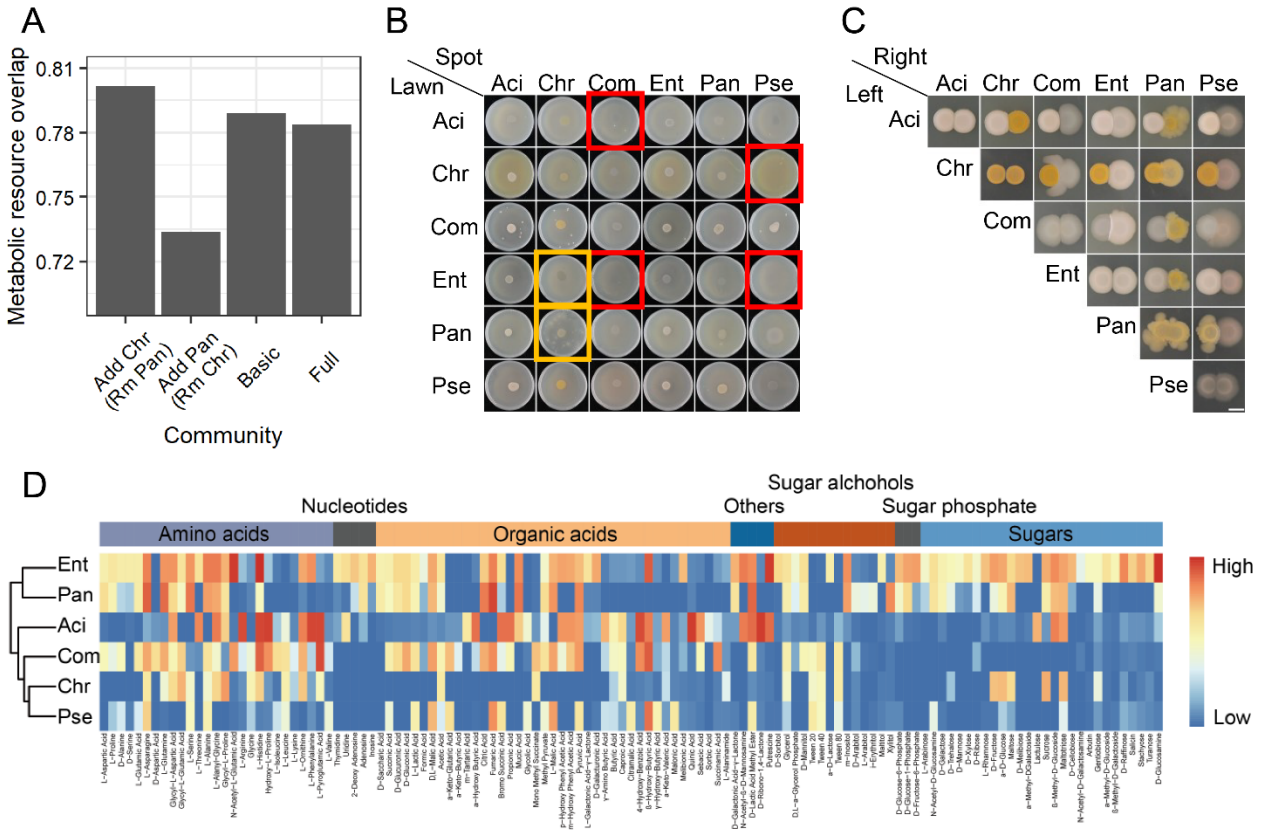


FIG 5 Resource competition. (**A**) Bar chart showing the metabolic resource overlap simulated by the metabolic modeling. (**B**) Spot-on-lawn assays. The lawn species were spread on TSB agar plates at an OD_{600} of 0.02 and dried, 5 µl of the spot species were spotted on the center at an OD_{600} of 0.4. The plate diameter is 6 cm. Photos were taken after 48 h incubation at 30°C. (**C**) Pair-wise spot assays. 5 µl of the two species were plot next to each other at an OD_{600} of 1. Scale bar represents 3 mm. Photos were taken after 48 h incubation at 30°C. (**D**) Carbon source metabolic ability measured by phenotype microarrays.

