Computational and Structural Biotechnology Journal 18 (2020) 4071–4081

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Experimental and computational approaches to unravel microbial community assembly

Daniel Aguirre de Cárcer *

Departamento de Biología, Universidad Autónoma de Madrid, Madrid, Spain

ARTICLE INFO

Article history:
Received 31 August 2020
Received in revised form 16 November 2020
Accepted 19 November 2020
Available online 3 December 2020

Keywords:
Microbial communities
Microbiome
Community assembly
Omics
Computational biology

ABSTRACT

Microbial communities have a preponderant role in the life support processes of our common home planet Earth. These extremely diverse communities drive global biogeochemical cycles, and develop intimate relationships with most multicellular organisms, with a significant impact on their fitness. Our understanding of their composition and function has enjoyed a significant thrust during the last decade thanks to the rise of high-throughput sequencing technologies. Intriguingly, the diversity patterns observed in nature point to the possible existence of fundamental community assembly rules. Unfortunately, these rules are still poorly understood, despite the fact that their knowledge could spur a scientific, technological, and economic revolution, impacting, for instance, agricultural, environmental, and health-related practices. In this minireview, I recapitulate the most important wet lab techniques and computational approaches currently employed in the study of microbial community assembly, and briefly discuss various experimental designs. Most of these approaches and considerations are also relevant to the study of microbial microevolution, as it has been shown that it can occur in ecological relevant timescales. Moreover, I provide a succinct review of various recent studies, chosen based on the diversity of ecological concepts addressed, experimental designs, and choice of wet lab and computational techniques. This piece aims to serve as a primer to those new to the field, as well as a source of new ideas to the more experienced researchers.

© 2020 The Author. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

1. Introduction

In nature, microorganisms most commonly appear as microbial communities; groups of potentially interacting populations that

* Address: Departamento de Biología, Universidad Autónoma de Madrid, Madrid 28049 Spain.
E-mail address: daniel.aguirre@uam.es

https://doi.org/10.1016/j.csbj.2020.11.031
2001-0370/© 2020 The Author. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

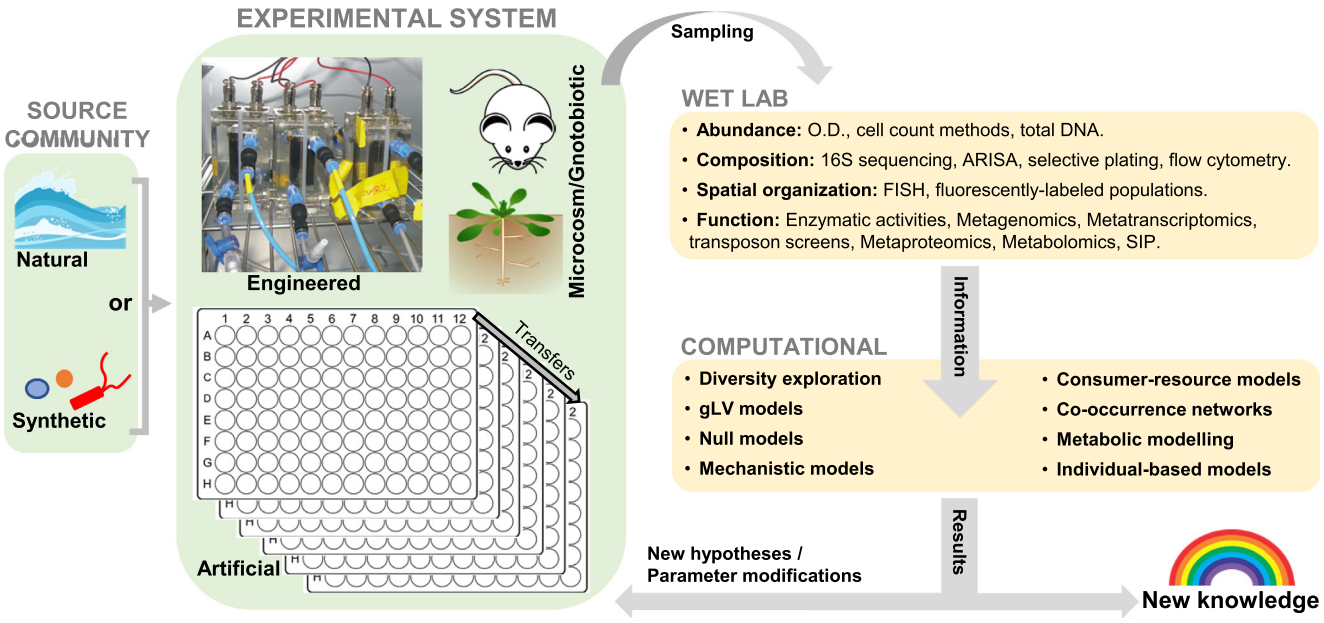

Fig. 1. Diagram depicting the general experimental scheme followed by microbial community assembly studies, pinpointing the most common design and analytical possibilities chosen by researchers.

co-exist in space and time [1], where each population is formed by genetically homogeneous individuals possessing specific genomic features which reflect on their associated life strategies, and to some extent determine their distribution in the environment [2]. The term "community assembly" commonly refers to the sum of all mechanisms that shape the composition of a microbial community, nowadays most often conceptualized as divisible into four basic processes; selection, dispersal, drift, and diversification [1].

Microbial communities account for a large fraction of the Earths biomass and biodiversity, and play essential roles in ecosystem processes [2]. Our understanding of the Earth's microbial communities' structure and function has increased significantly during the last decade as a result of a myriad of surveys, fostered by the arrival and subsequent democratization of high-throughput sequencing technologies. These surveys have uncovered common assembly patterns, such as high diversity and species richness, coexistence of theoretically competing populations, conspicuous functional stability despite large species turnover, and phylogenetic clustering. These common patterns of microbial community organization suggest the existence of fundamental community assembly rules [3]. However, these rules are still very poorly understood [4], despite the fact that their understanding would greatly enhance our ability to manage microbial communities [5]. Hence, understanding the common principles that govern microbial community assembly in nature is a major challenge in microbial ecology [6,7] with large economic and environmental implications.

In this minireview, I recapitulate the most important laboratory (henceforth wet lab) techniques and computational approaches currently employed in the study of microbial community assembly, briefly discuss various experimental setup considerations, and provide a succinct review of selected recent studies (Fig. 1). Many of these approaches, considerations, and studies are also relevant to the study of microbial microevolution, as it has been shown that it can occur in ecological timescales [8].

2. Wet lab methods

In addition to general techniques in microbiology and molecular biology, the study of microbial community assembly often employs a series of common tools and approaches to evaluate community structure and function, which I summarize here.

Absolute population counts. Most studies rely solely on 16S rRNA gene sequence profiling for their analysis. The observed relative abundance of sequence variants, or of groups of sequences clustered at specific phylogenetic depths (i.e. Operational Taxonomic Units; OTUs), indicate the existence of individuals with a shared phylogeny, commonly termed phylotypes. Nevertheless, measuring the actual number of individuals in experimental samples is required not only to fine-tune many experimental setups (e.g. inoculation loads, dilution factor in serial transfer experiments), but can also be used as a measure of overall community function (total biomass), or to feed various computational modelling approaches requiring absolute abundance information [9].

The most straightforward choice is the analysis of cell densities using optical density measurements. While certain cellular traits (e.g. cell adhesion, shape, size) can influence these measurements [10], it remains a fast and inexpensive proxy for cell density. However, the correlation between cell density and optical density holds true in a limited range of conditions, and thus the procedure requires proper calibration [10]. Another more precise yet time-consuming possibility is the use of direct cell counts using a variety of fluorescent stains and a cell-counting chamber [11] or flow cytometry, both of which can also feature live/dead discrimination modifications [12]. Counting colony forming units on solid media can be used as well in some circumstances as a proxy of total cell counts; the key is to work with synthetic communities of known composition whose members can be cultured and discriminated on chosen media. The abovementioned techniques are only suited for liquid samples, or samples were bacteria can be effectively detached from its solid matrix into a liquid sample. When this is

not possible, total DNA extracted from a sample has been used as a proxy of community biomass [13]. Nevertheless, here one needs to take into account the measurement threshold and error associated with the chosen DNA measuring technique, as well as the fact that DNA extraction efficiencies will vary between individual extractions. Finally, qPCR can also be used in certain settings, for instance with host-associated communities by normalizing bacterial gene counts against those of host-derived housekeeping genes [14].

