The ISME Journal, 2024, **18(1)**, wrae049
https://doi.org/10.1093/ismejo/wrae049
Advance access publication: 27 March 2024
Review Article

Strategies for tailoring functional microbial synthetic communities

Jiayi Jing[1,2], Paolina Garbeva[2,*], Jos M. Raaijmakers[2], Marnix H. Medema[1,*]

[1]Bioinformatics Group, Department of Plant Science, Wageningen University & Research, Droevendaalsesteeg 1, 6708PB Wageningen, The Netherlands
[2]Department of Microbial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Droevendaalsesteeg 10, 6708 PB Wageningen, The Netherlands

*Corresponding authors: Paolina Garbeva. Email: p.garbeva@nioo.knaw.nl and Marnix H. Medema. Email: marnix.medema@wur.nl

Abstract

Natural ecosystems harbor a huge reservoir of taxonomically diverse microbes that are important for plant growth and health. The vast diversity of soil microorganisms and their complex interactions make it challenging to pinpoint the main players important for the life support functions microbes can provide to plants, including enhanced tolerance to (a)biotic stress factors. Designing simplified microbial synthetic communities (SynComs) helps reduce this complexity to unravel the molecular and chemical basis and interplay of specific microbiome functions. While SynComs have been successfully employed to dissect microbial interactions or reproduce microbiome-associated phenotypes, the assembly and reconstitution of these communities have often been based on generic abundance patterns or taxonomic identities and co-occurrences but have only rarely been informed by functional traits. Here, we review recent studies on designing functional SynComs to reveal common principles and discuss multidimensional approaches for community design. We propose a strategy for tailoring the design of functional SynComs based on integration of high-throughput experimental assays with microbial strains and computational genomic analyses of their functional capabilities.

Keywords: synthetic communities, microbial ecology, microbial functions, bioinformatics, high-throughput screening

Introduction

Soil and plants are home to an impressive number of microorganisms pivotal for diverse ecosystem services, including degradation of pollutants, biogeochemical cycling, and supporting plant growth and health. A multitude of captivating natural phenomena, including plant disease suppression [1, 2], plant growth promotion [3, 4], and plant stress resilience [5], have been discovered to have a microbial basis, prompting extensive investigations into the intricate interactions between microorganisms, hosts, and environmental factors. Soil amendments that gave desirable phenotypes by altering soil microbial communities exemplified that fundamental understanding of the metabolic potential of microbial ecosystems can confer agronomic benefits [6, 7]. The development of culture-independent sequencing technologies and the explosion of bioinformatics tools to analyse the resulting meta'omic data have profoundly impacted the understanding of microbial communities in diverse environments. For example, the potential of unique microbes found in extreme environments can be leveraged to address challenges posed by climate change [8, 9]. Such methodologies have generated extensive datasets, offering a rich resource for generating numerous hypotheses. Still, it remains imperative to employ complementary experimental methods for rigorous testing of these hypotheses. Indeed, efforts to (re)construct microbial communities for applications [10-12], identify mechanisms and causality underlying microbiome-associated phenotypes [13-16], and analyse microbe–microbe interactions [17, 18] still strongly rely on culture-dependent microbiology, molecular biology, and plant biology methods due to the necessity of isolating and studying microbial strains and/or communities in a controlled environment (Fig. 1). While individual strains like *Bacillus amyloliquefaciens* and *Bacillus thuringiensis* have been used in biological control in agriculture for decades [19], their efficacy to confer specific phenotypes depends on complex interactions with the resident microbiota and their hosts [20]. Therefore, the design of synthetic communities (SynComs) composed of prioritized strains has become a key technology for studying complex microbiome-associated phenotypes in controlled conditions [16, 21]. This calls for diverse strategies, either for simplifying or deconstructing (drop-out approach) complex communities by identifying essential candidates (top-down) or for incrementally reconstructing a core microbial consortium responsible for specific phenotypes (bottom-up), starting from individual isolates that carry out specific functions [22, 23].

Central to the challenge of designing SynComs is the selection of candidates that are representative of the taxonomic and/or functional characteristics of a microbiome under study. One way to do that is by using taxonomic profiles such as high abundance/representativeness across samples [24, 25], co-occurrence with other community members [26], or differential abundance between samples with contrasting phenotypes [27]. There has been a growing focus in the last decade to explore the microbial biosynthetic potential through (meta)genome mining as a complementary approach to SynCom design in addition to traditional

Received: 8 January 2024. **Revised:** 26 February 2024

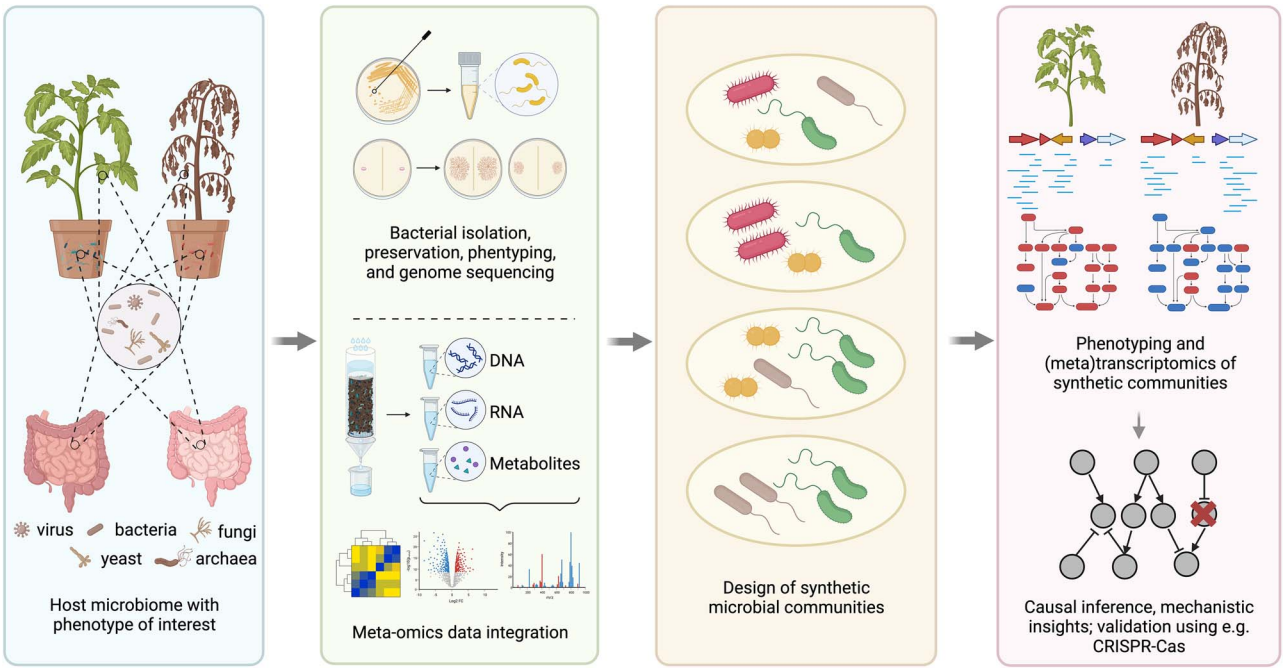

Figure 1. The importance of designing synthetic microbial communities to unravel microbiome-associated phenotypes. Starting often from a host with a phenotype of interest, bacterial strains are isolated and characterized using omics data and/or phenotypic assays. Based on taxonomic or functional traits, synthetic microbial communities with reduced community complexity are designed that can be used to study the mechanistic determinants of the phenotypes under study. Created with BioRender.com.

laboratory screening [28, 29]. Another frontier in this context is adopting *in silico* approaches for the prediction of metabolic interactions, e.g. using genome-scale metabolic models (GSMMs) [30-32].

In this mini-review, we will discuss the pros and cons of several past and present strategies for SynCom design. We will highlight approaches for SynCom design based on functional traits and propose a novel conceptual workflow that combines the strengths of computational (meta)genomic approaches with high-throughput phenotyping.