Discussion

A key concern in SynCom research is to understand how microbial interactions affect community composition and productivity. In our study, we evaluated the contribution of each individual to community productivity and revealed the metabolic interactions. Our SynCom system identified Pan (*P. eucrina*) and *Chr* (*C. rhizoplanae*) as important drivers of community interaction networks through metabolic cross-feeding and resource competition.

Manipulating microbial consortia has a variety of applications; however, attempts to engineer microbial consortia often fail to achieve the expected results owing to the unexpected effects of interactions within the community. Although the practical application of this SynCom is not yet clear, accurate prediction of the species composition is the first step toward manipulating microbial consortia (Widder et al., 2016). In this study, we assessed the SynCom composition using the amplicon sequencing method and qPCR. In the initial eleven-species biofilm community, Chr was the dominant species across all the timepoints. In the reduced six-species biofilm community, however, Ent and Pan became the most abundant species. These findings indicated that other species, although less abundant or not associated with community members, also affect the final

composition of the SynCom. This result was consistent with a previous study that showed rare species were indispensable to the synergistic effects in multi-species biofilms (Ren et al., 2015). Furthermore, higher-order interactions only occur in the presence of additional species, emphasizing the importance of studying microbial interactions in the context of communities (Friedman et al., 2017).

Resource competition plays a major role in shaping bacterial communities, while metabolic exchanges promote group survival (Goldford et al., 2018; Zelezniak et al., 2015). In the present study, we simulated the metabolic interactions using metabolic models and assessed their predictability using phenotypic assays. Specifically, communities with negative species exhibited higher resource competition, whereas communities with positive species displayed greater metabolite exchange. These results indicate that metabolic interactions play a key role in determining the composition of communities. Other mechanisms, such as niche partitioning created by spatial and temporal heterogeneity, and trade-offs between nutrient acquisition and environmental tolerance, can also promote coexistence between species (Louca et al., 2018). Our research provides insights into how metabolic competition and cooperation simultaneously shape the community composition. The methodology of network co-occurrence analysis combined with qPCR quantification, metabolic modeling, and pair-wise interaction can be applied to various SynComs studies. Ultimately, such studies should translate to a deeper understanding of how microbial communities behave in their native environments, and this knowledge may be applied to wastewater treatment, disease suppression, and crop yield enhancement.

Materials and methods

Strains and growth condition

As shown in Table 1, eleven bacterial isolates were selected from the rhizosphere of cucumber plants by (Sun et al., 2021) were selected for this study.

Table 1. Strains and abbreviations

Strain	Abbreviation
Achromobacter denitrificans XL100	Ach
Acinetobacter baumannii XL380	Aci
Bacillus velenzensis SQR9	Bac
Burkholderia contaminans XL73	Bur
Chryseobacterium rhizoplanae XL97	Chr
Comamonas odontotermitis WLL	Com
Enterobacter bugandensis	Ent
Pantoea eucrina XL123	Pan

Pseudomonas stutzeri XL272	Pse
Pseudoxanthomonas japonensis XL7	Pxa
Stenotrophomonas maltophilia XL133	Ste

237

238 The start inoculum was prepared by mixing equal volumes of species at an optical density at 600
239 nm (OD_{600}) of 1. The eleven-species biofilm was cultivated by mixing 4ml of the mixed-species
240 community (1%) with 400 ml of tryptic soy broth (TSB) and incubating at 30°C for two to eight days.
241 The four-, five- and six-species biofilms were grown in 6-well microtiter plates (VWR) insert with 100
242 µm sterile nylon mesh cell strainers (Biologix Cat #15-1100). 10 ml of TSB liquid medium and 100 µl
243 of start inoculum were added. The plates were incubated for 24 h at 30 °C to allow the biofilm to
244 grow on top of the nylon mesh cell strainer.