Community composition. Nowadays, microbial community composition is most commonly assessed through the NGS-mediated analysis of 16S rRNA phylogenetic marker genes present in the sample. This gold standard of microbial ecology features the high-throughput sequencing of amplicons obtained using universal primers targeting specific variable regions of the 16S rRNA gene, resulting in the identification and measurement of the relative abundance of phylotypes in a sample (for a recent review on best practices see [15]). Nowadays, such strategy represents a cost-effective method, especially when dealing with a large number of samples of unknown composition. However, due to the high fixed costs and often relatively slow turnaround, the approach is less useful when, for instance, tuning an experimental setup, producing initial results where the amount of samples is likely to be low and a quick turnaround desired, or when the phylotype composition is known *a priori* and its relative abundances can be tracked by other faster and less expensive methods. In this regard, low-complexity communities can be analyzed using low-throughput methods such as Terminal Restriction Fragment Length Polymorphism (T-RFLP) [16], Denaturing Gradient Gel Electrophoresis (DGGE) [17], or Automated Ribosomal Intergenic Spacer Analysis ARISA [18]. These methods present very low resolution when compared to 16S amplicon sequencing, but may represent a convenient way to analyze the overall community stability of complex communities, and will likely be able to provide a semi-quantitative assessment of the relative abundance of known members of simple communities. Overall, these methods have shown a comparable capacity to discern broad scale diversity patterns when compared to high-throughput sequencing methods, but seem to systematically underestimate community richness [19-22].

A common approach in the field when dealing with simple synthetic communities is to use plating on solid media when community members can be discerned by either selective media or colony morphology [23-26]. In this regard, it is worth noting that, more often than not, members of synthetic communities are at least partially selected on the basis of their discernibility by the abovementioned means. Such biased selection may cast a shadow on the actual ecological relevance of the observed community assembly behavior.

Community composition can also be studied using qPCRs with simple communities if specific probes are available or can be designed to discriminate community members [11,27]. Similarly, flow cytometry assays [28] can be used to quantify specifically labeled populations. Finally, for absolute, phylogeny-independent resolution, strains in a synthetic community can be individually tagged before experimentation, and tag-amplicon sequencing or tag-mediated Fluorescence *In Situ* Hybridization (FISH) used to measure their relative abundance and/or spatial organization [29,30].

Spatial organization. Bacterial communities in nature most commonly appear as biofilms [31]. Thus, the study of communities in a spatially structured microscale environment will be central to increase our understanding of ecologically-relevant community assembly. In most cases, omics approaches cannot be applied to microscale samples, providing instead an averaged representation of a composite of microscale sites. Nevertheless, the structured microscale environment can still be interrogated in terms of composition and co-occurrence patterns using fluorescence microscopy. Such spatial mapping allows to better identify interacting populations that might exchange metabolites or signals at scales relevant to bacteria [25,32]. Different fluorescent protein genes can be used to tag experimental populations. The approach also offers the possibility, depending on the experimental settings, of non-destructively imaging the communities and thus observing assembly in real time. However, it has been observed that the burden of fluorescent protein expression can alter fitness values [33], and thus could bias the resulting assembly patterns. On the other hand, FISH can be employed destructively with non-genetically modified strains, commonly allowing the visualization of very few phylotypes in a sample at a time [34,35]. Nevertheless, Valm et al [36] demonstrated that the use of a combinatorial labeling and spectral imaging approach to FISH (CLASI-FISH) could detect fifteen different phylotypes in a human oral microbial community sample.

Community function. Overall community function is an important trait to understand community assembly patterns in a given microbial ecosystem. Researchers commonly measure substrate consumption, biomass production, or respiration rates (e.g. [13]) as a proxy. Also, depending on the ecosystem being considered and the goals of the study, other specific important community functions such as changes on host phenotype [37], ecosystem-relevant enzymatic activities [13] or metabolites [38,39] can be measured. Moreover, it is also possible to gauge community function using a dedicated panel of ecologically-relevant enzymatic reactions [13,40].

Metagenomics, the direct sequencing of all DNA extracted from a sample, provides the opportunity to catalog the set of genes from an entire community, overcoming the well-known culturing approach bias. Although it can only say what functions can potentially be expressed, there is a general good correspondence between gene and transcript relative abundances in microbial communities [41]. Similarly, although sequencing-based approaches predominate in the high-throughput screening of community functions, functional gene-based microarrays can be employed as well (see e.g. [42]).

Metatranscriptomics provides a more accurate assessment of gene expression in the community, and thus a more precise evaluation of community functions. The approach follows similar wet lab and computational pipelines as shotgun metagenomics, but starting with an extraction of total RNA which is immediately copied to DNA. The approach thus pinpoints the genes being actively expressed at a given moment, and hence allows the observation of phenotypic adaptation by the studied community. For instance, a biocontrol rhizobacterium grown on the absence or presence of its target phytopathogenic oomycete showed the same growth rates, and thus a putative shotgun metagenomics approach would have found no differences between the two scenarios. However, the rhizobacterium did experience major transcriptional reprograming in the presence of its target [43].

Another technique that can assess the function of individual populations within a community is the use of genetic transposon screens. The strategy can provide per-gene fitness values for a strain in a community context, and thus has a great potential to reveal the genetic determinants of the strain's biotic and abiotic interactions in a given community and abiotic environment. While this approach can only be employed with strains amenable to genetic manipulation, there is continuing work on specific tools aiming to greatly expand the range of taxa open to such studies [44,45] and in a high-throughput manner [46]. Furthermore, a recent study has opened the road of genome editing within a microbial community context without requiring prior isolation [47]. This technique may be useful in future studies to assess the

role of particular genes on fitness, community function and community assembly.

Metaproteomics can be defined as the large-scale characterization of the entire protein complement of a microbial community at a given point in time [48]. Its goal is to identify and measure the proteins produced by a microbial community and pinpoint their phylotype of origin [49], thus enhancing our knowledge on the functional drivers of community assembly. As for metatranscriptomics, the success of metaproteomics relies heavily on the availability of relevant genomes to increase protein identification rates [50]. When they are unavailable, it is recommended to carry-out a combined metaproteomics and shotgun metagenomics approach, so that the former can employ the draft genomes derived from metagenomics to boost accurate protein identifications [51]. However, while there has been a remarkable development in related wet lab and computational techniques, the protein identification values obtained so far are still modest when considering the number of expected proteins in a given microbial community sample (for a recent review on metaproteomics with a focus on methodological considerations see [51]).

Metabolomics methods can also be applied to study complex microbial communities [52]. Targeted metabolomics approaches are able to adequately detect and quantify panels of metabolites defined *a priori*. On the other hand, untargeted metabolomics aims at describing the entire suite of metabolites in a sample [53]. Its results are commonly limited by the detection of large fractions of unknown metabolites, since metabolite identification relies on comparisons to still sparse reference databases. For instance, over 90% of detected metabolites were classified as unknown in acid mine drainage biofilm communities [54]. Moreover, detected metabolites cannot be directly linked to particular members of the community, and thus the approach is not very informative of biotic interactions *per se*. This shortcoming can be bridged by the use of Stable Isotope Probing (SIP), a technique specifically indicated to query microbial metabolic interactions and specific elemental fluxes [55]. Particular compounds whose fate is of interest to the research question can be labelled with ^{13}C, ^{15}N or ^{18}O, and introduced in the microbial ecosystem. Henceforth, SIP can be combined with any of the above omic approaches [55]. For the most common strategy, density-gradient centrifugation is used to separate the labelled and unlabeled DNA or RNA fractions. Subsequent enrichment analysis of both fractions via targeted metagenomics (e.g. 16S rRNA or specific functional genes) can inform on the identity of the members of the community involved in the degradation of the labeled compound [56].

Table 1
Characteristics of common computational methods used in microbial community assembly studies.

