



# Toward FAIR Representations of Microbial Interactions

Alan R. Pacheco,<sup>a</sup> Charlie Pauvert,<sup>b</sup> Dileep Kishore,<sup>c</sup> Daniel Segrè<sup>c,d</sup>

<sup>a</sup>Institute of Microbiology, ETH Zurich, Zurich, Switzerland

<sup>b</sup>Functional Microbiome Research Group, Institute of Medical Microbiology, University Hospital of RWTH, Aachen, Germany

<sup>c</sup>Bioinformatics Program and Biological Design Center, Boston University, Boston, Massachusetts, USA

<sup>d</sup>Department of Biology, Department of Biomedical Engineering, Department of Physics, Boston University, Boston Massachusetts, USA

Alan R. Pacheco and Charlie Pauvert contributed equally to this work. Author order was determined alphabetically.

**ABSTRACT** Despite an ever-growing number of data sets that catalog and characterize interactions between microbes in different environments and conditions, many of these data are neither easily accessible nor intercompatible. These limitations present a major challenge to microbiome research by hindering the streamlined drawing of inferences across studies. Here, we propose guiding principles to make microbial interaction data more findable, accessible, interoperable, and reusable (FAIR). We outline specific use cases for interaction data that span the diverse space of microbiome research, and discuss the untapped potential for new insights that can be fulfilled through broader integration of microbial interaction data. These include, among others, the design of intercompatible synthetic communities for environmental, industrial, or medical applications, and the inference of novel interactions from disparate studies. Lastly, we envision potential trajectories for the deployment of FAIR microbial interaction data based on existing resources, reporting standards, and current momentum within the community.

**KEYWORDS** microbiome, microbial interactions, microbial ecology, data sharing, accessibility, reproducibility, FAIR, metadata, co-occurrence, microbial networks

The enormous progress in biotechnological and computational techniques over the last few decades has revolutionized our understanding of microbial communities. In particular, studies based on amplicon and metagenomic sequencing have further clarified the fact that microbiomes are not static entities, but dynamic ecosystems whose constituent members interact in myriad ways with each other and with their environments (1–6). These sequence-based approaches have catalyzed early efforts to map microbial interrelationships and represent an invaluable first step in identifying the organisms participating in these interactions. However, these approaches have largely been limited to inferring microbial associations (e.g., via co-occurrences) (7, 8). While these associations may partially reflect causal interactions between microbes, the full landscape of interdependencies is likely much richer and more nuanced in ways that we are just starting to grasp (9–13). Indeed, interactions may be defined and measured in many different ways (14, 15), for example, by evaluating direct contact between cells (16, 17), physical proximity (18, 19), the cost of producing exchanged metabolites (20–22), and the type of chemical mediators involved (23, 24). These factors are crucial in determining the emergence and consequences of an interaction beyond its ecological classification (mutualism, competition, etc.), and can provide a more complete view of microbial ecosystem properties, which is helpful for building mathematical models of community dynamics.

Advances in metabolomics, transcriptomics, and high-throughput culturing platforms are beginning to produce a growing body of data on the mechanisms and environmental dependencies exhibited by microbial interactions (25–28). While this wealth

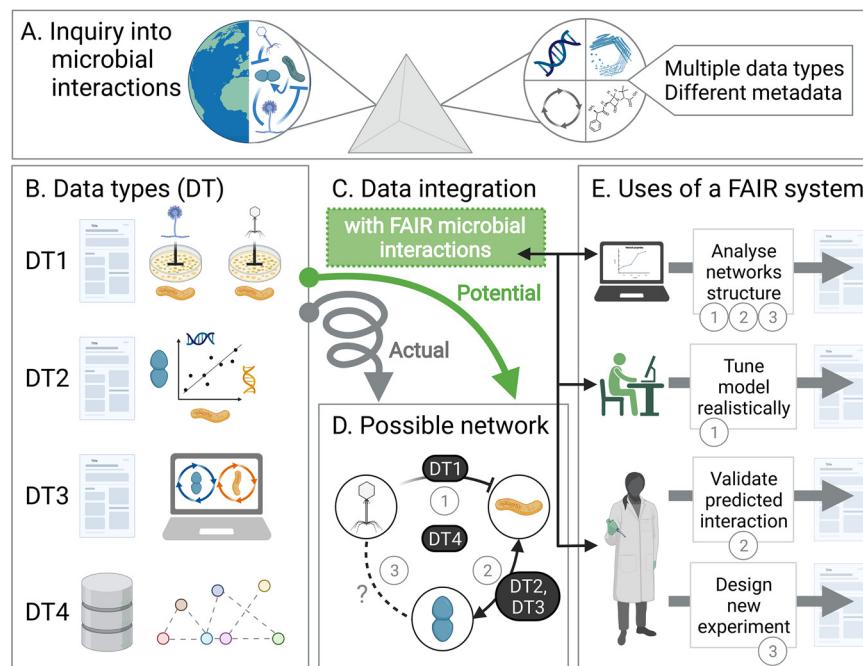
**Editor** Benjamin E. Wolfe, Tufts University

**Copyright** © 2022 Pacheco et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Daniel Segrè, dsegre@bu.edu.

The authors declare no conflict of interest.