245 ## Whole-genome sequencing, annotation, and comparison

246 The whole genomes of the six species were sequenced by different companies at different times.
247 The genomes of Aci and Com were sequenced using a combination of PacBio RS II and Illumina
248 HiSeq 4000 sequencing platforms. The genome of Aci was sequenced at the Beijing Genomics
249 Institute (BGI, Shenzhen, China). The Illumina data were used to evaluate the complexity of the
250 genome. Raw sequencing data and the assembled genome have been deposited to the National
251 Center for Biotechnology Information (NCBI) under the BioProject accession number PRJNA593376.
252 The genome of Com was sequenced at Majorbio Bio-Pharm Technology Co., Ltd. Raw sequencing
253 data and the assembled genome have been deposited to the NCBI under the BioProject accession
254 number PRJNA762695.

255 The genomes of Chr, Ent, and Pan were sequenced using PacBio Sequel platform and Illumina
256 NovaSeq PE150 at the Beijing Novogene Bioinformatics Technology Co., Ltd. Raw sequencing data
257 and the assembled genomes have been deposited to the NCBI under the BioProject accession
258 number PRJNA721858, PRJNA761942, and PRJNA762676. Pse was sequenced by (Sun et al.,
259 2021). Genomes were automatically annotated by NCBI PGAP. Genome comparison was
260 performed using Roary (Page et al., 2015). This generated core and accessory gene tables.

261 ## Eleven-species SynCom quantification and analysis

262 Biofilms formed at the air-liquid interface were collected on days 2, 4, 6, and 8. Each time point had
263 eight biological replicates. Genomic DNA of the biofilm samples was extracted using an E.Z.N.A.
264 Bacterial DNA Kit (Omega Bio-tek, Inc.) following the manufacturer's instructions. Universal primers
265 targeting the V3-V4 regions of the 16S rRNA gene were used to construct the DNA library for
266 sequencing. Paired-end sequencing of bacterial amplicons was performed on the Illumina MiSeq
267 instrument (300 bp paired-end reads). Raw sequencing data have been deposited to the NCBI SRA
268 database under BioProject accession number PRJNA739098. Reads were processed using the

USEARCH pipeline. The paired-end reads were merged using the “fastq_mergepairs” command. High-quality sequences were then selected using the “fastq_filter” command and dereplicated using the “derep_fulllength” command. The singletons were removed using “unoise3” algorithm and chimeric sequences were removed using “uchime_ref” command. The remaining sequences were used to create ASV table. Taxonomy of the ASVs was assigned to the species with a reference database consisting of the full 16S rRNA gene sequences of the eleven species. The compositional changes were visualized in Microsoft Office Excel 2019. Microbial co-occurrence networks were constructed to show the interactions among species during biofilm development. Spearman correlations among all taxa were calculated using the R *psych* package. Only edges with correlation scores > 0.6 were kept ($p < 0.05$, FDR-adjusted). Correlation networks were visualized via Gephi using the Fruchterman Reingold layout (Bastian et al., 2009).

Reduced SynCom biomass quantification

Biomass was defined by the fresh weight of biofilm. To measure the biomass, the cell strainer was taken out, removed visible drops with paper, and weighed. The fresh weight was total weight minus the weight of the nylon mesh. Each treatment had six biological replicates.

Cell numbers quantification by qPCR

To quantify the cell numbers of each species within the multi-species biofilm, strain-specific primers were designed for the selected six species based on the whole-genome comparison. These primers target the single-copy gene of the corresponding species. Conventional PCR and qPCR melt curves were used to evaluate the specificity of the primers obtained. Primer sequences and target genes were listed in Table 2. Standard curves were generated using plasmids containing corresponding fragments. qPCR was performed with Applied Biosystems Real-Time PCR Instrument. Reaction components are as follow 7.2 μl H_2O, 10 μl 2× ChamQ SYBR qPCR Master Mix (Vazyme), 0.4 μl 10 μM of each primer and 2 μl template DNA. The PCR programs were carried out under the following conditions: 95 °C for 10 min, 40 cycles of 95 °C for 30 s, 60 °C for 45 s, followed by a standard melting curve segment. Each treatment had six biological replicates, and each sample was run in triplicates (technical replicates).