Method	Input data	Goal
Community composition analysis	Community tables[1]	Exploration of diversity patterns
Metagenomics	Shotgun sequences from DNA	Cataloguing genes in community / Reconstruction of community genomes
Metatranscriptomics	Shotgun sequences from RNA	Evaluation of community-level gene expression
Null models	Commonly community tables[1]	Assessment of the stochasticity of selected assembly mechanism
Co-occurrence networks	Community tables[1]	Evaluation of species interactions, alternative community regimes, and keystone species.
Generalized Lotka-Volterra models	Time-series community tables[1]	Prediction of community dynamics
Mechanistic models of metabolite-mediate interactions	Prior knowledge of populations' interactions with metabolite pool	Prediction of community dynamics
Metabolic modelling	Genomic annotations[2]	Prediction of community metabolic interactions
Individual-based models	Pre-defined populations' attributes	Evaluation of community-level emergent properties and patterns
Consumer-resource models	Pre-defined populations' attributes	Prediction of community dynamics

[1]Community tables are most commonly derived from 16S rRNA amplicon sequencing data. [2] Augmented with experimental data when available.

3. Computational methods

Table 1 summarizes the characteristics of the computational methods most commonly used in microbial community assembly studies. As mentioned earlier, the use of 16S rRNA amplicon sequencing is the dominant strategy for the analysis of complex microbial communities. Prior to analysis, the sequences obtained from the chosen NGS platform need to be transformed into a community table describing the relative abundance of phylotypes (either exact sequence variants or OTUs) per sample. Depending on the needs of subsequent analyses, it is also common to obtain a description of the phylogenetic relationships among phylotypes in the form of phylogenetic trees or pairwise distance/dissimilarity values, as well as taxonomic assignments for each phylotype. The series of steps needed for such transformations can be carried out using various bioinformatic pipelines, each presenting different strengths and weaknesses (see [57]). The datasets obtained can then be used to summarize the relative abundance of phylotypes and taxonomic ranks along the experimental samples and groups, as well as to pinpoint statistical differences in particular taxa abundances related to a given experimental variable (e.g. [58]). Moreover, community tables are the raw input to construct co-abundance networks and generalized Lotka-Volterra models (gLV) among other downstream analysis possibilities. Nevertheless, most studies interested in microbial community assembly will commonly use a wide array of techniques to try to infer ecological processes from diversity patterns, which can be classified as multivariate analyses and null models.

Functional profiling is commonly carried out using translated search of metagenomic reads against a comprehensive protein database using dedicated software (e.g. [59]). On the other hand, genome assembly of metagenomic reads can lead to the recovery of draft genomes for the most abundant populations in a sample, which can in turn be used to feed downstream analyses (e.g. metabolic modelling) to increase our understanding of community assembly (for a recent review of best practices in shotgun metagenomics see [60]). If only 16S rRNA sequencing data is available, one can still use dedicated bioinformatic tools to provide a prediction of functional potential of a bacterial community [61], or even predict metabolomic profiles [62] and metabolic interactions [63,64]. The usefulness of metatranscriptomics is boosted if the genomes of the community members are available. In this case, each transcript can be accurately assigned to its original source genome using a read aligner (e.g. [65]), thus moving from an assessment of community function to an individualized functional assessment of each population in the community [27].

Many microbial ecology experiments nowadays produce a large number of variables such as phylotypes, genes, proteins, or metabolites. Thus, a common way to analyze such data is using multivariate statistics. While the number of multivariate statistical techniques is vast and continuously expanding, they can be categorized in three groups [66]; i) Exploratory methods; used to explore relationships between objects based on their variables profiles. ii)

Interpretative methods; where external explanatory variables are used to constrain the community table data. iii) Discriminatory methods; which aim to produce functions that distinguish between objects of different classes (for an extensive description of multivariate analyses see [66]).

Null models generate statistically expected random patterns through random permutations of ecological data by deliberately excluding certain mechanisms of interest (e.g. bioitic interactions or environmental filtering [67], phylogeny-based assembly). When the experimentally observed patterns are undistinguishable from the generated random patterns, community assembly is taken to be stochastic with respect to the mechanism excluded [68]. Neutral models represent a special case of mechanistic null models. Here, species are assumed to present equal fitness, and community structure is modeled on the basis of immigration from the metacommunity, random demographics, and speciation rate [69]. When the neutrality test is rejected, it is commonly taken as indicative of selection (abiotic and/or biotic) playing a significant role in community structure (for a comprehensive review on the use of null models see [4]).

Another common way to analyze community tables is building co-occurrence networks, were nodes represent taxa/phylotypes and edges co-abundance-derived metrics. These networks could be used to identify alternative community regimes, species interactions, and keystone species. However, their interpretation is not straightforward; while they represent an adequate tool to explore ecological associations, the implications of such associations are uncertain, and it is disputed to which degree biotic interactions can be recapitulated using available methods [70–72].

Community dynamics in a multi-species system can be studied using a gLV model; a system of ordinary differential equations with parameters commonly inferred from time-series data (e.g. [73]). Here, each population is parameterized with a growth rate and coefficients describing the strength of its interaction with every other population in the system (for a useful perspective on gLV see [74]). The approach assumes a constant abiotic environment and additive pair-wise interactions. While the researcher must consider whether or not the ecosystem under study represents a constant environment, it is yet unclear if additive pair-wise interactions can truly recapitulate community behavior. In this regard, Friedman et al [75] showed experimentally that additive pair-wise interactions could predict community dynamics to some extent in an eight species synthetic community. On the other hand, Momeni and co-workers [76] showed *in silico* that a gLV system cannot adequately model various common types of interaction mechanisms, and thus may often fail to predict community dynamics. An alternative to population dynamic models such as the gLVs are mechanistic models based on metabolite-mediate interactions. Here, community dynamics are simulated using prior knowledge of the interactions of individual microbes with a shared metabolite pool [76]. The approach has been shown to outperform gLV models [76,77]. However, the required prior knowledge is not easily obtained, but may be retrieved from *in vitro* studies, genome-scale metabolic models, or the literature [77]. Finally, Gaussian processes-based non-parametric modelling approaches are becoming increasingly popular in the analysis of longitudinal omics datasets and external variables for their flexibility and accuracy (e.g. [78]).

The metabolic modelling of microbial communities allows to simulate the metabolic interactions between its members, and can help explain observed diversity patterns and generate novel hypotheses regarding community assembly (e.g. [79]). The general pipeline for a single population starts with the use of its genome annotation to construct the corresponding metabolic network. The ensuing network is then constrained using knowledge from biophysics and environmental conditions under the steady-state assumption. The resulting constraint-based model [80] is commonly processed using Flux Balance Analysis (FBA) [81] by adding a declaration of an objective function. The particular choice (commonly maximization of growth rate) is subjective and will influence the resulting flux distribution [82]. For a given input list of possible community members, the same overall approach can be assayed to predict its community composition on a given abiotic environment, and the metabolic fluxes that sustain it [83]. However, the application of FBA to microbial communities is challenging, since its underlying assumptions (i.e. steady-state, single objective function) are oversimplistic in a community setting. This fact has led to the development of several specific tools for the metabolic modelling of microbial communities, which could be divided into quantitative and large-scale methods [84]. The former present significant predictive potential but are limited to simpler systems (e.g. Dynamic Flux Balance Analysis [85,86]), while the latter provide rather qualitative insights but can be used to model complex communities (e.g. [87,88]).

The metabolic modelling of microbial communities has, in many instances, allowed to accurately predict microbial interactions, dynamics, and evolution [79,89,90], but with exceptions [89]. Originally, the production of per-taxon high-quality metabolic models required time-consuming manual curation. Fortunately, enhanced automated tools that produce quality metabolic models from genome annotations alone are now available [91]. Finally, the metabolic modelling of a microbial community can be refined using metaproteomics and metabolomics data, if available [92].

The individual-based models simulate the discrete behavior of each individual in a community, and their collective action determines the emergent properties of the community as a whole. The approach seeks to model the properties, activities and interactions of all individuals in a community [93]. During simulation, individuals progress to the next iteration according to probabilistic rules arising from the model's parameters. In this regard, the approach can accommodate a wide range of parametrized mechanisms, such as mutation rates or metabolic states, and is suitable to describe the modeled community at different temporal and spatial resolutions [94]. Significantly, their ability to give rise to different assembly outcomes depending on local context is most useful to the study of microbial community assembly, since biotic interactions most commonly occur at the microscale [32]. For instance, Bauer et al [95] integrated individual based modelling with metabolic models to simulate the behavior of a seven species model community of the human gut. Among other things, their results indicated that spatial gradients of mucus glycans shaped community structure.