Strategies for the design of SynComs

Over the last decade, multiple principles in SynCom design and application were employed for diverse study objectives. One approach that is commonly used is taxonomy-based design, which relies on the exploration of microbiome composition in diverse natural samples and the identification of a core or representative microbiome. Exploring microbiome compositions across different geographic environments [33], host genotypes [34], or sampling times [35], (co-occurring sets of) microbial taxa that are persistently present can be selected to mimic the structure and function of the core microbiome. This approach has been frequently employed for the model plant *Arabidopsis* [36] and specific crops [37], as well as in gut microbiome studies [38, 39]. Recently, satellite-based measurements for the global grassland fields meta-data collection were integrated with microbiome data to identify taxa that are closely related to plant productivity [24]. Such principles could also be used for restoring damaged ecosystems by identifying and reconstituting the microbial consortia responsible for ecological stability [40]. Also, combined cross-kingdom SynComs have been constructed based on taxonomic co-occurrence networks that were able to protect tomato against Fusarium wilt disease [41]. In contexts beyond plants, over 100 common bacterial strains in the gut have been engineered into a synthetic community (hCom1), serving as a model system for in-depth exploration of causal inferences and disease mechanisms in the intestinal tract of experimental mice [39]. By iteratively identifying additional colonizing taxa after SynCom introduction into the mice gut and adding these taxa to the community, an expanded community (hCom2) could be created that was more diverse and stable compared with the original SynCom (hCom1) .

A variant of this taxonomy-based strategy that has been widely employed to design SynComs associated with particular phenotypes is based on comparing microbial taxa exhibiting significant abundance differences across samples with contrasting phenotypes. These comparisons can then be utilized to inform bottom-up strategies that involve assembling communities from relatively small numbers of individual microbial strains or species with relevant functional attributes and are likely to provide good starting points toward reconstitution of that phenotype. As an illustration, Zhuang et al. assessed rhizosphere microbiome compositions across different growth stages, soil types, and agricultural practices to identify taxa associated with growth/yield parameters, and used differential abundance analysis to select strains for the construction of a synthetic community that indeed conferred a growth-promoting phenotype to the host [42]. In a similar study analysing microbiome-mediated suppression of bacterial wilt, Kwak et al. could even identify a single flavobacterial strain through differential abundance analysis that was able to largely reconstitute the protective phenotype [43]. Instead of basing the SynCom design on community-level phenotypes, also phenotyping of individual isolates can be used to guide the reconstruction of microbial communities for disease management, as was successfully done to construct a SynCom of just seven strains suppressive against Fusarium wilt in banana [44]. In contrast, top-down approaches focus more on manipulating existing microbial communities through perturbations, such as community transplantation, selective heat treatment, or antimicrobials, that alter

community composition and dynamics. This principle can be a helpful first step in studying functional traits of complex natural microbial communities.

In addition to the foregoing principles, novel SynComs are increasingly established based on broad functional (metabolic) traits of the members of a natural community [18]. Metabolic interactions, including which and how efficiently microbes utilize substrates present in the environment or produced by other community members, drive the whole community's behavior, leading to various phenotypes. Such information has been used to construct a model consortium containing diverse taxa of chitin degraders and non-degraders to study the predicted and realized niches for each isolate; it turned out that the chitin-degrading or, more general, consuming behavior of microbial strains can differ between monoculture and mixed communities [22]. Moreover, predicting competition and substrate preferences by analysing the transcriptional and translational information allowed targeted manipulation of the activity of specific microbial members within natural communities by adding corresponding prebiotics or probiotics [45]. Function-based approaches can also be combined with taxonomic data associated with host phenotypes: for example, Carrion et al. identified taxa that were consistently differentially abundant between the endosphere microbiota of sugarbeet in disease-suppressive and conducive soils; guided by expression analysis of specific biosynthetic gene clusters and chitinase-encoding genes, they identified small SynComs that could largely reconstitute the disease-suppressive phenotype [28].

From the above, it is clear that the design of SynComs is no longer solely based on taxonomy but more and more involves selecting microbiome members that (i) show positive or negative interactions *in vitro* or *in vivo*, (ii) possess specific functional traits, and/or (iii) have complementary/similar niche preferences. However, integrating criteria such as microbial interactions, functional traits, and niche preferences introduces complexity, requiring comprehensive experimental validation and sophisticated analyses. Despite these challenges, this multifaceted approach can enhance SynCom functionality, enabling tailored designs of SynComs with increased resilience.

Prioritization of bioactive microbes or functional genes for SynCom design

For function-based SynCom design strategies, various genomic traits can be considered. Examples of such traits (Table 1) include CAZymes, secretion systems, antifungal metabolites, metallophores, biofilm-formation-associated exopolysaccharides, plant-immuno-stimulating metabolites, phytohormones, and more. How to prioritize functions and microbial members within a complex ecosystem is essential for community re-assembly. Interpreting the vast data generated by high-throughput sequencing technologies for this purpose can be challenging [72]. For example, the extent to which microbial networks constructed based on co-occurrence patterns represent the actual functional diversity in the spatio-temporal context of a given ecosystem is often unclear [73, 74]. The microbiome datasets generally only have relative (and not absolute) abundance data [75], and defining the roles of core and accessory taxa is difficult [76]. Adopting a multidimensional approach, through the integration of different types of 'omics and/or experimental (meta)data, could potentially provide a more accurate depiction of microbial diversity, dynamics, and functions.

A computational framework that adopts functional data for SynCom design was developed in 2018 and operates through top-down integration of metagenomic, metabolomic, and phenotypic datasets, enabling more reliable identification of putative mechanistic associations [77]. Relative to former approaches, this workflow accomplishes dimensionality reduction, filtering of false correlations and data integration through the standardization of data, binning of co-expressed genes and metabolites, and the assimilation of a priori (micro)biological knowledge. Another way of approaching computationally guided SynCom design is through visualizing the community function landscape through statistical learning, identifying potential associations between microbes and functional traits with the aim to better understand the dynamics and/or ecological context of natural or designed microbial communities [78-81]. Based on these function landscape conceptions, a modeling-based iteration provides possibilities to design a complex "high-function" community *in silico* by directed evolution based on carefully selected traits [82].

Knowledge about the spatial distribution and niches occupied by each community member is also an essential factor for keeping a stable community structure after restoration. Different ecological modeling approaches, including the Lotka–Volterra model, consumer-resource model, trait-based model, individual-based models, as well as genome-scale metabolic network models, can be employed for niche prediction [83]. Moreover, experimental approaches such as profiling the utilization of environmentally relevant substrates [84] offer predictions of potential metabolic niches that can be used to infer competitive or cooperative microbial interactions. Novel tools like TbasCO (Trait-based Comparative 'Omics) [85], focusing on expression of metabolic genes, can offer enhanced accuracy in capturing niche-differentiating traits over time. By discerning variations in the expression of genomically encoded functional traits among strains and species under diverse conditions, TbasCO provides nuanced insights into the regulation of genome-encoded functional potential in space and time. Indeed, utilizing combined transcriptional and translational information to predict competition demonstrated notably higher accuracy compared with inferring it from genomic data alone [45]. Genomic information integrated with metabolomic traits is also widely used to identify core genes and consortia that are related to essential metabolites [86]. All these strategies are expected to help analyse the utilization and production of primary and secondary metabolites of the host and co-occurring microbes. Specifically, the primary metabolic capability for abundantly available substrates in the selected environment closely correlated with successful colonization and rapid niche occupation [87-89]. When discussing resilience against stressors such as plant pathogens and parasites, the active role of secondary metabolites appeared to be the prioritized criterion [90, 91].

Computational approaches for trait-based SynCom design

A number of innovative computational approaches have been recently developed to address challenges in tailoring SynCom design based on massive (meta)genomics data, including prioritizing the most relevant microbial interactions, identifying key (ecological) functional traits, and optimizing functional community composition *in silico*. Some of the genome-based tools include antiSMASH [92], which predicts microbial secondary metabolite biosynthetic capabilities, MacSyFinder for the detection of macromolecular systems [93], and PHI-base [94] for pathogenicity identification. For secondary metabolite biosynthetic gene clusters, predicting their ecological functions is crucial to consider them for SynCom design. For example, gene clusters encoding the

Table 1. Examples of functional traits for SynCom design.