**Published** 25 August 2022



**FIG 1** Applying a FAIR system to the study of microbial interactions and correlations. (A–B) Studies that investigate microbial interactions gain relevant insights through the generation of multiple types of data and metadata (DTs). As representative examples of microbe–microbe interactions, we may consider fungi–bacteria or phage–bacteria cultivation experiments (DT1), correlations based on amplicon sequence or operational taxonomic unit counts (DT2), and flux balance models of genome-scale metabolic networks of two or more species (DT3). While specific databases (DT4 and Table 1) that compile these data sources exist, they lack a common reporting standard, which hinders downstream application and integration. (C) We envision a systematic approach for reporting microbial interactions following the principles of Findability, Accessibility, Interoperability, and Reusability (FAIR). (D) A FAIR representation of microbial interaction data, based on common identifiers for microorganisms and specific encodings of interactions and uncertainty, can enable new insights that bridge subdisciplines and generate predictions of new interaction networks. (E) Scientists spanning diverse areas of microbiome science can benefit from FAIR reporting of interactions. For example, a network scientist could identify common structures relying on a broader corpus of interactions, a modeler could more easily identify specific interactions to realistically simulate ecological dynamics, and an experimentalist could assess whether a novel interaction has been reported in other hosts or contexts. All these uses of the framework would lead not only to new scientific insights and more streamlined contribution to data collections, but also to growing interconnectedness within the diverse field of microbiome science.

of information has the potential to enhance our knowledge of specific microbial interrelationships, it poses the new challenge of finding an appropriate framework to describe interactions in a way that efficiently encompasses their diversity and complexity (Fig. 1A) (29–31).

Addressing this challenge would allow us to dynamically and continuously combine diverse sources of data (Fig. 1B and C) to yield insights that could not be obtainable from individual data sets. For example, one could use an open interaction database to easily fetch experimentally-grounded parameters (e.g., metabolite uptake/secretion rates) for simulating dynamics of microbial food webs and to predict the general conditions that determine the stability of a community (Fig. 1D) (32). As another example, network scientists and ecologists could use the same database to assess how widespread particular interactions are (33), or to identify network structures that are common across biomes or taxonomic groups, refining our understanding of how microbial ecosystems assemble (Fig. 1E) (34, 35). These inferences could then be used to clarify the connection between causal interactions and observed co-occurrence patterns—distinct data types whose combination can advance the understanding of how different microbial relationships affect community function, dynamics, and resilience. Finally, simultaneous mining of multiple data sets would enable searches for examples

**TABLE 1** Databases containing microbial interactions and their relevant FAIR features<sup>a</sup>

Database	Website	Reference	Relevant features
Microbial Interaction Network Database (MIND)	<a href="http://microbialnet.org">microbialnet.org</a>	43	Contains microbe–microbe interactions categorized by environmental contexts and metadata. Microorganisms involved are mapped to their NCBI Taxonomy ID to facilitate comparisons.
Protist Interaction DAtabase (PIDA)	<a href="https://github.com/ramalok/PIDA">github.com/ramalok/PIDA</a>	44	Contains literature-extracted microbe–microbe interactions categorized by interaction source and type (ecological). Microorganisms are mapped using Genbank accession nos. when appropriate.
Microbiota-Active Substance Interactions database (MASI)	<a href="http://www.aidlab.com/MASI/">http://www.aidlab.com/MASI/</a>	45	Contains microbe–drug and microbe–disease interactions with particular emphasis on linking to external resources. Microorganisms are mapped using NCBI Taxonomy ID, and molecules are mapped using PubChem ID.
The Bacterial Diversity Metadatabase (BacDive)	<a href="http://bacdive.com">bacdive.com</a>	57	Contains integrated bacterial information on culture conditions and physiology that are accessible interactively and programmatically. Deposited strains are mapped to their NCBI Taxonomy ID and culture collections persistent identifiers.
Web of Microbes (WoM)	<a href="http://webofmicrobes.org">webofmicrobes.org</a>	46	Contains microbe–metabolite–microbe interactions for multiple media conditions. Exptl information linking metabolite (environment) and microbe compatibility is displayed in a browsable and searchable manner.
Global Biotic Interactions database (GloBI)	<a href="http://globalbioticinteractions.org">globalbioticinteractions.org</a>	47	Contains descriptions of relationships between biological entities in general (i.e., species–species, species–habitat, etc.). Entities are indexed based on independent repositories that comply with light formatting and follow a strict ecology-based interaction ontology.