Closing this concise review of common computational methods used to unravel microbial community assembly are the consumer-resource models. Here, different taxa have different resource preferences, and community dynamics are defined on the basis of resource consumption rates. While this approach cannot be used to explicitly model any particular microbial community, it can be used to investigate if mechanistic assumptions about ecosystem workings can recapitulate experimentally observed patterns. For instance, Goldford et al [3] used a generic consumer-resource model to explain how secretion-mediated interactions could allow the experimentally-observed coexistence of a bacterial consortium on a single simple carbon source. More recently, Marsland and co-workers have developed an enhanced microbial consumer-resource model which incorporates metabolic cross-feeding and stochastic colonization [96].

4. Experimental systems

Microbial community assembly can be studied using different experimental scenarios that can be roughly classified into three categories; natural ecosystems, microcosms, and artificial systems. These categories can largely be positioned in that order along the representativity vs. flexibility-tractability axis. In the last few years, the study of community assembly in natural ecosystems has shifted its focus from a descriptive "who is there?" approach, to elucidating the relative contribution of stochastic vs. deterministic processes on community assembly (e.g. [97–99]). Both approaches are indeed pertinent to increase our understanding of the particular ecosystem under investigation. However, they are less useful for the goal of unraveling a set of common assembly rules in microbial communities. Nonetheless, these studies provide valuable coarse-grained assembly patterns (e.g. richness values, β-diversity patterns, phylogenetic signal) that need to be explainable by putative common assembly rules put forward by the research community (e.g. [100]). Microcosms, on the other hand, consist of environmental samples cultured in the lab under well-defined conditions [9]. These systems strive to be as similar to its natural counterpart as possible, yet allowing for increased experimental flexibility-tractability (e.g. [23]). A highly related concept is that of gnotobiotic organisms, whose microbiome only contains a defined experimental community (e.g. [101]). Lastly, artificial systems, while unrelated to microbial communities in nature, allow for high replication and tightly controlled experimental parameters.

Most experiments on microbial community assembly are carried out in a homogeneous liquid medium. To increase throughput, allow for increased replication and reduce costs, many studies are nowadays conducted using microtiter well plates. On the other hand, the small volume employed makes the control of abiotic factors more difficult, and imposes limitations on the amount of sampling that can be achieved. While in itself growth in a microtiter well represents batch growth, these containers have been employed in setups mimicking a continuous culture by performing serial transfers of the community into fresh media until a stable composition is reached [3,13,79,102,103]. During each growth-dilution cycle, cells are grown for a fixed period of time, then diluted using a constant factor and inoculated into fresh media. In this approach, the per-cycle incubation period and dilution factor need to be previously fine-tuned depending on the goal of the study. Most commonly, the sole criterion for this tuning is the ability to obtain a stable and diverse composition after a number of transfers. It is important to note that the community dynamics recorded during such serial transfer experiments will be influenced by fluctuating nutrient concentrations. For instance, Estrela et al [79] were able to obtain stable compositions dominated by *Pseudomonadaceae* and *Enterobacteriaceae* when performing *ex situ* cultivation of complex natural communities on glucose and minimal media. However, they also reported that each of these groups had a significant growth advantage at different periods of each cycle related to the abovementioned fluctuations. Using a different approach, the study of biofilms offers a wealth of research data and associated technical experience on specific model communities [11,34,38], as well as experience in evaluating spatially-structured local community assembly [25,104].

Microbial community assembly experimentation can be undertaken using natural complex communities or simple synthetic consortia. For the first strategy, different initial diversities can be attained by serial dilutions of the initial inoculant [39,105]. Migration can be artificially imposed by first pooling species from each replicate community to form a regional species pool, and then inoculate a percentage of this pool into each replicate community when desired [79]. Moreover, if replicate communities share the same abiotic space, such as when microbial communities are grown on beads introduced in the same reactor [6], replicate communities conform a natural metacommunity, with the down side that migration cannot be controlled. On the other hand, experimental setups featuring synthetic consortia can directly modify the initial diversity and migration rates [106] if desired. However, bear in mind that the use of synthetic consortia heavily influences the experimental representativity, and requires an initial culturomics effort to obtain the needed strains that will form the consortia. Interestingly, Kehe et al recently developed a droplets-based platform for the fast, massively parallel construction and screening of combinatorial synthetic communities [107]. Since natural communities can serve as the input species pool, their approach can somehow be regarded as standing in between the above two strategies.

When a mid-term goal of a research group is to build an increased understanding of a particular microbial ecosystem using microcosms, or when such ecosystem has merely been chosen as a model to study community assembly, obtaining a rich and representative collection of strains derived from the represented natural ecosystem using high-throughput cultivation (i.e. culturomics) may be a sound investment. Clearly, the choice of model community would incur in trade-offs related to various criteria such as financial constraints, representativity, tractability, stability, and size [108,109]. Nevertheless, such collection would permit experimentation using representative synthetic consortia, allowing for greatly increased tractability and monitoring of the community.

Culturomics approaches may employ various strategies based on different culturing media and conditions [110]. Furthermore, when a particular strain has been targeted for cultivation, there are bioinformatic approaches to reverse-engineer a suitable culturing media, such as using its 16S rRNA gene sequence [111]. For instance, Hartman et al. [112] reported the production of an isolate collection of 200 members spanning around 25% of the abundant members of the rhizosphere of clover using only one specific media. Bai and co-workers [113] employed various different media with the limiting dilution method, together with cell sorting and high-throughput sequencing of isolates, to obtain 7943 isolates from the *Arabidopsis* microbiome, while Zhang et al [114] developed an improved identification method using Illumina HiSeq sequencing which allowed them to characterize 70% of the rice root microbiome members. Following a different rationale, Lozano et al [115] proposed a three-member consortium of genetically and biochemically-tractable strains from the dominant rhizosphere bacterial phyla showing various emergent properties as a means to further explore community ecology.

More common in the literature is the use of *post hoc* culturomics, where researchers focus on isolating a few strains selected on the basis of previous experimental results with complex natural communities, with the goal of performing follow-up experiments to complement or substantiate previous observations and hypotheses [3,6,13]. One such common follow-up experiment is the use of conditioned media to probe pairwise interspecies interactions [3,6,39]. Here the spent media obtained after growth of a donor strain, depleted for specific resources and containing new secreted metabolites, is used to measure growth of a recipient strain.

5. Representative studies

What follows is a series of cherry-picked recent studies, chosen on the basis of the variety and relevance of the experimental setups developed and the wet lab and computational tools employed (see Table 2 for a summary). As an example of a comprehensive

Table 2
Characteristics of selected studies focusing on microbial community assembly.

Experimental setup	Wet lab methods	Computational methods
• Artificial **Ref:** [3,79] • Microtiter plate-based • High replication • Simple media • Serial transfers • Different natural source communities • Migration (forced) • Time series	• 16S rRNA gene profiling • Targeted metabolomics • Conditioned media • Culturomics (*post hoc*)	• Diversity exploration • Null model • Consumer-resource model • Metabolic model • Functional prediction
• Natural (Soil) **Ref:** [97] • High replication • Wide range of environmental conditions	• 16S rRNA gene profiling • ITS profiling	• Diversity exploration • Null model • Neutral model • Co-occurrence Network
• Microcosm (Pitcher plant) **Ref:** [13] • Microtiter plate-based • Realistic complex medium • Serial transfers • Different natural source communities • Time series	• 16S rRNA gene profiling • EcoPlates • Culturomics (*post hoc*)	• Diversity exploration • Null models
• Artificial (Phycosphere) **Ref:** [102] • Microtiter plate-based • Realistic simple media combinations • Serial transfers • Single natural source community	• 16S rRNA gene profiling • Metagenome sequencing and assembly • Untargeted metabolomics	• Diversity exploration • Weighted sum model
• Artificial (Human gut) **Ref:** [103] • Microtiter plate-based • Rich medium • Serial transfers • Strain collection combinations • Time series	• 16S rRNA gene profiling • Conditioned media • Targeted Metabolomics	• Co-occurrence Network • gLV
• Artificial (Marine particles) **Ref:** [6] • Hydrogel beads in shared reactor • Realistic simple media combinations • Single natural source community • Migration (natural) • Time series	• 16S rRNA gene profiling • Targeted Metabolomics • Conditioned media • Culturomics (*post hoc*)	• Diversity exploration
• Artificial (Biofilm) **Ref:** [118] • Drip-flow biofilm • Rich medium • Strain collection combinations	• Metatranscriptomics • FISH	• Differential gene expression • Co-localization analysis
• Microcosm (Biofilm) **Ref:** [38] • Carriers in shared reactor • Realistic medium • Physical constraints • Single natural source community • Migration (natural)	• 16S rRNA gene profiling • FISH	• Diversity exploration • OTU significance testing • Null model

study of a natural microbial ecosystem, Jiao et al analyzed microbial diversity in adjacent pairs of rice (flooded) and maize (not flooded) fields in 127 sites distributed along a wide range of habitats and regions [97]. In addition to the more common β-diversity analyses, they used null model analysis, neutral models, and co-occurrence networks to assess the relative importance of community assembly processes in soils linked to measured abiotic environments and geographic patterns. Following similar goals and overall analytical framework, Logares and co-workers [99] assessed the drivers of prokaryotic and picoeukaryotic community structure in the sunlit ocean.