Functional trait categories	Example genes/ pathways/compounds	Relevance in SynCom design	Assessment methods/tools	References
Nutrient acquisition	Amino acid, organic acid, sugar and plant polymer catabolic pathways	Influence colonization ability; The potential competition for niches	Eco-plate; experimental testing using specific substrates as the sole C or N source; GEMs	[46, 47]
	Chitinases	Degradation of fungal cell walls	The Carbohydrate-Active EnZymes database (CAZy)	[22, 48]
	Phytase	Improvement of phosphorus availability through phytate degradation	Phytase activity assay; gene expression analysis	[49]
	Phosphate solubilizing genes (e.g. pqq)	Enhancement of plant nutrient availability through phosphate solubilization	Pikovskaya's agar assay for phosphate solubilization; gene expression analysis	[50]
	Nitrogen fixation genes (e.g. *nif* genes)	Contribution to plant growth by fixing atmospheric nitrogen	Acetylene reduction assay for nitrogen fixation; gene expression analysis	[51]
Protein secretion systems	Type VI secretion systems	Potential for secreting bioactive substances	Macromolecular System Finder (MacSyFinder), SecReT6	[52, 53]
Biosynthetic potential	Antifungal or antibacterial compounds (e.g. 2, 4-DAPG)	Growth inhibition or killing of (pathogenic) fungi or bacteria	Genomic prediction using antiSMASH/fungiSMASH	[54, 55]
	Siderophore/Metallophore	Iron/metal competition with other microbial members or pathogens	Genomic prediction using antiSMASH; experimental testing with Chrome Azurol S (CAS) Medium	[56]
	VOCs production	VOCs can influence plant growth and act as signaling molecules	Gas chromatography for VOCs analysis; genomic analysis	[57, 58]
Secretion of plant-immunostimulating primary metabolites	Indole-3-Acetic Acid (IAA)	Stimulate plant growth, development, and can influence the plant's immune response	Liquid/Gas chromatography–mass spectrometry (LC–MS) for quantifying IAA production; gene expression analysis	[59]
	1-Aminocyclopropane-1-Carboxylate (ACC) Deaminase	Modulate ethylene levels in plants, influencing their response to stress	Polymerase chain reaction (PCR) for gene detection; gas chromatography for measuring ethylene levels	[60]
	Exopolysaccharides (EPS), biofilm formation	EPS produced by microbes can act as immunostimulants, influencing plant defense responses, and form a physically protective biofilm	Staining methods for visualizing biofilm formation; genetic analysis of EPS biosynthetic genes	[61, 62]
Secretion of phytohormones	Cytokinin	Cytokinins regulate cell division and differentiation in plants	Enzyme-linked immunosorbent assay (ELISA) for cytokinin detection; genetic analysis	[63, 64]
	Gibberellin	Gibberellins influence plant growth and development, especially stem elongation	High-performance liquid chromatography (HPLC) for gibberellin quantification; gene expression analysis	[65, 66]
	Abscisic acid (ABA)	ABA is involved in plant stress responses and regulates various physiological processes	ELISA for ABA detection; gene expression analysis	[67, 68]
	Ethylene	Ethylene regulates plant growth, fruit ripening, and responses to stress	Gas chromatography for ethylene measurement; gene expression analysis	[69]
Antibiotic resistance genes	Genes associated with antibiotic resistance	Understanding microbial interactions and competition in the community	PCR or metagenomic analysis for antibiotic resistance genes	[70, 71]

For each trait category, examples are provided, relevance is briefly explained, and assessment tools are indicated.

production of metallophores, which are key functional determinants in disease-suppressive soils [95], can be annotated automatically through the identification of genes encoding the biosynthesis of metal-ion-chelating substructures [96]. Carbohydrate-acting enzymes involved in the breakdown of fungal cell walls and plant-derived polymers, can be annotated with automated systems such as dbCAN [97]. Additionally, gene clusters encoding the biosynthesis of antifungals, antibiotics, toxins, or biofilm-associated exopolysaccharides can be identified through comparison with reference biosynthetic gene clusters encoding products of known function, such as those deposited in the MIBiG database [98]. Similarly, reference databases of virulence factors (e.g. VFDB [99]) or secretion systems (e.g. SecReT6 [100]) can aid in the identification of pathogenicity-related functional traits.

Genome-scale metabolic network models (GSMM/GEMs) have experienced a notable rise in microbiome studies and are particularly advantageous in the context of predicting functional interactions within microbial communities [31, 32, 101-103]. Moreover, alongside the rise of GSMM, graph-theoretic approaches offer valuable insights into microbial community dynamics, particularly in predicting biotic interactions and understanding the influence of nutrients and the environment [104, 105]. Such approaches were employed in identifying minimal sets of species for desired metabolic potential [106], and/or elucidating metabolic exchanges between organisms [102, 107]. An exciting study employed GSMMs to estimate the competitive and corporate potential across thousands of habitats. The results indicated competitive communities resist species invasion but struggle with nutrient shifts, while cooperative communities show the opposite pattern [108]. Multiple tools have been created for automated metabolic network reconstruction of microbial species as well as communities [109-114]. MiMiC is one of the most straightforward tools for designing functional representative SynComs by utilizing shotgun metagenomic data for protein family annotations and aims to cover a maximum number of functions within the community with a minimum number of microbial taxa [115]. Similar to MiMiC, CoMiDA identifies potential metabolic pathways from substrates to products instead of using protein families and aims to find minimal combinations to perform these processes [106]. However, critical factors like inter-member growth compatibility, exchange of metabolites, cross-feeding, differential regulation of metabolic traits, and co-cultivation conditions still require incorporation within these algorithms. In efforts to narrow this gap, FLYCOP utilizes GEMs to assign metabolic potentials and COMETS (Computation of Microbial Ecosystems in Time and Space) [116] to predict microbial interactions and their dynamic flux balance to further simulate community dynamics thru iterative algorithms and identify the optimal combinations between multiple consortium configurations [30].

Artificial intelligence for SynCom design

Machine learning (ML) and artificial intelligence (AI) are increasingly used for (iterative) experimental optimization of SynComs, as they can help to navigate the highly dimensional combinatorial space of taxa and functions. For example, BacterAI, a novel automated science platform, allowed the design and use of an experimental platform to generate growth data as a "reward" dataset for further optimizing the model to improve the experimental design. Microbial metabolic activity prediction was efficiently generated through active learning on iterative designs without prior knowledge [117]. However, there are still challenges regarding the use of these approaches for tailoring SynComs because of the limitations of available dataset sizes and the lack of evaluation standards for measuring SynCom quality. Moreover, AI and/or ML approaches should be used with caution, since they can give false or invalid associations when used without validation. A recent study identified extremely accurate predictions of tumor types and presence using microbial abundance patterns [118], whereas these correlations were demonstrated to be fictional upon further analysis, thus illustrating risks due to inadvertently training on contamination, batch effect, or false positive classifications [119, 120]. An innovative attempt has been made to utilize the prediction of causal relationships between microbial members and host phenotypes to develop novel SynComs using deep learning methods [14]. Specifically, their approach involves characterizing the relationship between bioassays (i.e. growth on *Arabidopsis* root exudate for each strain), defining functional blocks by grouping the strains based on their effects on plant Pi content, creating partially overlapping SynComs, and utilizing a neural network model to design novel microbial combinations for predicting Pi content in plant. The experimental validation results suggested that nearly all of these predicted Pi content was indeed realized in the *in planta* assay. Another data-driven framework to identify keystone species (microbial taxa that are essential for a stable community structure) employed deep learning to quantify the importance of each species by conducting drop-out assays [121]. Such assays were widely used to systematically eliminate SynCom members and check if/how this "drop-out" affected the microbiome-associated phenotypes [122].