<sup>a</sup>In addition to databases describing intermicrobial interactions, this table also contains databases that contain frameworks to compile a variety of biological relationships (e.g., GloBI), as well as microbial databases aggregating scattered knowledge and promoting data sharing (e.g., MASI and BacDive).

of specific interactions (Fig. 1E), to (i) identify interactions that occur robustly irrespective of biome and experimental details, (ii) facilitate the bottom-up design of synthetic consortia by complementing existing approaches (36–38), or (iii) help create experimentally-verified data sets for benchmarking microbial inference methods (39).

More generally, the prevalence and effects of specific interaction attributes (e.g., dependence on specific resources (40), strain-level physiological differences (41), and definitions of ecological outcomes (8, 42)) across organisms and ecosystems could be quantified and compared, contributing to an enhanced understanding of the general ecological principles that govern living systems.

Despite this promising prospect, several factors complicate the integration of microbial interaction data. Among these, of particular importance is the fact that the majority of interaction data are not accessible outside the original study in which they were reported and often appear in the form of arbitrarily formatted tables. For this reason, efforts have been made to create interaction databases with standardized formatting (Table 1) (43–47). While these represent useful resources for finding specific interacting participants from diverse microbiomes, such databases are often limited to one

or a few types of data. Nonetheless, these efforts follow in the rich history of endeavors within biology aimed at standardizing and sharing data and computational models of biological systems (47–51). For example, one of the most prominent early cases for the need for standardization was the explosion of high-throughput data generated from DNA microarrays at the end of the 1990s. It is particularly telling that a commentary article accompanying the paper that proposed MIAME (Minimum Information About a Microarray Experiment) (50) was titled “Microarray Standards at Last” (52), capturing the acknowledged need for appropriate reporting standards at the time. As suggested by the title, MIAME was not created overnight, but rather entailed a careful process that integrated viewpoints from multiple stakeholders to create a useful and accepted reporting framework that enhanced the reproducibility of results and the drawing of broader insights from integrated sources of data. To begin a similar journey, we propose that a greater focus on reporting interaction attributes and mechanisms using standardized formats will open up important opportunities for the microbiome field, and outline specific steps that can be taken to reach this potential.

## A FAIR REPRESENTATION OF MICROBIAL INTERACTIONS

We specifically envision the adoption of data sharing and stewardship practices that would enable microbial interaction data to fulfill the principles of Findability, Accessibility, Interoperability, and Reusability (FAIR). These principles, first formally presented in 2016 to address growing challenges in scientific data management, serve to guide efforts that aim to improve access to reliable and reproducible scholarly data (53) and have already been adopted as an important component of microbiome data management (1, 6, 54–58). We therefore focus our present discussion on two concrete efforts that can be initiated in order to make data on intermicrobial interactions more FAIR, namely (i) the creation and/or adoption of open web infrastructures for cataloging and making data from disparate sources available to the community (hence Findable and Accessible), and (ii) the adoption of a minimal set of metadata requirements that are human- and machine-readable (i.e., Interoperable and Reusable).

**1: An open web catalog for findable and accessible microbial interaction data.** It is difficult to imagine how to make progress on Findability and Accessibility without a centralized resource that is capable of capturing the wide breadth of interactions currently available only in individual publications or split into type-specific repositories (Fig. 1B and Table 1), and designed to be able to grow to accommodate newly generated data. Such a resource could be generated via multiple strategies, including the following:

- A. **Integrating existing repositories of microbial interaction data into an established infrastructure.** Several existing infrastructures could in principle serve this purpose. To illustrate the potential advantages and challenges of this strategy, we may consider as an example Global Biotic Interactions (GloBI) (47), a metadatabase for sharing and analyzing species interaction data. It provides a searchable platform to identify specific interactions based on the organisms involved and the relationships they experience (e.g., X preys on, hosts, is symbiont of Y), and can therefore serve as a structure for integrating a wider breadth of microbial interaction data and their attributes. Nonetheless, as GloBI considers the species level as the most phylogenetically precise, it may not easily capture strain- or mutant-specific interactions common in microbial ecology research (59). Integration into GloBI would also require amending existing metadata items not applicable to microbes, as well as the ontologies of interactions (apart from “interacts with,” “parasitizes,” “ecologically co-occurs with”) to match features of microbial interactions.
- B. **Using existing database-building tools and available metadata.** For example, one could consider using the recently-published tool mako (35) to create a database by importing network files or deposited sequences to be analyzed. If these input files contain metadata on nodes, edges, or samples from which sequences were obtained, the metadata in question can be readily propagated

into the database. However, mako is currently limited to undirected interactions and to creating a local interaction database, which complicates continuous online access and editing.