Cui et al studied the apple flower stigma [116], representing an interesting experimental system to probe community assembly hypotheses in a natural setting. When flowers bloom, immigrating microbes are received by a nutrient-rich virgin environment. Sampling the stigma from individual flowers likely provided a closer representation of a local community than most other natural studies. Moreover, samples represented relevant host-associated communities that can be more easily manipulated and monitored than, for instance, the rhizosphere or the gut environments.

Martinez and co-workers used an effective design to study historical contingencies in the mice gut [117]. Germ-free mice were sequentially inoculated in alternating order with different complex source communities, or a simplified synthetic community and a complex source community. Using this simple approach, the authors uncovered both historical contingency effects and priority effects in the assembly of mice gut communities. Carlström et al also focused on assessing historical contingency and priority effects in the phylosphere of the model plant *Arabidopsis thaliana* [58]. However, they chose a different approach including only synthetic communities built from a pool of 62 native strains. Here, they performed drop-out experiments, measuring the effect of the absence of the selected taxonomic group or strain on the rest of the community, as well as late introduction experiments where the taxonomic group or strain that had been left out was later introduced.

As mentioned earlier, biofilms represent an excellent resource in the study of community assembly. For instance, Liu et al [118] studied the interactions within three and four-species biofilms using metatranscriptomics and FISH, and concluded that the observed increase in biomass productivity of the four-species consortium emerged from a fine-tuned microscale distribution of each species driven by concerted biotic interactions. Another example is Suarez and co-workers' study of community assembly on nitrifying

biofilms [38]. Their setup featured the same source community and (uncontrolled) immigration on their replicated experimental communities. The main variable assessed was the thickness constraint physically imposed on the community. In addition to community composition exploration through the use of β-diversity analyses and null models, they also studied the 3D assembly of the communities using FISH.

Venturelli et al [103] employed a robot-assisted serial transfer setup on microtiter plates to study a twelve-members bacterial collection derived from the human gut. The authors first measured individual growth rates in mono and co-cultures. Subsequently, they measured each member's growth rate in the complete community, as well as in drop-out experiments. With dense time-resolved data, they developed a model to predict higher-dimensional community structure from lower-order assembly data. Their results support the notion that pairwise interactions are major drivers of complex community dynamics, an idea previously put forward by Friedman et al's experiments [75].

Enke and co-workers [6] employed minute (*ca.* 100 μm) paramagnetic hydrogel beads made with chosen substrates to study microbial community assembly on marine polysaccharide particles. Their experimental setup had various interesting characteristics; the small size of the community growing onto each bead could mean that each bead contained a single local community. Thus, a single-bead sampling technique could be devised allowing for a greatly reduced noise in studies of microbial interactions. The beads where immersed in the same milieu, and hence were subjected to the same abiotic conditions. Moreover, all beads within the recipient naturally formed an actual metacommunity, with the possibility of immigration between local communities and a shared regional species pool. Also, sampling at different times allowed following successional dynamics. Featuring essentially the same characteristics, Leventhal et al studied community assembly on granular biofilms [119].

Goldford et al [3] and Estrela et al [79] employed the same microtiter plate serial transfer strategy to study microbial community assembly of large and complex microbial communities on single carbon and energy sources. The system allowed for easy high-replication and handling due to the microtiter plate support. It also allowed to modify the abiotic selective pressures [3], as well as the use of various different initial community sources, and the ability to impose different levels of migration from a regional species pool [79]. In this regard, the experimental communities behaved as a metacommunity except for the fact that immigration was not dependent on the microbes' dispersal capabilities or natural mass transfers, but on experimentally producing a regional pool and forced immigration at each transfer. With such simple yet flexible setup, both complementing studies showed how the observed reproducible community assembly results presented significant phylogenetic signal and reflected an emergent metabolic structure which could be predicted by genome-scale metabolic models.

Oliphant et al [120] employed replicate bioreactors with different media mimicking different human dietary regimes to grow a defined synthetic community formed by bacterial strains obtained from the same human faecal sample, which then were profiled by 16S sequencing and metabolomics. Once the behavior of this initial control community was established, the authors recreated the same defined community but with each strain arising from different donors, with the goal of assessing whether coadaptation between strains, as expected to be present in the first community, influenced community dynamics. In this elegant yet simple manner, their results pointed, for instance, to polysaccharide utilization by Firmicutes being dependent on coadaptation.

Bittleston et al [13] studied the microbial community of wild carnivorous pitcher plants. The rationale for such unconventional choice was that it represents a plant associated microbiome with an evident community function, degradation of insect biomass, that can be evaluated in the lab as the activity of a series of specific enzymatic reactions (e.g. chitinase). The authors produced synthetic pitcher plants microcosms using a realistic complex media (ground crickets) in a serial transfers setting. With this approach and an exhaustive analysis of time series data, they found that community structure was contingent on early community states; the same strain could have different dynamics in different community contexts. Also, early community dynamics could translate into different microbial regimes that, despite convergence on overall community function (i.e. respiration rates), had different functional repertoires.

Closing this brief list is Fu and co-workers' study of community assembly on the phycosphere [102]; the diffusive boundary surrounding phytoplankton cells. Instead of following the same approach as Bittleston et al [13], and thus employ a realistic complex media derived from phytoplankton, the authors used synthetic media composed of different combinations of five exometabolites common in the phycosphere. These different media were used to grow a single natural community in what was essentially a serial transfer experiment. With this setup, they showed that community assembly could be predicted as a simple sum of assemblages supported by each individual metabolite, hence supporting similar previous results by Enke et al [6].

6. Summary and outlook

There is a constant rise in the number of studies featuring lab-based microbial community assembly experimentation. As succinctly reviewed above, these studies are ever more sophisticated in their experimental design and analytical strategies. The field has shown a noticeable increasing interest in adequately combining wet lab and computational approaches, but still requires further collaborative efforts between microbial ecologists, physiologists, and computational biologists.

Since biotic interactions mostly occur at the local scale [32], future experimental designs and development of wet lab techniques should strive to obtain experimental data from discrete local communities. Similarly, immigration and micro-evolution are potentially important drivers of community structure, and as such should be taken into account when considering experimental design (e.g. [121]). Furthermore, future experimental designs should strive towards increased experimental replication and highly time-resolved data to better understand microbial community dynamics, and the potential role of higher order interactions in determining community assembly. Finally, since in nature the bacterial kingdom seldom appears alone, microbial community assembly studies should take more into account potential multi-kingdom interactions, significantly phage-bacteria interactions where density-dependent effects and co-evolution are important drivers of community dynamics [122,123].

A promising path rests in the use of microfluidics or microdroplet-based systems [124,125] as vehicles for local community assembly and sampling. These systems can also effectively allow the analysis of structured community assembly, thus facilitating the study of biotic interactions and local assembly [126]. Moreover, their increased throughput and sampling possibilities may allow to sample a larger parameter space in a single experiment [127]. Another innovative possibility is the use of advanced functional materials, not only as a physical container and support for the community or to facilitate other experimental features such as sampling or migration, but also as a means to manipulate and interact with the communities *in vivo* [128].