In an era of rapid advancements in AI, the establishment of community-level GEMs is poised to become increasingly efficient and reliable for predicting metabolic interactions among microbes and how they cooperatively utilize substrates both pre-existing and generated during microbial activities. Combined with AI-driven cycles between computational designs and experimental assays to iteratively validate interactions and improve SynComs, the associations generated by these tools can be further employed to tailor a wider range of SynComs with pre-defined functions. These computationally tailored SynComs may exhibit superior colonization capabilities and metabolic potential compared with manually designed ones.

Aspects affecting the reconstitution of SynComs

The utilization of different tools for crafting microbiota communities responsible for specific (metabolic) functions in the context of microbiota transplantation strategies holds great promise for the future. Nonetheless, the ability of the predicted communities to successfully colonize true hosts will remain enigmatic until subjected to validation in wet lab, greenhouse, and field/host experiments. As the transition from the selection and combination of SynCom members to their reconstitution, a myriad of additional challenges are faced, including the need to reconcile disparate growth rates among microbes, the determination of the order of inoculation (i.e. priority effects), the amount of cell density of each candidate strain [123], and the evaluation of potential interactions that could result in the loss of certain SynCom members during the process. Furthermore, variations in initial concentrations for strains that have different growth rates may have a substantial impact on the ultimate structure and stability of the assembled community [124]. All these variations are expected to lead to increased functional stochasticity when employing SynComs for investigating interactions or causal inferences. This underscores

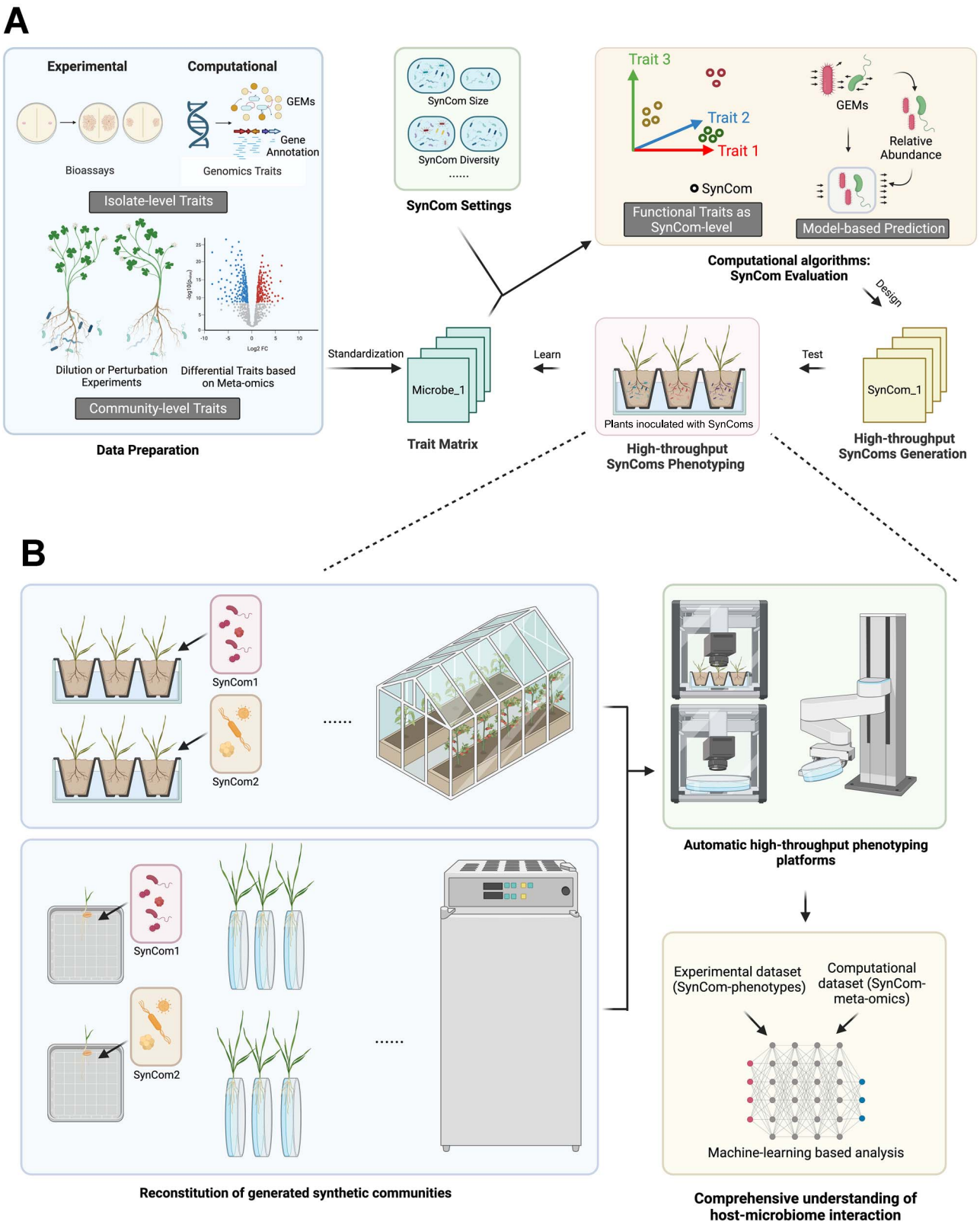

Figure 2. Proposed conceptual workflow for SynCom design. (a) Computational high-throughput SynCom design and validation. Functional traits at both the isolate and the community level will first be identified by experimental/computational strategies. The resulting trait matrix will then be used for high-throughput SynCom generation and validation, using an iterative design-test-learn cycle. (b) High-throughput SynCom screening and ML-based analysis. The generated SynComs will be reconstituted and screened for phenotypes using automated high-throughput phenotyping platforms. The observed phenotyping dataset as well as correlated meta-omics, i.e. rhizosphere meta-transcriptomics data, can be used as (extended) training data for ML-based analysis to obtain an enhanced understanding of host-microbiome interactions and design increasingly more effective and stable SynComs.

the necessity of monitoring the community composition and structural stability through low-pass metagenomic sequencing, qPCR data, or fluorescent markers during different stages of the reconstitution process. Alternatively, metabolic modeling may be able to predict niche complementarity and community stability in the future, especially if it can be fine-tuned by experimental data such as those mentioned above.

Priority effects, which refer to the timing of introduction of the microbial taxa and the advantages to establish themselves in specific ecological niches (principle of "first come first serve"), have been studied across various host systems [125]. This phenomenon has also been widely employed to modulate competition in the restoration of microbial communities [126]. When addressing the restoration of SynComs in the lab, a new strategy involves grouping microbes with similar functions or taxonomies, enabling the inference of interactions or associations between certain groups and host phenotypes by introducing or eliminating each separately [14]. This top-down strategy demands considerable lab work including high-throughput automated phenotyping [127, 128], as well as controlled gnotobiotic experimental systems [129, 130] that mimic natural complexity. Amidst numerous related endeavors, the development of EcoFABs (reproducible fabricated ecosystems) stands out as a significant attempt toward standardizing microbial community model systems [131]. This system facilitates standardization of every step in the process, with defined microbiota, laboratory habitats, and reproducible protocols for cultivation and spatiotemporal analysis.

Synergizing bioinformatics and high-throughput validation for Syncom design

The evolution of high-throughput phenotypic platforms as well as the development of cloud laboratories have significantly mitigated the constraints associated with phenotyping. In recent investigations, researchers restored 136 randomly assembled SynComs of diverse scales into plant systems [132]. The experimental data derived from these trials were employed as a dataset for ML, leading to the successful identification of microbial strains predictive of phenotypic outcomes. While traditional SynCom design methodologies may remain effective for specific functions or as a simplified model system, these novel conceptual frameworks are needed to process and extract meaningful insights from big data. We propose that computational data processing should encompass the integration of functional traits across diverse dimensions, including phenotypes from both large-scale functional assays and *in silico* predictions that can be calibrated and recalibrated against experimental data (Fig. 2). This will result in a standardized trait matrix for each candidate microbe. Together with different SynCom design parameters, including the size of the communities, the desired taxonomic diversity among others, the generated SynComs can be evaluated by calculating functional traits at the SynCom level and/or using model-based strategies to predict SynCom functions. From this, multiple alternative SynComs can be constructed having similar functional trait compositions from different taxonomic origins, which allows us to explore multiple possible solutions in parallel. High-throughput phenotypic systems will then yield tractable sample information post-inoculation of such diverse SynComs, encompassing parameters such as plant biomass (via 3D scanning), stress protective effects, growth form, alterations in plant root exudates including volatile organic compounds (VOCs), and gene expression differences (via meta-transcriptomics). The generated combinations, along with their phenotypic data, could then be reused as input data for AI-based tools to learn and model SynCom functionality and predict community-level phenotypes, and help select new SynCom designs to iteratively improve performance. In the future, it may be feasible to build databases for SynCom-related datasets and explore correlations based on massive SynCom datasets associated with different hosts and phenotypes to identify genotype–phenotype patterns across laboratories. Overall, our proposed conceptual workflow presents a different perspective for the design of SynComs by incorporating multidimensional data information from *in vitro* and *in vivo* assays as well as computational predictions. We anticipate that this will accelerate the adoption of SynComs as potent experimental tools in the forthcoming era of microbial ecology research.