C. **Establishing a new database specific to microbial interactions.** Such a database could be designed from the ground up to more flexibly store multiple types of microbial interactions (e.g., co-occurrences, causal interactions, and higher-order interactions), as well as their attributes. This approach could apply several modalities for importing and organizing data in a standardized or automated way, facilitating the incorporation of data from individual studies as they are published. Recent manually-compiled catalogs of interactions and their attributes (14, 15) as well as of tools to convert them into searchable online resources ([cpauvert.github.io/mi-atlas](https://cpauvert.github.io/mi-atlas)) may serve as small-scale examples for planning such a larger-scale resource.

**2: A minimal set of metadata requirements for interoperability and reusability.**

While a centralized database would lay a foundation for FAIR microbial interaction data, its impact would remain limited if its contents cannot be easily updated by scientists and accessed by humans and machines. We therefore also advocate for the inclusion of metadata with reports of interactions as a way to promote interoperability and reusability. While convergence to specific guidelines will require significant community discussions and buy-in from stakeholders, we propose that the following four categories of metadata could serve as a starting point: microbial entities, interaction inference methods, interaction context, and attributes. These are described in detail in Table 2 and are outlined as follows:

A. **Microbial entities.** The species (and strain, if relevant) names of each of the microbes participating in an interaction should be provided, e.g., in a comma-separated list, along with their taxonomic accession numbers and eventually their sequence identifier (Table 2A). These lists would also accommodate interactions that cannot be easily described via a pairwise representation. Interaction attributes and effects specific to each participant could be matched with each identifier.

B. **Interaction inference methods.** Despite being challenging to standardize, documentation of the methods that were used to identify an interaction represent highly relevant metadata. As a first step, the evidence for the interaction in question can be broadly categorized using the Evidence and Conclusion Ontology (60), which would indicate whether experimental or computational methods (or both) were used. We also propose a more specific metadata item for the type of computational or experimental method used (e.g., simulation, microscopy, cultivation, and sampling) (Table 2B). Lastly, the relevant publication, code, detailed protocols, and other literature-based evidence should be accessible via persistent identifiers (e.g., DOIs).

C. **Interaction context.** The environmental context of the interaction—such as the biome (e.g., host-associated, synthetic)—could be documented using the Environment Ontology (61) or propagated from the samples used to infer the interaction. Cultivation conditions could also be integrated following the standards established by databases of bacterial isolates (57) and extended to co-cultures. Relevant metadata are proposed in Table 2C with an emphasis on linking values to existing resources such as the Gene Ontology for cellular components, the Chemical Entities of Biological Interest Ontology for compounds, and the Genome Standards Consortium (51) for the oxygen status of the environment.

D. **Interaction attributes.** Defining an interaction's type (e.g., cooperation, antagonism, association, pairwise or higher-order, etc.) is also not trivial, but could be guided by incorporating existing ontologies such as the active list maintained by the OBO Foundry (62). Several other frameworks exist to describe interaction types such as GloBI, Population and Community Ontology (63), and Interaction Network Ontology (64). It nonetheless remains to be seen whether these ontologies are appropriate for