Finally, a much awaited technique still in need of robust protocols is the mass spectrometry imaging of microbial communities.

The approach is capable of interrogating chemically complex systems with high spatial specificity, and thus is well positioned to provide complementary information to many of the wet lab and computational approaches enumerated in this piece [129]. Another emerging technique allows for the chronological labelling of living bacteria. The approach is based on the incorporation of fluorescent D-amino acids (FDAAs) into the peptidoglycans of metabolically active bacteria. Following this approach, Wang et al assessed the viability of transplanted microbiotas [130]; the authors first administered a FDAA to the donor mice, the labeled microbiota was transplanted to recipient mice which then received a second FDAA with a different fluorophore. The surviving transplanted populations could be differentiated by presenting both fluorophores simultaneously.

CRediT authorship contribution statement

Daniel Aguirre de Cárcer: Conceptualization, Writing - original draft, Funding acquisition.

Declaration of Competing Interest

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

References

[1] Vellend M. Conceptual synthesis in community ecology. Q Rev Biol 2010;85:183–206.
[2] Falkowski PG, Fenchel T, Delong EF. The microbial engines that drive Earth's biogeochemical cycles. Science 2008;320:1034–9.
[3] Goldford JE, Lu N, Bajić D, Estrela S, Tikhonov M, et al. Emergent simplicity in microbial community assembly. Science 2018;361:469–74.
[4] Zhou J, Ning D. Stochastic community assembly: does it matter in microbial ecology?. Microbiol Mol Biol Rev 2017;81:00002–17.
[5] Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, et al. A communal catalogue reveals Earth's multiscale microbial diversity. Nature 2017;551:457–63.
[6] Enke TN, Datta MS, Schwartzman J, Cermak N, Schmitz D, et al. Modular assembly of polysaccharide-degrading marine microbial communities. Curr Biol 2019;29:1528–35.
[7] Lurgi M, Thomas T, Wemheuer B, Webster NS, Montoya JM. Modularity and predicted functions of the global sponge-microbiome network. Nat Commun 2019;10:992.
[8] Castledine M, Sierocinski P, Padfield D, Buckling A. Community coalescence: an eco-evolutionary perspective. Philos Trans R Soc Lond B Biol Sci 2020;375:23.
[9] Widder S, Allen RJ, Pfeiffer T, Curtis TP, Wiuf C, et al. Challenges in microbial ecology: building predictive understanding of community function and dynamics. ISME J 2016;10:2557–68.
[10] Stevenson K, McVey AF, Clark IBN, Swain PS, Pilizota T. General calibration of microbial growth in microplate readers. Sci Rep 2016;6:38828.
[11] Kinnunen M, Dechesne A, Albrechtsen H-J, Smets BF. Stochastic processes govern invasion success in microbial communities when the invader is phylogenetically close to resident bacteria. ISME J 2018;12:2748–56.
[12] Robertson J, McGoverin C, Vanholsbeeck F, Swift S. Optimisation of the protocol for the LIVE/DEAD® BacLight™ bacterial viability kit for rapid determination of bacterial load. Front Microbiol 2019;10:801.
[13] Bittleston LS, Gralka M, Leventhal GE, Mizrahi I, Cordero OX. Context-dependent dynamics lead to the assembly of functionally distinct microbial communities. Nat Commun 2020;11:1440.
[14] Bodenhausen N, Bortfeld-Miller M, Ackermann M, Vorholt JA. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. PLoS Genet 2014;10:e1004283.
[15] Pollock J, Glendinning L, Wisedchanwet T, Watson M. The madness of microbiome: attempting to find consensus "Best Practice" for 16S microbiome studies. Appl Environ Microbiol 2018;84:e02627–e2717.
[16] Prakash O, Pandey PK, Kulkarni GJ, Mahale KN, Shouche YS. Technicalities and glitches of terminal restriction fragment length polymorphism (T-RFLP). Indian J Microbiol 2014;54:255–61.
[17] Siqueira JF, Jr., Sakamoto M, Rosado AS (2017) Microbial Community Profiling Using Terminal Restriction Fragment Length Polymorphism (T-RFLP) and Denaturing Gradient Gel Electrophoresis (DGGE). Methods Mol Biol: 6685-1_8.
[18] Bodenhausen N, Bortfeld-Miller M, Ackermann M, Vorholt JA. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. PLoS Genet 2014;10:e1004283-e.
[19] Jami E, Shterzer N, Mizrahi I. Evaluation of automated ribosomal intergenic spacer analysis for bacterial fingerprinting of rumen microbiome compared to pyrosequencing technology. Pathogens 2014;3:109–20.
[20] Gobet A, Boetius A, Ramette A. Ecological coherence of diversity patterns derived from classical fingerprinting and Next Generation Sequencing techniques. Environ Microbiol 2014;16:2672–81.
[21] Johnston-Monje D, Lopez Mejia J. Botanical microbiomes on the cheap: inexpensive molecular fingerprinting methods to study plant-associated communities of bacteria and fungi. Appl Plant Sci 2020;8:e11334.
[22] van Dorst J, Bissett A, Palmer AS, Brown M, Snape I, et al. Community fingerprinting in a sequencing world. FEMS Microbiol Ecol 2014;89:316–30.
[23] Lawrence D, Fiegna F, Behrends V, Bundy JG, Phillimore AB, et al. Species interactions alter evolutionary responses to a novel environment. PLoS Biol 2012;10:e1001330.
[24] Liu W, Jacquiod S, Brejnrod A, Russel J, Burmølle M, et al. Deciphering links between bacterial interactions and spatial organization in multispecies biofilms. ISME J 2019;13:3054–66.
[25] Tan CH, Lee KWK, Burmølle M, Kjelleberg S, Rice SA. All together now: experimental multispecies biofilm model systems. Environ Microbiol 2017;19:42–53.
[26] Celiker H, Gore J. Clustering in community structure across replicate ecosystems following a long-term bacterial evolution experiment. Nat Commun 2014;5:4643.
[27] D'Hoe K, Vet S, Faust K, Moens F, Falony G, et al. Integrated culturing, modeling and transcriptomics uncovers complex interactions and emergent behavior in a three-species synthetic gut community. Elife 2018;7:e37090.
[28] Koch C, Harms H, Müller S. Dynamics in the microbial cytome—single cell analytics in natural systems. Curr Opin Biotechnol 2014;27:134–41.
[29] Kebschull JM, Zador AM. Cellular barcoding: lineage tracing, screening and beyond. Nat Meth 2018;15:871–9.
[30] Cira NJ, Pearce MT, Quake SR. Neutral and selective dynamics in a synthetic microbial community. Proc Natl Acad Sci U S A 2018;115:E9842–8.
[31] Flemming HC, Wuertz S. Bacteria and archaea on Earth and their abundance in biofilms. Nat Rev Microbiol 2019;17:247–60.
[32] Cordero OX, Datta MS. Microbial interactions and community assembly at microscales. Curr Opin Microbiol 2016;31:227–34.
[33] Yang D-D, Alexander A, Kinnersley M, Cook E, Caudy A, et al. Fitness and productivity increase with ecotypic diversity among Escherichia coli strains that coevolved in a simple, constant environment. Appl Environ Microbiol 2020;86:e00051–e120.
[34] Liu W, Russel J, Burmølle M, Sørensen SJ, Madsen JS. Micro-scale intermixing: a requisite for stable and synergistic co-establishment in a four-species biofilm. ISME J 2018;12:1940–51.
[35] Røder HL, Herschend J, Russel J, Andersen MF, Madsen JS, et al. Enhanced bacterial mutualism through an evolved biofilm phenotype. ISME J 2018;12:2608–18.
[36] Valm AM, Mark Welch JL, Borisy GG. CLASI-FISH: principles of combinatorial labeling and spectral imaging. Syst Appl Microbiol 2012;35:496–502.
[37] Lynch JB, Hsiao EY. Microbiomes as sources of emergent host phenotypes. Science 2019;365:1405.
[38] Suarez C, Piculell M, Modin O, Langenheder S, Persson F, et al. Thickness determines microbial community structure and function in nitrifying biofilms via deterministic assembly. Sci Rep 2019;9:019–41542.
[39] Zhang Z, Deng Y, Feng K, Cai W, Li S, et al. Deterministic assembly and diversity gradient altered the biofilm community performances of bioreactors. Environ Sci Technol 2019;53:1315–24.
[40] Aguirre de Cárcer D, Martín M, Mackova M, Macek T, Karlson U, et al. The introduction of genetically modified microorganisms designed for rhizoremediation induces changes on native bacteria in the rhizosphere but not in the surrounding soil. Isme J 2007;1:215–23.
[41] Franzosa EA, Morgan XC, Segata N, Waldron L, Reyes J, et al. Relating the metatranscriptome and metagenome of the human gut. Proc Natl Acad Sci U S A 2014;111:19.
[42] Tu Q, Yu H, He Z, Deng Y, Wu L, et al. GeoChip 4: a functional gene-array-based high-throughput environmental technology for microbial community analysis. Mol Ecol Resour 2014;14:914–28.
[43] Tomada S, Sonego P, Moretto M, Engelen K, Pertot I, et al. Dual RNA-Seq of Lysobacter capsici AZ78 – Phytophthora infestans interaction shows the implementation of attack strategies by the bacterium and unsuccessful oomycete defense responses. Environ Microbiol 2017;19:4113–25.
[44] Wiles TJ, Wall ES, Schlomann BH, Hay EA, Parthasarathy R, et al. Modernized tools for streamlined genetic manipulation and comparative study of wild and diverse proteobacterial lineages. MBio 2018;9:e01877–e1918.
[45] Liu H, Price MN, Waters RJ, Ray J, Carlson HK, et al. Magic pools: parallel assessment of transposon delivery vectors in bacteria. mSystems 2018;3:00143–217.
[46] Wetmore KM, Price MN, Waters RJ, Lamson JS, He J, et al. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. MBio 2015;6:00306–315.
[47] Rubin BE, Diamond S, Cress BF, Crits-Christoph A, He C, et al. Targeted genome editing of bacteria within microbial communities. bioRxiv 2020. 2020.07.17.209189.