Conflicts of interest

The authors declare no conflict of interest.

Data availability

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Schlatter D, Kinkel L, Thomashow L *et al.* Disease suppressive soils: new insights from the soil microbiome. *Phytopathology* 2017;**107**:1284–97. https://doi.org/10.1094/PHYTO-03-17-0111-RVW
2. Mendes R, Kruijt M, De Bruijn I *et al.* Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* 2011;**332**:1097–100. https://doi.org/10.1126/science.1203980
3. Bhattacharyya PN, Jha DK. Plant growth-promoting rhizobacteria (PGPR): emergence in agriculture. *World J Microbiol Biotechnol* 2012;**28**:1327–50. https://doi.org/10.1007/s11274-011-0979-9
4. Kaymak HC. Potential of PGPR in agricultural innovations. In: Maheshwari D.K. (ed.), *Plant Growth and Health Promoting Bacteria*. Berlin, Heidelberg. Berlin, Heidelberg: Springer, 2010, 45–79
5. Liu H, Brettell LE, Qiu Z *et al.* Microbiome-mediated stress resistance in plants. *Trends Plant Sci* 2020;**25**:733–43. https://doi.org/10.1016/j.tplants.2020.03.014
6. Morales-Salmerón L, Fernández-Boy E, Madejón E *et al.* Soil legacy and organic amendment role in promoting the resistance of contaminated soils to drought. *Appl Soil Ecol* 2024;**195**:105226. https://doi.org/10.1016/j.apsoil.2023.105226
7. Solanki MK, Joshi NC, Singh PK *et al.* From concept to reality: transforming agriculture through innovative rhizosphere engineering for plant health and productivity. *Microbiol Res* 2024;**279**:127553. https://doi.org/10.1016/j.micres.2023.127553
8. Rodriguez R, Durán P. Natural Holobiome engineering by using native extreme microbiome to counteract the climate change effects. *Front Bioeng Biotechnol* 2020;**8**:568. https://doi.org/10.3389/fbioe.2020.00568
9. Mukhtar S, Mehnaz S, Malik KA. Microbial diversity in the rhizosphere of plants growing under extreme environments and its impact on crop improvement. *Environ Sustain* 2019;**2**:329–38. https://doi.org/10.1007/s42398-019-00061-5
10. Suman A, Govindasamy V, Ramakrishnan B *et al.* Microbial community and function-based synthetic bioinoculants: a perspective for sustainable agriculture. *Front Microbiol* 2022;**12**:805498. https://doi.org/10.3389/fmicb.2021.805498

11. Mazza Rodrigues JL, Melotto M. Naturally engineered plant microbiomes in resource-limited ecosystems. *Trends Microbiol* 2023;**31**:329–31. https://doi.org/10.1016/j.tim.2023.02.006
12. Perreault R, Laforest-Lapointe I. Plant-microbe interactions in the phyllosphere: facing challenges of the anthropocene. *ISME J* 2022;**16**:339–45. https://doi.org/10.1038/s41396-021-01109-3
13. Nobori T, Cao Y, Entila F *et al.* Dissecting the cotranscriptome landscape of plants and their microbiota. *EMBO Rep* 2022;**23**:e55380. https://doi.org/10.15252/embr.202255380
14. Herrera Paredes S, Gao T, Law TF *et al.* Design of synthetic bacterial communities for predictable plant phenotypes. *PLoS Biol* 2018;**16**:e2003962. https://doi.org/10.1371/journal.pbio.2003962
15. Müller DB, Vogel C, Bai Y *et al.* The plant microbiota: systems-level insights and perspectives. *Annu Rev Genet* 2016;**50**:211–34. https://doi.org/10.1146/annurev-genet-120215-034952
16. Vorholt JA, Vogel C, Carlström CI *et al.* Establishing causality: opportunities of synthetic communities for plant microbiome research. *Cell Host Microbe* 2017;**22**:142–55. https://doi.org/10.1016/j.chom.2017.07.004
17. Dundore-Arias JP, Michalska-Smith M, Millican M *et al.* More than the sum of its parts: unlocking the power of network structure for understanding organization and function in microbiomes. *Annu Rev Phytopathol* 2023;**61**:403–23. https://doi.org/10.1146/annurev-phyto-021021-041457
18. Wang M, Osborn LJ, Jain S *et al.* Strain dropouts reveal interactions that govern the metabolic output of the gut microbiome. *Cell* 2023;**186**:2839–2852.e21. https://doi.org/10.1016/j.cell.2023.05.037
19. Jacobsen BJ, Zidack NK, Larson BJ. The role of *Bacillus* -based biological control agents in integrated pest management systems: plant diseases. *Phytopathology* 2004;**94**:1272–5. https://doi.org/10.1094/PHYTO.2004.94.11.1272
20. Trivedi P, Leach JE, Tringe SG *et al.* Author correction: plant–microbiome interactions: from community assembly to plant health. *Nat Rev Microbiol* 2021;**19**:72–2. https://doi.org/10.1038/s41579-020-00490-8
21. Song C, Jin K, Raaijmakers JM. Designing a home for beneficial plant microbiomes. *Curr Opin Plant Biol* 2021;**62**:102025. https://doi.org/10.1016/j.pbi.2021.102025
22. McClure R, Farris Y, Danczak R *et al.* Interaction networks are driven by community-responsive phenotypes in a chitin-degrading consortium of soil microbes. *mSystems* 2022;**7**:e00372–22. https://doi.org/10.1128/msystems.00372-22
23. Huet S, Romdhane S, Breuil M-C *et al.* Experimental community coalescence sheds light on microbial interactions in soil and restores impaired functions. *Microbiome* 2023;**11**:42. https://doi.org/10.1186/s40168-023-01480-7
24. Delgado-Baquerizo M. Simplifying the complexity of the soil microbiome to guide the development of next-generation SynComs. *J Sustain Agric Environ* 2022;**1**:9–15. https://doi.org/10.1002/sae2.12012
25. Camargo AP, De Souza RSC, Jose J *et al.* Plant microbiomes harbor potential to promote nutrient turnover in impoverished substrates of a Brazilian biodiversity hotspot. *ISME J* 2023;**17**:354–70. https://doi.org/10.1038/s41396-022-01345-1
26. Bell TH, Bell T. Many roads to bacterial generalism. *FEMS Microbiol Ecol* 2020;**97**:fiaa240. https://doi.org/10.1093/femsec/fiaa240
27. Berihu M, Somera TS, Malik A *et al.* A framework for the targeted recruitment of crop-beneficial soil taxa based on network analysis of metagenomics data. *Microbiome* 2023;**11**:8. https://doi.org/10.1186/s40168-022-01438-1
28. Carrión VJ, Perez-Jaramillo J, Cordovez V *et al.* Pathogen-induced activation of disease-suppressive functions in the endophytic root microbiome. *Science* 2019;**366**:606–12. https://doi.org/10.1126/science.aaw9285
29. Getzke F, Hassani MA, Crüsemann M *et al.* Cofunctioning of bacterial exometabolites drives root microbiota establishment. *Proc Natl Acad Sci* 2023;**120**:e2221508120. https://doi.org/10.1073/pnas.2221508120
30. García-Jiménez B, García JL, Nogales J. FLYCOP: metabolic modeling-based analysis and engineering microbial communities. *Bioinformatics* 2018;**34**:i954–63. https://doi.org/10.1093/bioinformatics/bty561
31. Ye C, Wei X, Shi T *et al.* Genome-scale metabolic network models: from first-generation to next-generation. *Appl Microbiol Biotechnol* 2022;**106**:4907–20. https://doi.org/10.1007/s00253-022-12066-y
32. Mataigne V, Vannier N, Vandenkoornhuyse P *et al.* Multi-genome metabolic modeling predicts functional inter-dependencies in the Arabidopsis root microbiome. *Microbiome* 2022;**10**:217. https://doi.org/10.1186/s40168-022-01383-z
33. Mittelstrass J, Sperone FG, Horton MW. Using transects to disentangle the environmental drivers of plant-microbiome assembly. *Plant Cell Environ* 2021;**44**:3745–55. https://doi.org/10.1111/pce.14190
34. van Leeuwen PT, Brul S, Zhang J *et al.* Synthetic microbial communities (SynComs) of the human gut: design, assembly, and applications. *FEMS Microbiol Rev* 2023;**47**:fuad012. https://doi.org/10.1093/femsre/fuad012
35. Spor A, Koren O, Ley R. Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat Rev Microbiol* 2011;**9**:279–90. https://doi.org/10.1038/nrmicro2540
36. Bai Y, Müller DB, Srinivas G *et al.* Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* 2015;**528**:364–9. https://doi.org/10.1038/nature16192
37. Niu B, Paulson JN, Zheng X *et al.* Simplified and representative bacterial community of maize roots. *Proc Natl Acad Sci* 2017;**114**:114. https://doi.org/10.1073/pnas.1616148114
38. Dill-McFarland KA, Weimer PJ, Pauli JN *et al.* Diet specialization selects for an unusual and simplified gut microbiota in two- and three-toed sloths. *Environ Microbiol* 2016;**18**:1391–402. https://doi.org/10.1111/1462-2920.13022
39. Cheng AG, Ho P-Y, Aranda-Díaz A *et al.* Design, construction, and in vivo augmentation of a complex gut microbiome. *Cell* 2022;**185**:3617–3636.e19. https://doi.org/10.1016/j.cell.2022.08.003
40. Shade A. Microbiome rescue: directing resilience of environmental microbial communities. *Curr Opin Microbiol* 2023;**72**:102263. https://doi.org/10.1016/j.mib.2022.102263
41. Zhou X, Wang J, Liu F *et al.* Cross-kingdom synthetic microbiota supports tomato suppression of Fusarium wilt disease. *Nat Commun* 2022;**13**:7890. https://doi.org/10.1038/s41467-022-35452-6
42. Zhuang L, Li Y, Wang Z *et al.* Synthetic community with six *Pseudomonas* strains screened from garlic rhizosphere microbiome promotes plant growth. *Microb Biotechnol* 2021;**14**:488–502. https://doi.org/10.1111/1751-7915.13640
43. Kwak M-J, Kong HG, Choi K *et al.* Rhizosphere microbiome structure alters to enable wilt resistance in tomato. *Nat Biotechnol* 2018;**36**:1100–9. https://doi.org/10.1038/nbt.4232
44. Prigigallo MI, Gómez-Lama Cabanás C, Mercado-Blanco J *et al.* Designing a synthetic microbial community devoted to biological control: the case study of Fusarium wilt of