**TABLE 2** List of proposed metadata for minimum information about intermicrobial interaction data<sup>a</sup>

Metadata	Level	Description
A. Which microbial entities are involved?		
participants	M	Comma-separated list of the microbial entities' names, with descriptions of any genetic manipulations performed.
tax_id	M	Comma-separated list of the matching identifiers from the NCBI Taxonomy at the relevant taxonomic level. (e.g., NCBI:txid1043002, NCBI:txid411903). Novel taxa lacking identifiers are denoted by N/A <sup>b</sup> .
sequence_id	R	Comma-separated list of the accessions to the matching sequence data (e.g., genome, marker gene sequence). Taxa from presequencing era articles could be denoted by N/A.
env_origin	X	Term from the Environmental Ontology indicating from which biome the microbial entities originate (e.g., soil [ENVO:00001998]).
source_collection	X	Comma-separated list of the source of the participants engaging in this interaction: isolation, commercial collection, academic collection
B. How was the interaction uncovered?		
evidence_type	M	Type of evidence used to determine the interaction using the Evidence and Conclusion Ontology. At least one of the following broader terms are required: exptl [ECO:0000006], computational [ECO:0007672], or both [ECO:0007661].
method_type	R	One or several of the following types of methods used to determine the interaction: <ul style="list-style-type: none"> <li>• Simulation-based (e.g., generalized Lotka-Volterra model, genome-scale metabolic model)</li> <li>• Microscopy-based (e.g., co-localization with fluorescent markers, assisted motility)</li> <li>• Cultivation-based (e.g., continuous co-culture in bioreactor, co-plating on solid media)</li> <li>• Sample-based (e.g., co-occurrences drawn from analyses of abundances obtained from <i>in situ</i> or <i>in vivo</i> sampling).</li> </ul>
reference	M	Persistent identifier (e.g., DOI or URL) to a resource, script, or article, documenting the method.
software_parameters	X	Name, version, and parameters of the software used, using the following syntax: {software}:{version}:{parameters}.
C. What is the environmental context of the experiment?		
env_broad_scale	R	Biome term from the Environmental Ontology. Engineered ecosystems such as bioreactors, agar plates, or computational models use N/A.
Site	X	Cellular component (from the Gene Ontology) involved in the interaction: cytoplasm [GO:0005737], membrane [GO:0016020], or the extracellular region [GO:0005576].
compounds	X	One or several chemical entities involved in the interaction using either broad or precise terms from the ChEBI ontology with their identifiers. Example: short-chain fatty acid [CHEBI:26666], bacterial metabolite [CHEBI:76969], or penicillin [CHEBI:17334].
medium_name	X	Name of the ( <i>in vitro</i> or <i>in silico</i> ) cultivation medium used, with URL to the composition. Example: BHI [ <a href="https://bacmedia.dsmz.de/medium/215">https://bacmedia.dsmz.de/medium/215</a> ].
rel_to_oxygen	X	Term indicating the oxygen status of the environment using terms from the MiS: aerobic or anaerobic.
ph	X	Measurement of the pH in the environment.
temp	X	Measurement of the temp of the environment (in °C).
carbon_source	X	Term from the ChEBI indicating the specific carbon source(s) used.
nitrogen_source	X	Term from the ChEBI indicating the specific nitrogen source(s) used.
inoculation_densities	X	Comma-separated list of densities for cultivation experiments measured with the associated units in brackets: with optical density [OD 600], with colony forming units [CFU/mL].
D. What are the attributes of the interaction?		
participant_outcomes	R	Comma-separated list of the outcome for each participant: 0 (not affected), 1 (positively affected), -1 (negatively affected), N/A.
ecological_outcome	X	For pairwise interactions, one or several terms describing the overall outcome ( <a href="https://doi.org/10.2307/1307540">https://doi.org/10.2307/1307540</a> ): co-occurrence, cooperation, commensalism, exploitation, amensalism, competition, neutralism.
strength	X	Numerical value quantifying the intensity of the interaction. Example: the inhibition score after co-plating, the correlation value between relative abundances, or estimate of generalized Lotka-Volterra model parameters.
dependencies	X	One or several of the following terms: contact, time, space.
keywords	X	Comma-separated list of terms providing more detail on the broader context of the interaction (e.g., disease-related, biofuel production, uncultivable organisms, metabolic engineering).
notes	X	Open text field for additional relevant comments.

<sup>a</sup>The level of requirement of the metadata is either mandatory (M), recommended (R), or optional (X). Four general questions regarding interactions divide the metadata into four categories (A–D).

<sup>b</sup>N/A, not applicable.

describing all known attributes of microbial interactions, or if a larger set defined by the community is needed. In the meantime, we propose the inclusion of the ecological effect experienced by each participant or by each set of participants (positive/negative/neutral) (42), as well as of information on whether the association described is a co-occurrence by providing the associated metric strength. Lastly, we propose the inclusion of any known interaction dependencies (e.g., on spatial structure or physical contact) and any additional user-defined keywords that provide further relevant information not captured in the previous items (Table 2D).

As an example of how these metadata can be compiled for different data types (Fig. 1B), we have used them to describe three interactions gathered from the literature (Table 3).

## OUTLOOK

The practices, standards, and use cases we have outlined here are by no means exhaustive, but are rather meant to catalyze further discussion on ways to improve the access to and usability of data on microbial interactions and their attributes. We believe the time is opportune for such discussions to take place, not only due to the rapidly growing body of data on microbial interactions and their mechanisms, but also because of a growing momentum within the microbiome community to improve the reliability and reproducibility of research outputs. These are exemplified by government-funded initiatives such as the National Microbiome Data Collaborative (NMDC, USA (56, 65) and the National Research Data Infrastructure (NFDI4Microbiota, Germany; (<https://nfdi4microbiota.de>), which advocate for the adoption of reporting standards for microbiome data. As with existing accepted data reporting standards, any proposed global framework for describing microbial interactions must be shaped by its various stakeholders, including computational and empirical researchers, industry representatives, funding agencies, educational users, and publishers. Such involvement would enable any formalism to be flexible and broadly embraced, as opposed to a rigid standard with little endorsement or room for growth.