[48] Wilmes P, Bond PL. The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. Environ Microbiol 2004;6:911–20.
[49] Hagen LH, Frank JA, Zamanzadeh M, Eijsink VGH, Pope PB, et al. Quantitative metaproteomics highlight the metabolic contributions of uncultured phylotypes in a thermophilic anaerobic digester. Appl Environ Microbiol 2017;83:e01955–e2016.
[50] Siggins A, Gunnigle E, Abram F. Exploring mixed microbial community functioning: recent advances in metaproteomics. FEMS Microbiol Ecol 2012;80:265–80.
[51] Kunath BJ, Minniti G, Skaugen M, Hagen LH, Vaaje-Kolstad G, et al. Metaproteomics: sample preparation and methodological considerations. Adv Exp Med Biol 2019;1073:187–215.
[52] Malik AA, Swenson T, Weihe C, Morrison EW, Martiny JBH, et al. Drought and plant litter chemistry alter microbial gene expression and metabolite production. ISME J 2020;14:2236–47.
[53] Orešič M. Metabolomics, a novel tool for studies of nutrition, metabolism and lipid dysfunction. Nutr Metab Cardiovas Dis 2009;19:816–24.
[54] Mosier AC, Justice NB, Bowen BP, Baran R, Thomas BC, et al. Metabolites associated with adaptation of microorganisms to an acidophilic, metal-rich environment identified by stable-isotope-enabled metabolomics. MBio 2013;4:e00484–e512.
[55] Abram F. Systems-based approaches to unravel multi-species microbial community functioning. Comp Struct Biotechnol J 2015;13:24–32.
[56] Neufeld JD, Wagner M, Murrell JC. Who eats what, where and when? Isotope-labelling experiments are coming of age. ISME J 2007;1:103–10.
[57] Prodan A, Tremaroli V, Brolin H, Zwinderman AH, Nieuwdorp M, et al. Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. PLoS One 2020;15:e0227434.
[58] Carlström CI, Field CM, Bortfeld-Miller M, Müller B, Sunagawa S, et al. Synthetic microbiota reveal priority effects and keystone strains in the Arabidopsis phyllosphere. Nat Ecol Evol 2019;3:1445–54.
[59] Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Meth 2018;15:962–8.
[60] Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 2017;35:833–44.
[61] Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, et al. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol. 2020;38:685–8.
[62] Mallick H, Franzosa EA, McLver LJ, Banerjee S, Sirota-Madi A, et al. Predictive metabolomic profiling of microbial communities using amplicon or metagenomic sequences. Nat Commun 2019;10:3136.
[63] Mendes-Soares H, Mundy M, Soares LM, Chia N. MMinte: an application for predicting metabolic interactions among the microbial species in a community. BMC Bioinf 2016;17:343.
[64] Diener C, Gibbons SM, Resendis-Antonio O. MICOM: metagenome-scale modeling to infer metabolic interactions in the gut microbiota. mSystems 2020;5. e00606-e619.
[65] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Meth 2012;9:357–9.
[66] Paliy O, Shankar V. Application of multivariate statistical techniques in microbial ecology. Mol Ecol 2016;25:1032–57.
[67] Gotelli NJ, Ulrich W. Statistical challenges in null model analysis. Oikos 2012;121:171–80.
[68] Chase JM, Myers JA. Disentangling the importance of ecological niches from stochastic processes across scales. Philos Trans R Soc Lond B Biol Sci 2011;366:2351–63.
[69] Hubbell SP. Neutral theory and the evolution of ecological equivalence. Ecology 2006;87:1387–98.
[70] Hirano H, Takemoto K. Difficulty in inferring microbial community structure based on co-occurrence network approaches. BMC Bioinf 2019;20:019–2915.
[71] Lv X, Zhao K, Xue R, Liu Y, Xu J, et al. Strengthening insights in microbial ecological networks from theory to applications. mSystems 2019;4: e00124–e219.
[72] Röttjers L, Faust K. From hairballs to hypotheses-biological insights from microbial networks. FEMS Microbiol Rev 2018;42:761–80.
[73] Kuntal BK, Gadgil C, Mande SS. Web-gLV: a web based platform for Lotka-Volterra based modeling and simulation of microbial populations. Front Microbiol 2019;10:288.
[74] Gonze D, Coyte KZ, Lahti L, Faust K. Microbial communities as dynamical systems. Curr Opin Microbiol 2018;44:41–9.
[75] Friedman J, Higgins LM, Gore J. Community structure follows simple assembly rules in microbial microcosms. Nat Ecol Evol 2017;1:017–0109.
[76] Momeni B, Xie L, Shou W. Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions. Elife 2017;6:e25051.
[77] Brunner JD, Chia N. Metabolite-mediated modelling of microbial community dynamics captures emergent behaviour more effectively than species–species modelling. JRSoc Interface 2019;16:20190423.
[78] Cheng L, Ramchandran S, Vatanen T, Lietzén N, Lahesmaa R, et al. An additive Gaussian process regression model for interpretable non-parametric analysis of longitudinal data. Nat Commun 2019;10:019–09785.
[79] Estrela S, Vila JCC, Lu N, Bajic D, Rebolleda-Gomez M, et al. (2020) Metabolic rules of microbial community assembly. bioRxiv: 2020.03.09.984278.
[80] Varma A, Palsson BO. Metabolic capabilities of Escherichia coli II. Optimal growth patterns. J Theor Biol 1993;165:503–22.
[81] Orth JD, Thiele I, Palsson B. What is flux balance analysis?. Nat Biotechnol 2010;28:245–8.
[82] Schuetz R, Kuepfer L, Sauer U. Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. Mol Syst Biol 2007;3:10.
[83] Succurro A, Ebenhöh O. Review and perspective on mathematical modeling of microbial ecosystems. Biochem Soc Trans 2018;46:403–12.
[84] Bosi E, Bacci G, Mengoni A, Fondi M. Perspectives and challenges in microbial communities metabolic modeling. Front Genet 2017;8.
[85] Harcombe WR, Riehl WJ, Dukovski I, Granger BR, Betts A, et al. Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. Cell Rep 2014;7:1104–15.
[86] Brunner JD, Chia N. Minimizing the number of optimizations for efficient community dynamic flux balance analysis. PLoS Comput Biol 2020;16: e1007786.
[87] Tobalina L, Bargiela R, Pey J, Herbst FA, Lores I, et al. Context-specific metabolic network reconstruction of a naphthalene-degrading bacterial community guided by metaproteomic data. Bioinformatics 2015;31:1771–9.
[88] Zelezniak A, Andrejev S, Ponomarova O, Mende DR, Bork P, et al. Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc Natl Acad Sci U S A 2015;112:6449–54.
[89] Chacón JM, Harcombe WR. The power of metabolism for predicting microbial community dynamics. mSystems 2019;4:e00146–e219.
[90] Pacheco AR, Moel M, Segrè D. Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. Nat Commun 2019;10:103.
[91] Machado D, Andrejev S, Tramontano M, Patil KR. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. Nucl Acids Res 2018;46:7542–53.
[92] Yizhak K, Benyamini T, Liebermeister W, Ruppin E, Shlomi T. Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. Bioinformatics 2010;26:i255–60.
[93] Hellweger FL, Clegg RJ, Clark JR, Plugge CM, Kreft J-U. Advancing microbial sciences by individual-based modelling. Nat Rev Microbiol 2016;14:461–71.
[94] Succurro A, Moejes FW, Ebenhöh O. A diverse community to study communities: integration of experiments and mathematical models to study microbial consortia. J Bacteriol 2017;199:e00865–e916.
[95] Bauer E, Zimmermann J, Baldini F, Thiele I, Kaleta C. BacArena: individual-based metabolic modeling of heterogeneous microbes in complex communities. PLoS Comput Biol 2017;13:e1005544.
[96] Marsland 3rd R, Cui W, Mehta P. A minimal model for microbial biodiversity can reproduce experimentally observed ecological patterns. Sci Rep 2020;10:3308.
[97] Jiao S, Yang Y, Xu Y, Zhang J, Lu Y. Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. ISME J 2020;14:202–16.
[98] Zhao P, Bao J, Wang X, Liu Y, Li C, et al. Deterministic processes dominate soil microbial community assembly in subalpine coniferous forests on the Loess Plateau. PeerJ 2019;7:e6746.
[99] Logares R, Deutschmann IM, Junger PC, Giner CR, Krabberød AK, et al. Disentangling the mechanisms shaping the surface ocean microbiota. Microbiome 2020;8:55.
[100] Aguirre de Cárcer D. A conceptual framework for the phylogenetically constrained assembly of microbial communities. Microbiome 2019;7:142.
[101] Burns AR, Guillemin K. The scales of the zebrafish: host-microbiota interactions from proteins to populations. Curr Opin Microbiol 2017;38:137–41.
[102] Fu H, Uchimiya M, Gore J, Moran MA. Ecological drivers of bacterial community assembly in synthetic phycospheres. Proc Natl Acad Sci U S A 2020;117:3656–62.
[103] Venturelli OS, Carr AC, Fisher G, Hsu RH, Lau R, et al. Deciphering microbial interactions in synthetic human gut microbiome communities. Mol Syst Biol 2018;14:20178157.
[104] Røder HL, Sørensen SJ, Burmølle M. Studying bacterial multispecies biofilms: where to start?. Trends Microbiol 2016;24:503–13.
[105] Justice NB, Sczesnak A, Hazen TC, Arkin AP. Environmental selection, dispersal, and organism interactions shape community assembly in high-throughput enrichment culturing. Appl Environ Microbiol 2017;83: e01253–e1317.
[106] Amor DR, Ratzke C, Gore J. Transient invaders can induce shifts between alternative stable states of microbial communities. Sci Adv 2020;6:eaay8676.
[107] Kehe J, Kulesa A, Ortiz A, Ackerman CM, Thakku SG, et al. Massively parallel screening of synthetic microbial communities. Proc Nat Acad Sci U S A 2019;116:12804.
[108] Blasche S, Kim Y, Oliveira AP, Patil KR. Model microbial communities for ecosystems biology. Curr Opin Sys Biol 2017;6:51–7.
[109] Wolfe BE. Using cultivated microbial communities to dissect microbiome assembly: challenges, limitations, and the path ahead. mSystems 2018;3: e00161–e217.
[110] Epstein S, Lewis K, Nichols D, E G. New Approaches to Microbial Isolation. In: Baltz R, Demain A, Davies J, Bull A, Junker B, Katz L, et al., editors. Manual of Industrial Microbiology and Biotechnology. Washington, DC: ASM Press; 2010. p. 3-12.