banana. *Front Microbiol* 2022;**13**:967885. https://doi.org/10.3389/fmicb.2022.967885
45. Moyne O, Al-Bassam M, Lieng C *et al.* Guild and niche determination enable targeted alteration of the microbiome. *BioRxiv* 2023. https://doi.org/10.1101/2023.05.11.540389
46. Park H, Patel A, Hunt KA *et al.* Artificial consortium demonstrates emergent properties of enhanced cellulosic-sugar degradation and biofuel synthesis. *Npj Biofilms Microbiomes* 2020;**6**:59. https://doi.org/10.1038/s41522-020-00170-8
47. Zomorrodi AR, Segrè D. Synthetic ecology of microbes: mathematical models and applications. *J Mol Biol* 2016;**428**:837–61. https://doi.org/10.1016/j.jmb.2015.10.019
48. McClure R, Naylor D, Farris Y *et al.* Development and analysis of a stable, reduced complexity model soil microbiome. *Front Microbiol* 2020;**11**:1987. https://doi.org/10.3389/fmicb.2020.01987
49. Shulse CN, Chovatia M, Agosto C *et al.* Engineered root bacteria release plant-available phosphate from phytate. *Appl Environ Microbiol* 2019;**85**:e01210–9. https://doi.org/10.1128/AEM.01210-19
50. De Zutter N, Ameye M, Debode J *et al.* Shifts in the rhizobiome during consecutive *in planta* enrichment for phosphate-solubilizing bacteria differentially affect maize P status. *Microb Biotechnol* 2021;**14**:1594–612. https://doi.org/10.1111/1751-7915.13824
51. Venkataraman M, Yñigez-Gutierrez A, Infante V *et al.* Synthetic biology toolbox for nitrogen-fixing soil microbes. *ACS Synth Biol* 2023;**12**:3623–34. https://doi.org/10.1021/acssynbio.3c00414
52. Russell AB, Peterson SB, Mougous JD. Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol* 2014;**12**:137–48. https://doi.org/10.1038/nrmicro3185
53. Wang B, Zhang Z, Xu F *et al.* Soil bacterium manipulates antifungal weapons by sensing intracellular type IVA secretion system effectors of a competitor. *ISME J* 2023;**17**:2232–46. https://doi.org/10.1038/s41396-023-01533-7
54. Gong L, Tan H, Chen F *et al.* Novel synthesized 2, 4-DAPG analogues: antifungal activity, mechanism and toxicology. *Sci Rep* 2016;**6**:32266. https://doi.org/10.1038/srep32266
55. Jousset A, Becker J, Chatterjee S *et al.* Biodiversity and species identity shape the antifungal activity of bacterial communities. *Ecology* 2014;**95**:1184–90. https://doi.org/10.1890/13-1215.1
56. Feng Z, Sun H, Qin Y *et al.* A synthetic community of siderophore-producing bacteria increases soil selenium bioavailability and plant uptake through regulation of the soil microbiome. *Sci Total Environ* 2023;**871**:162076. https://doi.org/10.1016/j.scitotenv.2023.162076
57. de Boer W, Li X, Meisner A *et al.* Pathogen suppression by microbial volatile organic compounds in soils. *FEMS Microbiol Ecol* 2019;**95**:fiz105. https://doi.org/10.1093/femsec/fiz105
58. Weisskopf L, Schulz S, Garbeva P. Microbial volatile organic compounds in intra-kingdom and inter-kingdom interactions. *Nat Rev Microbiol* 2021;**19**:391–404. https://doi.org/10.1038/s41579-020-00508-1
59. Liu H, Qiu Z, Ye J *et al.* Effective colonisation by a bacterial synthetic community promotes plant growth and alters soil microbial community. *J Sustain Agric Environ* 2022;**1**:30–42. https://doi.org/10.1002/sae2.12008
60. Penrose DM, Glick BR. Methods for isolating and characterizing ACC deaminase-containing plant growth-promoting rhizobacteria. *Physiol Plant* 2003;**118**:10–5. https://doi.org/10.1034/j.1399-3054.2003.00086.x
61. Song C, Zhao C, Wang Q *et al.* Impact of carbon/nitrogen ratio on the performance and microbial community of sequencing batch biofilm reactor treating synthetic mariculture wastewater. *J Environ Manag* 2021;**298**:113528. https://doi.org/10.1016/j.jenvman.2021.113528
62. Karygianni L, Ren Z, Koo H *et al.* Biofilm Matrixome: extracellular components in structured microbial communities. *Trends Microbiol* 2020;**28**:668–81. https://doi.org/10.1016/j.tim.2020.03.016
63. Giron D, Frago E, Glevarec G *et al.* Cytokinins as key regulators in plant–microbe–insect interactions: connecting plant growth and defence. *Funct Ecol* 2013;**27**:599–609. https://doi.org/10.1111/1365-2435.12042
64. Gupta R, Elkabetz D, Leibman-Markus M *et al.* Cytokinin drives assembly of the phyllosphere microbiome and promotes disease resistance through structural and chemical cues. *ISME J* 2022;**16**:122–37. https://doi.org/10.1038/s41396-021-01060-3
65. Keswani C, Singh SP, García-Estrada C *et al.* Biosynthesis and beneficial effects of microbial gibberellins on crops for sustainable agriculture. *J Appl Microbiol* 2022;**132**:1597–615. https://doi.org/10.1111/jam.15348
66. Nett RS, Bender KS, Peters RJ. Production of the plant hormone gibberellin by rhizobia increases host legume nodule size. *ISME J* 2022;**16**:1809–17. https://doi.org/10.1038/s41396-022-01236-5
67. Shi T-Q, Peng H, Zeng S-Y *et al.* Microbial production of plant hormones: opportunities and challenges. *Bioengineered* 2017;**8**:124–8. https://doi.org/10.1080/21655979.2016.1212138
68. Shahzad R, Khan AL, Bilal S *et al.* Inoculation of abscisic acid-producing endophytic bacteria enhances salinity stress tolerance in Oryza sativa. *Environ Exp Bot* 2017;**136**:68–77. https://doi.org/10.1016/j.envexpbot.2017.01.010
69. Ravanbakhsh M, Sasidharan R, Voesenek LACJ *et al.* Microbial modulation of plant ethylene signaling: ecological and evolutionary consequences. *Microbiome* 2018;**6**:52. https://doi.org/10.1186/s40168-018-0436-1
70. Gibson MK, Forsberg KJ, Dantas G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J* 2015;**9**:207–16. https://doi.org/10.1038/ismej.2014.106
71. Wang Z, Han M, Li E *et al.* Distribution of antibiotic resistance genes in an agriculturally disturbed lake in China: their links with microbial communities, antibiotics, and water quality. *J Hazard Mater* 2020;**393**:122426. https://doi.org/10.1016/j.jhazmat.2020.122426
72. Faust K. Open challenges for microbial network construction and analysis. *ISME J* 2021;**15**:3111–8. https://doi.org/10.1038/s41396-021-01027-4
73. Faust K, Raes J. Microbial interactions: from networks to models. *Nat Rev Microbiol* 2012;**10**:538–50. https://doi.org/10.1038/nrmicro2832
74. Dini-Andreote F, Kowalchuk GA, Prosser JI *et al.* Towards meaningful scales in ecosystem microbiome research. *Environ Microbiol* 2021;**23**:1–4. https://doi.org/10.1111/1462-2920.15276
75. Gloor GB, Macklaim JM, Pawlowsky-Glahn V *et al.* Microbiome datasets are compositional: and this is not optional. *Front Microbiol* 2017;**8**:2224. https://doi.org/10.3389/fmicb.2017.02224
76. Escalas A, Paula FS, Guilhaumon F *et al.* Macroecological distributions of gene variants highlight the functional organization of soil microbial systems. *ISME J* 2022;**16**:726–37. https://doi.org/10.1038/s41396-021-01120-8
77. Pedersen HK, Forslund SK, Gudmundsdottir V *et al.* A computational framework to integrate high-throughput '-omics' datasets for the identification of potential mechanistic links. *Nat Protoc* 2018;**13**:2781–800. https://doi.org/10.1038/s41596-018-0064-z