Bearing these considerations in mind, we suggest the following roadmap toward FAIR microbial interaction data. First, we call for increased discussions within the scientific community to select and prioritize the interaction features that are most useful to report. These can be carried out via dedicated workshops that, in addition to biologists, could include philosophers of biology interested in microorganisms, as well as physicists and mathematicians who can help define the qualitative and quantitative nature of intermicrobial interactions and their important attributes. This first community-driven effort could lead to the creation of a reporting standard, extending our suggestions in Table 2 to a more mature “Minimal Information for Intermicrobial Interactions” definition similar to those for publishing microarray data (50) and genome sequences (51, 66, 67), or for assessing the quality of genome-scale metabolic models (68). Second, these metadata suggestions could be further implemented as usable formats such as SBML (49, 69) or BIOM (70), which enable the standardized export and sharing of genome-scale models and count data, respectively. As such, data scientists and bioinformaticians could take part in hackathons to develop such a toolbox with standardized file formats, converter scripts, and validators to streamline the adoption of microbial interaction metadata. Third, we envision teams of investigators and students gathering for “annota-thons” to collaboratively extract knowledge from the microbial interaction literature and use the aforementioned toolboxes to compile the relevant metadata, ensuring that an open web catalog of microbial interactions truly relies on known published material. Last, the rise of an open community, willing to quickly share protocols and methods of scientific projects enabled by FAIR microbial interaction data resources, would provide further incentives for adoption of standard formats, creating a positive feedback loop that could accelerate benefits for the whole community and pave the way for major integrative and collaborative advances in microbiome research.

**TABLE 3** Three microbial interaction data types represented according to the proposed metadata guidelines in Table 2<sup>a</sup>

Metadata	Example A	Example B	Example C
Microbial entities			
participants	<i>Azotobacter vinelandii</i> DSM 85, <i>Chlamydomonas reinhardtii</i> strain 187, <i>Alternaria</i> sp. GY-05-1221 NCBItaxonid <sup>b</sup> 354, NCBItaxonid3055, NCBItaxonid667197 N/A, N/A, FJ627005.1	<i>Acidobacteria</i> , <i>Gamma proteobacteria</i> NCBItaxonid57723, NCBItaxonid1236	<i>Bacteroides caccae</i> ATCC 43185, <i>Lactobacillus rhamnosus</i> GG NCBItaxonid411901, NCBItaxonid568703 NZ_CP022412.2, NZ_CP031290.1 digestive tract environment [ENVO:01001033], digestive tract environment [ENVO:01001033]
sequence_id			
env_origin			
source_collection	commercial collection, commercial collection, isolation		
Interaction inference methods			
evidence_type	experimental evidence used in manual assertion [ECO:0000269]	high throughput evidence used in automatic assertion [ECO:00006057]	computational evidence [ECO:0007672]
method_type	cultivation-based microscopy-based	sample-based	simulation-based
reference	<a href="https://doi.org/10.1007/s12223-010-0067-9">https://doi.org/10.1007/s12223-010-0067-9</a>		<a href="https://doi.org/10.1038/ismej.2011.119">https://doi.org/10.1038/ismej.2011.119</a>
Interaction context			
env_broad_scale		soil environment [ENVO:01001044]	
compounds	chlorophyll [CHEBI:28966], cystathioneine [CHEBI:17755]	alanine [CHEBI:16449], asparagine [CHEBI:22653], nicotinic acid [CHEBI:15940], lactate [CHEBI:24996]	
medium_name	Azotobacter Medium [ <a href="https://bacmedia.dsmz.de/medium/3">https://bacmedia.dsmz.de/medium/3</a> ]	DMEM 6429 (+ vitamin K, hemin, and arabinogalactan)	
rel_to_oxygen	aerobic	anaerobic	
temp	25	37	
Interaction attributes			
participants_outcome	1,1,1	N/A <sup>b</sup>	
ecological_outcome	cooperation	co-occurrence	
strength	0.66	0.66	
dependencies	contact		

<sup>a</sup>Included here are an interaction from an alga-bacteria-fungi cultivation experiment (A), a co-occurrence analysis of bacterial operational taxonomic units (B), and a genome-scale metabolic model (C).

<sup>b</sup>N/A, not applicable.

## ACKNOWLEDGMENTS

We thank the editors and the reviewers for their positive comments and for their helpful recommendations. A.R.P. is funded by a James S. McDonnell Postdoctoral Fellowship. C.P. is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—460129525. D.K. was partially funded by the Kilachand Multicellular Design Program graduate fellowship. D.S. acknowledges funding from the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research through the Microbial Community Analysis and Functional Evaluation in Soils Science Focus Area Program (m-CAFEs) under contract number DE-AC02-05CH11231 to Lawrence Berkeley National Laboratory, from the Human Frontier Science Program (HFSP Research Grant RGP0060/2021), from the NSF Center for Chemical Currencies of a Microbial Planet (C-CoMP, publication #011), from the NIH National Cancer Institute (grant R21CA260382), and from the NIH National Institute on Aging, award number UH2AG064704. Figure elements created with Biorender.com.