[111] Oberhardt MA, Zarecki R, Gronow S, Lang E, Klenk H-P, et al. Harnessing the landscape of microbial culture media to predict new organism–media pairings. Nat Commun 2015;6:8493.
[112] Hartman K, van der Heijden MGA, Roussely-Provent V, Walser J-C, Schlaeppi K. Deciphering composition and function of the root microbiome of a legume plant. Microbiome 2017;5:2.
[113] Bai Y, Müller DB, Srinivas G, Garrido-Oter R, Potthoff E, et al. Functional overlap of the Arabidopsis leaf and root microbiota. Nature 2015;528:364–9.
[114] Zhang J, Liu Y-X, Zhang N, Hu B, Jin T, et al. NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. Nat Biotechnol 2019;37:676–84.
[115] Lozano GL, Bravo JI, Garavito Diago MF, Park HB, Hurley A, et al. Introducing THOR, a model microbiome for genetic dissection of community behavior. MBio 2019;10:02846–2918.
[116] Cui Z, Huntley RB, Zeng Q, Steven B (2020) Temporal and spatial dynamics in the apple flower microbiome in the presence of the phytopathogen Erwinia amylovora. bioRxiv: 2020.02.19.956078.
[117] Martínez I, Maldonado-Gomez MX, Gomes-Neto JC, Kittana H, Ding H, et al. Experimental evaluation of the importance of colonization history in early-life gut microbiota assembly. Elife 2018;18:36521.
[118] Liu W, Jacquiod S, Brejnrod A, Russel J, Burmolle M, et al. Deciphering links between bacterial interactions and spatial organization in multispecies biofilms. ISME J 2019;13:3054–66.
[119] Leventhal GE, Boix C, Kuechler U, Enke TN, Sliwerska E, et al. Strain-level diversity drives alternative community types in millimetre-scale granular biofilms. Nat Microbiol 2018;3:1295–303.
[120] Oliphant K, Parreira VR, Cochrane K, Allen-Vercoe E. Drivers of human gut microbial community assembly: coadaptation, determinism and stochasticity. ISME J 2019;13:3080–92.
[121] Cairns J, Jokela R, Becks L, Mustonen V, Hiltunen T. Repeatable ecological dynamics govern the response of experimental communities to antibiotic pulse perturbation. Nat Ecol Evol 2020;4:1385–94.
[122] Koskella B, Brockhurst MA. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS Microbiol Rev 2014;38:916–31.
[123] De Sordi L, Lourenço M, Debarbieux L. The battle within: interactions of bacteriophages and bacteria in the gastrointestinal tract. Cell Host Microbe 2019;25:210–8.
[124] Hsu RH, Clark RL, Tan JW, Ahn JC, Gupta S, et al. (2019) Microbial Interaction Network Inference in Microfluidic Droplets. Cell Systems 9: 229-42.e4.
[125] Gupta S, Ross TD, Gomez MM, Grant JL, Romero PA, et al. Investigating the dynamics of microbial consortia in spatially structured environments. Nat Commun 2020;11:2418.
[126] Madsen JS, Sørensen SJ, Burmølle M. Bacterial social interactions and the emergence of community-intrinsic properties. Curr Opin Microbiol 2018;42:104–9.
[127] Hansen RH, Timm AC, Timm CM, Bible AN, Morrell-Falvey JL, et al. Stochastic assembly of bacteria in microwell arrays reveals the importance of confinement in community development. PLoS One 2016;11:e0155080.
[128] Wondraczek L, Pohnert G, Schacher FH, Köhler A, Gottschaldt M, et al. Artificial microbial arenas: materials for observing and manipulating microbial consortia. Adv Mater 2019;31:16.
[129] Dunham SJ, Ellis JF, Li B, Sweedler JV. Mass spectrometry imaging of complex microbial communities. Acc Chem Res 2017;50:96–104.
[130] Wang W, Lin L, Du Y, Song Y, Peng X, et al. Assessing the viability of transplanted gut microbiota by sequential tagging with D-amino acid-based metabolic probes. Nat Commun 2019;10:019–09267.