Table 2. Primers used in this study

Primer	Sequence	Product size (bp)	Target gene locus_tag (Genbank)
Aci_F	ATTTAGTATCTGGTGAAGTCATCCGTA	92	GOD87_RS00290
Aci_R	CCGACAAATAAAGCTTGAGTAACTCC		
Com_F	CTCAAAACCAGTGTGATCGTGGAA	109	LAD35_RS00960
Com_R	TATTGCCCATCAGCAGAGTGTAGC		

Chr_F	GAACATCAGTTATCTTGTGAGCGGTA	94	KB553_RS00055
Chr_R	CATACAGGCTCCCATTCCTATTGTG		
Ent_F	AGCGTTACAGCAGCTACAGGATATTCACC	96	K9O83_RS00400
Ent_R	CTTTTCACCATCACCCCATCCCTCGGTA		
Pan_F	TTAACATCGAAAAGCCTTCCCACCGTA	101	LAC65_RS01220
Pan_R	ATTCATCAGAAGCGCATGTATTACACT		
Pse_F	GAAATTCATCTTCGAACACAGCACAC	124	GOM96_RS02845
Pse_R	CTAGCTAACGGGGTTAAGTGCTTC		

Pair-wise assay

The direct competition of these isolates against each other was evaluated using the spot-on-lawn assay and the pair-wise spot assay. Spot-on-lawn assay: 5 ml of lawn species (OD_{600} ~ 0.02) grown in TSB medium was spread onto a 25 ml TSB plate (1.5% agar) and removed by pipetting. Plates were dried for 20 min. 5 µl of spot species (OD_{600}~ 0.4) grown in TSB medium was spot on the center of the plates. Pair-wise spot assay: 5 µl of the dual-species (OD_{600}~ 1) grown in TSB medium were spot on TSB plate (1.5% agar) at 5 mm between the center of each colony. Plates were grown at 30 °C and imaged at 48 h. The experiments were performed twice.

Growth promotion assay

The potential growth promotion of the bacterial metabolites to another species was evaluated using a spent medium growth curve assay (Sun et al., 2021). Donor bacteria were grown in the M9 medium with 1% glucose till the glucose was under detection. The cell culture was spun down, then the spent medium was filter-sterilized and directly used as the medium of growth curve assay. 2 µl of recipient species (OD_{600}~ 1) was inoculated to 200 µl spent medium or M9 glucose medium in a 10×10 well Honeycomb Microplate. OD_{600} was measured every 30 minutes at 30 °C with Bioscreen C Automated Microbiology Growth Curve Analysis System. Each treatment has 3 replicates. The experiments were performed twice. The carrying capacity (maximum population size) was compared.

Genome-scale metabolic modeling

Metabolic models were reconstructed using the CarveMe pipeline (Machado et al., 2018). The quality of the metabolic models was validated using MEMOTE (Lieven et al., 2020). The metabolic interaction potential and metabolic resource overlap for each community were analyzed using SMETANA (Zelezniak et al., 2015; Zorrilla et al., 2021). The simulated cross-feeding results were summarized as SMETANA score, which estimates the strength of metabolic exchanging (Zelezniak et al., 2015).

Carbon source metabolic activity measurement

We used PM1 (BIOLOG Cat #13101) and PM2 (BIOLOG Cat #13102) phenotypic microarrays to assess the carbon source utilization ability of the community members (Bochner et al., 2001). The assays were performed following the manufacturer's instructions. Briefly, 100 µl of diluted cell suspension of each species mixed with the BiOLOG redox dyes were added to each well of the PM plates, and the plates inoculated with water and dyes were used as negative controls. All of the plates were then incubated at 30 °C for up to 48 h. If the species could utilize the carbon source in a well, the colorless tetrazolium dye will be reduced to purple formazan by cell respiration. The color changes were measured by an endpoint absorbance at 590 nm with a microplate reader. The variable level of color changes indicates the carbon source metabolic activity.