## REFERENCES

- Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, Navas-Molina JA, Janssen S, Kopylova E, Vázquez-Baeza Y, González A, Morton JT, Mirarab S, Zech Xu Z, Jiang L, Haroon MF, Kanbar J, Zhu Q, Jin Song S, Koscolek T, Bokulich NA, Lefler J, Brislawn CJ, Humphrey G, Owens SM, Hampton-Marcell J, Berg-Lyons D, McKenzie V, Fierer N, Fuhrman JA, Claeset A, Stevens RL, Shade A, Pollard KS, Goodwin KD, Jansson JK, Gilbert JA, Knight R, Earth Microbiome Project Consortium. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551:457–463. <https://doi.org/10.1038/nature24621>.
- The Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486:207–214. <https://doi.org/10.1038/nature11234>.
- Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, Chaffron S, Ignacio-Espinoza JC, Roux S, Vincent F, Bittner L, Darzi Y, Wang J, Audic S, Berline L, Bontempi G, Cabello AM, Coppola L, Cornejo-Castillo FM, d'Ovidio F, De Meester L, Ferrera I, Garet-Delmas M-J, Guidi L, Lara E, Pesant S, Royo-Llonch M, Salazar G, Sánchez P, Sebastian M, Souffreau C, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Gorsky G, Not F, Ogata H, Speich S, Stemann L, Weissenbach J, Wincker P, Acinas SG, Sunagawa S, Bork P, Sullivan MB, Karsenti E, Bowler C, de Vargas C, Raes J. 2015. Determinants of community structure in the global plankton interactome. *Science* 348:1262073. <https://doi.org/10.1126/science.1262073>.
- Royo-Llonch M, Sánchez P, Ruiz-González C, Salazar G, Pedrós-Alio C, Sebastián M, Labadie K, Paoli L, Ibarbalz F, Zinger L, Churchward B, Chaffron S, Eveillard D, Karsenti E, Sunagawa S, Wincker P, Karp-Boss L, Bowler C, Acinas SG. 2021. Compendium of 530 metagenome-assembled bacterial and archaeal genomes from the polar Arctic Ocean. *Nat Microbiol* 6:1561–1574. <https://doi.org/10.1038/s41564-021-00979-9>.
- Tedersoo L, Bahram M, Pöhlme S, Köljalg U, Yorou NS, Wijesundera R, Villarreal Ruiz L, Vasco-Palacios AM, Thu PQ, Suija A, Smith ME, Sharp C, Saluveer E, Saitta A, Rosas M, Riit T, Ratkowsky D, Pritsch K, Pöldmaa K, Piepenbring M, Phosri C, Peterson M, Parts K, Pärtel K, Otsing E, Nouhra E, Njouonkou AL, Nilsson RH, Morgado LN, Mayor J, May TW, Majuakim L, Lodge DJ, Lee SS, Larsson K-H, Kohout P, Hosaka K, Hiiiesalu I, Henkel TW, Harend H, Guo L-D, Greslebin A, Grelet G, Geml J, Gates G, Dunstan W, Dunk C, Drenkhan R, Dearnaley J, De Kesel A, et al. 2014. Global diversity and geography of soil fungi. *Science* 346:1256688. <https://doi.org/10.1126/science.1256688>.
- Gilbert JA, Meyer F, Antonopoulos D, Balaji P, Brown CT, Brown CT, Desai N, Eisen JA, Evers D, Field D, Feng W, Huson D, Jansson J, Knight R, Knight J, Kolker E, Konstantindis K, Kostka J, Kyprides N, Mackelprang R, McHardy A, Quince C, Raes J, Sczyrba A, Shade A, Stevens R. 2010. Meeting report: the terabase metagenomics workshop and the vision of an Earth microbiome project. *Stand Genomic Sci* 3:243–248. <https://doi.org/10.4056/sigs.1433550>.
- Ma B, Wang Y, Ye S, Liu S, Stirling E, Gilbert JA, Faust K, Knight R, Jansson JK, Cardona C, Röttjers L, Xu J. 2020. Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome* 8: 82. <https://doi.org/10.1186/s40168-020-00857-2>.
- Faust K, Raes J. 2012. Microbial interactions: from networks to models. *Nat Rev Microbiol* 10:538–550. <https://doi.org/10.1038/nrmicro2832>.
- Blanchet FG, Cazelles K, Gravel D. 2020. Co-occurrence is not evidence of ecological interactions. *Ecol Lett* 23:1050–1063. <https://doi.org/10.1111/ele.13525>.
- Gupta G, Ndiaye A, Filteau M. 2021. Leveraging experimental strategies to capture different dimensions of microbial interactions. *Front Microbiol* 12:700752. <https://doi.org/10.3389/fmicb.2021.700752>.
- Carr A, Diener C, Baliga NS, Gibbons SM. 2019. Use and abuse of correlation analyses in microbial ecology. *ISME J* 13:2647–2655. <https://doi.org/10.1038/s41396-019-0459-z>.
- Hirano H, Takemoto K. 2019. Difficulty in inferring microbial community structure based on co-occurrence network approaches. *BMC Bioinformatics* 20:329. <https://doi.org/10.1186/s12859-019-2915-1>.
- Berry D, Widder S. 2014. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front Microbiol* 5:219.
- D'Souza G, Shitut S, Preussger D, Yousif G, Waschna S, Kost C. 2018. Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Nat Prod Rep* 35:455–488. <https://doi.org/10.1039/c8np00009c>.
- Pacheco AR, Segré D. 2019. A multidimensional perspective on microbial interactions. *FEMS Microbiol Lett* 366:fnz125. <https://doi.org/10.1093/femsle/fnz125>.
- Konovalova A, Søgaard-Andersen L. 2011. Close encounters: contact-dependent interactions in bacteria. *Mol Microbiol* 81:297–301. <https://doi.org/10.1111/j.1365-2958.2011.07711.x>.
- Schroekh V, Scherlach K, Nützmann H-W, Shelest E, Schmidt-Heck W, Schuemann J, Martin K, Hertweck C, Brakhage AA. 2009. Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. *Proc Natl Acad Sci U S A* 106:14558–14563. <https://doi.org/10.1073/pnas.0901870106>.
- Gupta S, Ross TD, Gomez MM, Grant JL, Romero PA, Venturelli OS. 2020. Investigating the dynamics of microbial consortia in spatially structured environments. *Nat Commun* 11:2418. <https://doi.org/10.1038/s41467-020-16200-0>.
- Wang G, Or D. 2014. Trophic interactions induce spatial self-organization of microbial consortia on rough surfaces. *Sci Rep* 4:6757. <https://doi.org/10.1038/srep06757>.
- Figueiredo ART, Kramer J. 2020. Cooperation and conflict within the Microbiota and their effects on animal hosts. *Front Ecol Evol* 8:132. <https://doi.org/10.3389/fevo.2020.00132>.
- Zampieri M, Sauer U. 2016. Model-based media selection to minimize the cost of metabolic cooperation in microbial ecosystems. *Bioinformatics* 32: 1733–1739. <https://doi.org/10.1093/bioinformatics/btw062>.
- Pacheco AR, Moel M, Segré D. 2019. Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nat Commun* 10: 103. <https://doi.org/10.1038/s41467-018-07946-9>.
- Niehaus L, Boland I, Liu M, Chen K, Fu D, Henckel C, Chaung K, Miranda SE, Dyckman S, Crum M, Dedrick S, Shou W, Momeni B. 2019. Microbial coexistence through chemical-mediated interactions. *Nat Commun* 10: 2052. <https://doi.org/10.1038/s41467-019-10062-x>.