78. Xie L, Shou W. Steering ecological-evolutionary dynamics to improve artificial selection of microbial communities. *Nat Commun* 2021;**12**:6799. https://doi.org/10.1038/s41467-021-26647-4
79. Sánchez Á, Vila JCC, Chang C-Y *et al.* Directed evolution of microbial communities. *Annu Rev Biophys* 2021;**50**:323–41. https://doi.org/10.1146/annurev-biophys-101220-072829
80. Amor DR. Smooth functional landscapes in microcosms. *Nat Ecol Evol* 2023;**7**:1754–5. https://doi.org/10.1038/s41559-023-02214-6
81. Skwara A, Gowda K, Yousef M *et al.* Statistically learning the functional landscape of microbial communities. *Nat Ecol Evol* 2023;**7**:1823–33. https://doi.org/10.1038/s41559-023-02197-4
82. Chang C-Y, Vila JCC, Bender M *et al.* Engineering complex communities by directed evolution. *Nat Ecol Evol* 2021;**5**:1011–23. https://doi.org/10.1038/s41559-021-01457-5
83. Van Den Berg NI, Machado D, Santos S *et al.* Ecological modelling approaches for predicting emergent properties in microbial communities. *Nat Ecol Evol* 2022;**6**:855–65. https://doi.org/10.1038/s41559-022-01746-7
84. Rutgers M, Wouterse M, Drost SM *et al.* Monitoring soil bacteria with community-level physiological profiles using biolog™ ECO-plates in the Netherlands and Europe. *Appl Soil Ecol* 2016;**97**:23–35. https://doi.org/10.1016/j.apsoil.2015.06.007
85. McDaniel EA, Van Steenbrugge JJM, Noguera DR *et al.* TbasCO: trait-based comparative 'omics identifies ecosystem-level and niche-differentiating adaptations of an engineered microbiome. *ISME Commun* 2022;**2**:111. https://doi.org/10.1038/s43705-022-00189-2
86. Bauermeister A, Mannochio-Russo H, Costa-Lotufo LV *et al.* Mass spectrometry-based metabolomics in microbiome investigations. *Nat Rev Microbiol* 2022;**20**:143–60. https://doi.org/10.1038/s41579-021-00621-9
87. Ghirardi S, Dessaint F, Mazurier S *et al.* Identification of traits shared by rhizosphere-competent strains of fluorescent pseudomonads. *Microb Ecol* 2012;**64**:725–37. https://doi.org/10.1007/s00248-012-0065-3
88. Sohn SI, Ahn JH, Pandian S *et al.* Dynamics of bacterial community structure in the rhizosphere and root nodule of soybean: impacts of growth stages and varieties. *Int J Mol Sci* 2021;**22**:5577–7. https://doi.org/10.3390/ijms22115577
89. Zboralski A, Biessy A, Savoie MC *et al.* Metabolic and genomic traits of phytobeneficial phenazine-producing pseudomonas spp. are linked to rhizosphere colonization in arabidopsis thaliana and solanum tuberosum. *Appl Environ Microbiol* 2020;**86**:86. https://doi.org/10.1128/AEM.02443-19
90. Adedeji AA, Babalola OO. Secondary metabolites as plant defensive strategy: a large role for small molecules in the near root region. *Planta* 2020;**252**:61. https://doi.org/10.1007/s00425-020-03468-1
91. Buddhika UVA, Abeysinghe S. Secondary metabolites from microbes for plant disease management. In: Singh K.P., Jahagirdar S., Sarma B.K. (eds.), *Emerging Trends in Plant Pathology*. Singapore: Springer Singapore, 2021, 331–42
92. Blin K, Shaw S, Augustijn HE *et al.* antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* 2023;**51**:W46–50. https://doi.org/10.1093/nar/gkad344
93. Néron B, Denise R, Coluzzi C *et al.* MacSyFinder v2: improved modelling and search engine to identify molecular systems in genomes. *Peer Community J* 2023;**3**:e28. https://doi.org/10.24072/pcjournal.250
94. Urban M, Cuzick A, Seager J *et al.* PHI-base in 2022: a multi-species phenotype database for pathogen–host interactions. *Nucleic Acids Res* 2022;**50**:D837–47. https://doi.org/10.1093/nar/gkab1037
95. Gu S, Wei Z, Shao Z *et al.* Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. *Nat Microbiol* 2020;**5**:1002–10. https://doi.org/10.1038/s41564-020-0719-8
96. Reitz ZL, Butler A, Medema MH. Automated genome mining predicts combinatorial diversity and taxonomic distribution of peptide metallophore structures. *bioRxiv* 2022. https://doi.org/10.1101/2022.12.14.519525
97. Zheng J, Ge Q, Yan Y *et al.* dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res* 2023;**51**:W115–21. https://doi.org/10.1093/nar/gkad328
98. Terlouw BR, Blin K, Navarro-Muñoz JC *et al.* MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Res* 2023;**51**:D603–10. https://doi.org/10.1093/nar/gkac1049
99. Liu B, Zheng D, Zhou S *et al.* VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res* 2022;**50**:D912–7. https://doi.org/10.1093/nar/gkab1107
100. Zhang J, Guan J, Wang M *et al.* SecReT6 update: a comprehensive resource of bacterial type VI secretion systems. *Sci China Life Sci* 2023;**66**:626–34. https://doi.org/10.1007/s11427-022-2172-x
101. Ibrahim M, Raajaraam L, Raman K. Modelling microbial communities: harnessing consortia for biotechnological applications. *Comput Struct Biotechnol J* 2021;**19**:3892–907. https://doi.org/10.1016/j.csbj.2021.06.048
102. Zelezniak A, Andrejev S, Ponomarova O *et al.* Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc Natl Acad Sci* 2015;**112**:6449–54. https://doi.org/10.1073/pnas.1421834112
103. Zomorrodi AR, Segrè D. Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities. *Nat Commun* 2017;**8**:1563. https://doi.org/10.1038/s41467-017-01407-5
104. Kim S, Thapa I, Zhang L *et al.* A novel graph theoretical approach for modeling microbiomes and inferring microbial ecological relationships. *BMC Genomics* 2019;**20**:945. https://doi.org/10.1186/s12864-019-6288-7
105. Hankeln W, Buttigieg PL, Kostadinov I, *et al.* Applying graph theoretic approaches to microbial metagenomes: ecological perspectives on function. *Proc. First ACM Int. Conf. Bioinforma. Comput. Biol.*pp. 478–480. Niagara Falls New York: ACM, 2010.
106. Eng A, Borenstein E. An algorithm for designing minimal microbial communities with desired metabolic capacities. *Bioinformatics* 2016;**32**:2008–16. https://doi.org/10.1093/bioinformatics/btw107
107. Ravikrishnan A, Blank LM, Srivastava S *et al.* Investigating metabolic interactions in a microbial co-culture through integrated modelling and experiments. *Comput Struct Biotechnol J* 2020;**18**:1249–58. https://doi.org/10.1016/j.csbj.2020.03.019
108. Machado D, Maistrenko OM, Andrejev S *et al.* Polarization of microbial communities between competitive and cooperative metabolism. *Nat Ecol Evol* 2021;**5**:195–203. https://doi.org/10.1038/s41559-020-01353-4
109. Machado D, Andrejev S, Tramontano M *et al.* Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res* 2018;**46**:7542–53. https://doi.org/10.1093/nar/gky537
110. Aite M, Chevallier M, Frioux C *et al.* Traceability, reproducibility and wiki-exploration for "à-la-carte" reconstructions of genome-scale metabolic models. *PLoS*