Data analysis and figures

All the data needed to evaluate the conclusions in this paper are provided in the figures. Source data related to this paper would be available online upon publication. Data were analyzed using R studio. A random forest approach was used to identify the contribution of the six species responsible for the differences in community composition. The analysis was performed using the R *randomForest* package (Andy Liaw et al., 2018). Plots were generated using Microsoft Office Excel 2019 (stacked bar plot), Graphpad prism 8 (heatmap), R *ggplot2*, *ggpubr, ggalluvial, pheatmap* packages, and Adobe Illustrator CC 2020 (Adobe Inc.). Schematic diagrams were generated using BioRender.

Acknowledgments

This work was financially supported by the National Nature Science Foundation of China (31972512 and 42090064), the Innovative Research Team Development Plan of the Ministry of Education of China (Grant No. IRT_17R56), the Fundamental Research Funds for the Central Universities (KYXK202009). XS was supported by a Chinese Scholarship Council fellowship. ÁTK was supported by the Danish National Research Foundation (DNRF137) for the Center for Microbial Secondary Metabolites and the Novo Nordisk Foundation via the INTERACT project (grant number NNF19SA0059360). During her unexpected trap in Denmark caused by the pandemic of Omicron, author XS is extremely grateful to ZX for his financial support, and to Danish guitarist Niklas Johansen for his weekly lessons and constant spiritual support.

Competing Interests

The authors declare that there are no competing financial interests related to the work described.

Author contributions

XS, JX, ZX, ÁTK designed the study, XS, JX, RX performed the experiments. DZ and WW performed the metabolic modeling. XS and JX analyzed the data and created the figures. XS wrote the first draft of the manuscript, all the authors revised the manuscript.

References

Abram F. 2015. Systems-based approaches to unravel multi-species microbial community functioning. *Computational and Structural Biotechnology Journal* **13**:24–32. doi:10.1016/J.CSBJ.2014.11.009

Barberán A, Bates ST, Casamayor EO, Fierer N. 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME Journal* **6**:343–351. doi:10.1038/ismej.2011.119

Bengtsson-Palme J. 2020. Microbial model communities: To understand complexity, harness the power of simplicity. *Computational and Structural Biotechnology Journal* **18**:3987–4001. doi:10.1016/j.csbj.2020.11.043

Blasche S, Kim Y, Oliveira AP, Patil KR. 2017. Model microbial communities for ecosystems biology. *Current Opinion in Systems Biology* **6**:51–57. doi:10.1016/J.COISB.2017.09.002

Bochner BR, Gadzinski P, Panomitros E. 2001. Phenotype Microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Research* **11**:1246–1255. doi:10.1101/gr.186501

Breugelmans P, Barken KB, Tolker-Nielsen T, Hofkens J, Dejonghe W, Springael D. 2008. Architecture and spatial organization in a triple-species bacterial biofilm synergistically degrading the phenylurea herbicide linuron. *FEMS Microbiology Ecology* **64**:271–282. doi:10.1111/j.1574-6941.2008.00470.x

Cavaliere M, Feng S, Soyer OS, Jiménez JI. 2017. Cooperation in microbial communities and their biotechnological applications. *Environmental Microbiology* **19**:2949–2963. doi:10.1111/1462-2920.13767

Chan SHJ, Simons MN, Maranas CD. 2017. SteadyCom: Predicting microbial abundances while ensuring community stability. *PLOS Computational Biology* **13**:e1005539. doi:10.1371/JOURNAL.PCBI.1005539

Cho I, Blaser MJ. 2012. The human microbiome: at the interface of health and disease. *Nature Reviews Genetics 2012 13:4* **13**:260–270. doi:10.1038/nrg3182

Durán P, Thiergart T, Garrido-Oter R, Agler M, Kemen E, Schulze-Lefert P, Hacquard S. 2018. Microbial Interkingdom Interactions in Roots Promote Arabidopsis Survival. *Cell* **175**:973-983.e14. doi:10.1016/j.cell.2018.10.020

Estrela S, Sanchez-Gorostiaga A, Vila JCC, Sanchez A. 2021. Nutrient dominance governs the assembly of microbial communities in mixed nutrient environments. *eLife* **10**. doi:10.7554/ELIFE.65948

Faust K. 2021. Open challenges for microbial network construction and analysis. *The ISME Journal* 1–8. doi:10.1038/s41396-021-01027-4

Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, Huttenhower C. 2012. Microbial co-occurrence relationships in the Human Microbiome. *PLoS Computational Biology* **8**. doi:10.1371/journal.pcbi.1002606

389 Fitzpatrick CR, Salas-González I, Conway JM, Finkel OM, Gilbert S, Russ D, Teixeira PJPL, Dangl JL. 2020. The
390 Plant Microbiome: From Ecology to Reductionism and beyond. *Annual Review of Microbiology* **74**:81–
391 100. doi:10.1146/annurev-micro-022620-014327

392 Foster KR, Bell T. 2012. Competition, not cooperation, dominates interactions among culturable microbial
393 species. *Current Biology* **22**:1845–1850. doi:10.1016/j.cub.2012.08.005

394 Friedman J, Higgins LM, Gore J. 2017. Community structure follows simple assembly rules in microbial
395 microcosms. *Nature Ecology and Evolution* **1**:1–7. doi:10.1038/s41559-017-0109

396 Fu H, Uchimiya M, Gore J, Moran MA. 2020. Ecological drivers of bacterial community assembly in synthetic
397 phycospheres. *Proceedings of the National Academy of Sciences of the United States of America*
398 **117**:3656–3662. doi:10.1073/PNAS.1917265117/-/DCSUPPLEMENTAL

399 Gao C-H, Cao H, Cai P, Sørensen SJ. 2020. The initial inoculation ratio regulates bacterial coculture
400 interactions and metabolic capacity. *The ISME Journal* 1–12. doi:10.1038/s41396-020-00751-7

401 Gao CH, Zhang M, Wu Y, Huang Q, Cai P. 2019. Divergent Influence to a Pathogen Invader by Resident
402 Bacteria with Different Social Interactions. *Microbial Ecology* **77**:76–86. doi:10.1007/s00248-018-1207-z

403 Goldford JE, Lu N, Bajić D, Estrela S, Tikhonov M, Sanchez-Gorostiaga A, Segrè D, Mehta P, Sanchez A. 2018.
404 Emergent simplicity in microbial community assembly. *Science* **361**:469–474.
405 doi:10.1126/science.aat1168

406 Gómez-Godínez LJ, Martínez-Romero E, Trejo JB, Arteaga-Garibay RI. 2021. Tools and challenges to exploit
407 microbial communities in agriculture. *Current Research in Microbial Sciences* **2**:100062.
408 doi:10.1016/j.crmicr.2021.100062

409 Heinken A, Basile A, Hertel J, Thinnes C, Thiele I. 2021. Genome-scale metabolic modeling of the human
410 microbiome in the era of personalized medicine. *https://doi.org/101146/annurev-micro-060221-*
411 *012134* **75**:199–222. doi:10.1146/ANNUREV-MICRO-060221-012134

412 Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald M, Malfatti S, del Rio TG, Jones CD,
413 Tringe SG, Dangl JL. 2015. Salicylic acid modulates colonization of the root microbiome by specific
414 bacterial taxa. *Science* **349**:860–864. doi:10.1126/science.aaa8764