Comput Biol 2018;**14**:e1006146. https://doi.org/10.1371/journal.pcbi.1006146
111. Karp PD, Latendresse M, Paley SM *et al.* Pathway tools version 19.0 update: software for pathway/genome informatics and systems biology. *Brief Bioinform* 2016;**17**:877–90. https://doi.org/10.1093/bib/bbv079
112. Wang H, Marcišauskas S, Sánchez BJ *et al.* RAVEN 2.0: a versatile toolbox for metabolic network reconstruction and a case study on Streptomyces coelicolor. *PLoS Comput Biol* 2018;**14**:e1006541. https://doi.org/10.1371/journal.pcbi.1006541
113. Henry CS, DeJongh M, Best AA *et al.* High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat Biotechnol* 2010;**28**:977–82. https://doi.org/10.1038/nbt.1672
114. Wendering P, Nikoloski Z. COMMIT: consideration of metabolite leakage and community composition improves microbial community reconstructions. *PLoS Comput Biol* 2022;**18**:e1009906. https://doi.org/10.1371/journal.pcbi.1009906
115. Kumar N, Hitch TCA, Haller D *et al.* MiMiC: a bioinformatic approach for generation of synthetic communities from metagenomes. *Microb Biotechnol* 2021;**14**:1757–70. https://doi.org/10.1111/1751-7915.13845
116. Harcombe WR, Riehl WJ, Dukovski I *et al.* Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Rep* 2014;**7**:1104–15. https://doi.org/10.1016/j.celrep.2014.03.070
117. Dama AC, Kim KS, Leyva DM *et al.* BacterAI maps microbial metabolism without prior knowledge. *Nat Microbiol* 2023;**8**:1018–25. https://doi.org/10.1038/s41564-023-01376-0
118. Poore GD, Kopylova E, Zhu Q *et al.* Microbiome analyses of blood and tissues suggest cancerdiagnostic approach. *Nature* 2020;**579**:567–74. https://doi.org/10.1038/s41586-020-2095-1
119. Gihawi A, Ge Y, Lu J *et al.* Major data analysis errors invalidate cancer microbiome findings. *MBio* 2023;**14**:e01607–23. https://doi.org/10.1128/mbio.01607-23
120. Gihawi A, Cooper CS, Brewer DS. Caution regarding the specificities of pan-cancer microbial structure. *Microb Genomics* 2023;**9**:9. https://doi.org/10.1099/mgen.0.001088
121. Wang X-W, Sun Z, Jia H *et al.* Identifying keystone species in microbial communities using deep learning. *Nat Ecol Evol* 2023;**8**:22–31. https://doi.org/10.1038/s41559-023-02250-2
122. Finkel OM, Salas-González I, Castrillo G *et al.* A single bacterial genus maintains root growth in a complex microbiome. *Nature* 2020;**587**:103–8. https://doi.org/10.1038/s41586-020-2778-7
123. Zuñiga C, Li C-T, Yu G *et al.* Environmental stimuli drive a transition from cooperation to competition in synthetic phototrophic communities. *Nat Microbiol* 2019;**4**:2184–91. https://doi.org/10.1038/s41564-019-0567-6
124. Coker J, Zhalnina K, Marotz C *et al.* A reproducible and Tunable synthetic soil microbial community provides new insights into microbial ecology. *mSystems* 2022;**7**:e00951–22. https://doi.org/10.1128/msystems.00951-22
125. Debray R, Herbert RA, Jaffe AL *et al.* Priority effects in microbiome assembly. *Nat Rev Microbiol* 2022;**20**:109–21. https://doi.org/10.1038/s41579-021-00604-w
126. Young TP, Stuble KL, Balachowski JA *et al.* Using priority effects to manipulate competitive relationships in restoration. *Restor Ecol* 2017;25
127. Araus JL, Cairns JE. Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci* 2014;**19**:52–61. https://doi.org/10.1016/j.tplants.2013.09.008
128. Shakoor N, Lee S, Mockler TC. High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. *Curr Opin Plant Biol* 2017;**38**:184–92. https://doi.org/10.1016/j.pbi.2017.05.006
129. Stecher B, Berry D, Loy A. Colonization resistance and microbial ecophysiology: using gnotobiotic mouse models and single-cell technology to explore the intestinal jungle. *FEMS Microbiol Rev* 2013;**37**:793–829. https://doi.org/10.1111/1574-6976.12024
130. Basic M, Bleich A. Gnotobiotics: past, present and future. *Lab Anim* 2019;**53**:232–43. https://doi.org/10.1177/0023677219836715
131. Zengler K, Hofmockel K, Baliga NS *et al.* EcoFABs: advancing microbiome science through standardized fabricated ecosystems. *Nat Methods* 2019;**16**:567–71. https://doi.org/10.1038/s41592-019-0465-0
132. Emmenegger B, Massoni J, Pestalozzi CM *et al.* Identifying microbiota community patterns important for plant protection using synthetic communities and machine learning. *Nat Commun* 2023;**14**:7983. https://doi.org/10.1038/s41467-023-43793-z