415 Lee K, Yoon SS. 2017. Pseudomonas aeruginosa biofilm, a programmed bacterial life for fitness. *J Microbiol*
416 *Biotechnol* **27**:1053–1064. doi:10.4014/jmb.1611.11056

417 Lee KWK, Periasamy S, Mukherjee M, Xie C, Kjelleberg S, Rice SA. 2014. Biofilm development and enhanced
418 stress resistance of a model, mixed-species community biofilm. *ISME Journal* **8**:894–907.
419 doi:10.1038/ismej.2013.194

420 Lieven C, Beber ME, Olivier BG, Bergmann FT, Ataman M, Babaei P, Bartell JA, Blank LM, Chauhan S, Correia
421 K, Diener C, Dräger A, Ebert BE, Edirisinghe JN, Faria JP, Feist AM, Fengos G, Fleming RMT, García-
422 Jiménez B, Hatzimanikatis V, van Helvoirt W, Henry CS, Hermjakob H, Herrgård MJ, Kaafarani A, Kim HU,
423 King Z, Klamt S, Klipp E, Koehorst JJ, König M, Lakshmanan M, Lee D-Y, Lee SY, Lee S, Lewis NE, Liu F, Ma
424 H, Machado D, Mahadevan R, Maia P, Mardinoglu A, Medlock GL, Monk JM, Nielsen J, Nielsen LK,
425 Nogales J, Nookaew I, Palsson BO, Papin JA, Patil KR, Poolman M, Price ND, Resendis-Antonio O,
426 Richelle A, Rocha I, Sánchez BJ, Schaap PJ, Malik Sheriff RS, Shoaie S, Sonnenschein N, Teusink B, Vilaça
427 P, Vik JO, Wodke JAH, Xavier JC, Yuan Q, Zakhartsev M, Zhang C. 2020. MEMOTE for standardized
428 genome-scale metabolic model testing. *Nature Biotechnology 2020 38:3* **38**:272–276.
429 doi:10.1038/s41587-020-0446-y

Liu W, Jacquiod S, Brejnrod A, Russel J, Burmølle M, Sørensen SJ. 2019. Deciphering links between bacterial interactions and spatial organization in multispecies biofilms. *ISME Journal* **13**:3054–3066. doi:10.1038/s41396-019-0494-9

Louca S, Polz MF, Mazel F, Albright MBN, Huber JA, O'Connor MI, Ackermann M, Hahn AS, Srivastava DS, Crowe SA, Doebeli M, Parfrey LW. 2018. Function and functional redundancy in microbial systems. *Nature Ecology & Evolution 2018 2:6* **2**:936–943. doi:10.1038/s41559-018-0519-1

Machado D, Andrejev S, Tramontano M, Patil KR. 2018. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Research* **46**:7542–7553. doi:10.1093/NAR/GKY537

Nielsen AT, Tolker-Nielsen T, Barken KB, Molin S. 2000. Role of commensal relationships on the spatial structure of a surface-attached microbial consortium. *Environmental Microbiology* **2**:59–68. doi:10.1046/j.1462-2920.2000.00084.x

Niu B, Paulson JN, Zheng X, Kolter R. 2017. Simplified and representative bacterial community of maize roots. *Proceedings of the National Academy of Sciences of the United States of America* **114**:E2450–E2459. doi:10.1073/pnas.1616148114

Oliveira NM, Martinez-Garcia E, Xavier J, Durham WM, Kolter R, Kim W, Foster KR. 2015. Biofilm formation as a response to ecological competition. *PLoS Biology* **13**:1–23. doi:10.1371/journal.pbio.1002191

Ortiz A, Vega NM, Ratzke C, Gore J. 2021. Interspecies bacterial competition regulates community assembly in the C. elegans intestine. *ISME Journal* 1–12. doi:10.1038/s41396-021-00910-4

Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**:3691–3693. doi:10.1093/BIOINFORMATICS/BTV421

Poudel R, Jumpponen A, Schlatter DC, Paulitz TC, McSpadden Gardener BB, Kinkel LL, Garrett KA. 2016. Microbiome networks: A systems framework for identifying candidate microbial assemblages for disease management. *Phytopathology* **106**:1083–1096. doi:10.1094/PHYTO-02-16-0058-FI

Raghupathi PK, Liu W, Sabbe K, Houf K, Burmølle M, Sørensen SJ. 2018. Synergistic Interactions within a Multispecies biofilm enhance individual species protection against grazing by a pelagic protozoan. *Frontiers in Microbiology* **8**:2649. doi:10.3389/fmicb.2017.02649

Ratzke C, Barrere J, Gore J. 2020. Strength of species interactions determines biodiversity and stability in microbial communities. *Nature Ecology and Evolution* **4**:376–383. doi:10.1038/s41559-020-1099-4

Ren D, Madsen JS, Sørensen SJ, Burmølle M. 2015. High prevalence of biofilm synergy among bacterial soil isolates in cocultures indicates bacterial interspecific cooperation. *ISME Journal* **9**:81–89. doi:10.1038/ismej.2014.96

Rendueles O, Ghigo JM. 2012. Multi-species biofilms: How to avoid unfriendly neighbors. *FEMS Microbiology Reviews* **36**:972–989. doi:10.1111/j.1574-6976.2012.00328.x

Sadiq FA, Burmølle M, Heyndrickx M, Flint S, Lu W, Chen W, Zhao J, Zhang H. 2021. Community-wide changes reflecting bacterial interspecific interactions in multispecies biofilms. *Critical Reviews in Microbiology* 1–21. doi:10.1080/1040841x.2021.1887079

Sun X, Xu Z, Xie J, Hesselberg-Thomsen V, Tan T, Zheng D, Strube ML, Dragoš A, Shen Q, Zhang R, Kovács ÁT. 2021. *Bacillus velezensis* stimulates resident rhizosphere *Pseudomonas stutzeri* for plant health through metabolic interactions. *The ISME Journal 2021* 1–14. doi:10.1038/s41396-021-01125-3

Wei Z, Gu Y, Friman VP, Kowalchuk GA, Xu Y, Shen Q, Jousset A. 2019. Initial soil microbiome composition and functioning predetermine future plant health. *Science Advances* **5**:759–784. doi:10.1126/sciadv.aaw0759

Weiss AS, Burrichter AG, Chakravarthy A, Raj D, von Strempel A, Meng C, Kleigrewe K, Münch PC, Rössler L, Huber C, Eisenreich W, Jochum LM, Göing S, Jung K, Lincetto C, Hübner J, Marinos G, Zimmermann J, Kaleta C, Sanchez A, Stecher B. 2021. In vitro interaction network of a synthetic gut bacterial community. *The ISME Journal 2021* 1–15. doi:10.1038/s41396-021-01153-z

Widder S, Allen RJ, Pfeiffer T, Curtis TP, Wiuf C, Sloan WT, Cordero OX, Brown SP, Momeni B, Shou W, Kettle H, Flint HJ, Haas AF, Laroche B, Kreft J-U, Rainey PB, Freilich S, Schuster S, Milferstedt K, van der Meer JR, Groβkopf T, Huisman J, Free A, Picioreanu C, Quince C, Klapper I, Labarthe S, Smets BF, Wang H, Soyer OS. 2016. Challenges in microbial ecology: building predictive understanding of community function and dynamics. *The ISME Journal 2016 10:11* **10**:2557–2568. doi:10.1038/ismej.2016.45

Zelezniak A, Andrejev S, Ponomarova O, Mende DR, Bork P, Patil KR. 2015. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proceedings of the National Academy of Sciences* **112**:6449–6454. doi:10.1073/PNAS.1421834112

Zorrilla F, Buric F, Patil KR, Zelezniak A. 2021. metaGEM: reconstruction of genome scale metabolic models directly from metagenomes. *Nucleic Acids Research* **49**:e126–e126. doi:10.1093/NAR/GKAB815