24. Momeni B, Xie L, Shou W. 2017. Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions. *Elife* 6:e25051. <https://doi.org/10.7554/eLife.25051>.

25. Antwi RE, Griffiths SM, Harrison XA, Aranega-Bou P, Arce A, Bettridge AS, Brailsford FL, de Menezes A, Devaynes A, Forbes KM, Fry EL, Goodhead I, Haskell E, Heys C, James C, Johnston SR, Lewis GR, Lewis Z, Macey MC, McCarthy A, McDonald JE, Mejia-Florez NL, O'Brien D, Orland C, Pautasso M, Reid WDK, Robinson HA, Wilson K, Sutherland WJ. 2017. Fifty important research questions in microbial ecology. *FEMS Microbiol Ecol* 93:fix044.

26. Hsu RH, Clark RL, Tan JW, Ahn JC, Gupta S, Romero PA, Venturelli OS. 2019. Microbial Interaction Network inference in microfluidic droplets. *Cell Syst* 9:229–242.e4. <https://doi.org/10.1016/j.cels.2019.06.008>.

27. Blasche S, Kim Y, Mars RAT, Machado D, Maansson M, Kafkia E, Milanese A, Zeller G, Teusink B, Nielsen J, Benes V, Neves R, Sauer U, Patil KR. 2021. Metabolic cooperation and spatiotemporal niche partitioning in a kefir microbial community. *Nat Microbiol* 6:196–208. <https://doi.org/10.1038/s41564-020-00816-5>.

28. Zhang Y, Kastman EK, Guasto JS, Wolfe BE. 2018. Fungal networks shape dynamics of bacterial dispersal and community assembly in cheese rind microbiomes. *Nat Commun* 9:336. <https://doi.org/10.1038/s41467-017-02522-z>.

29. Cordero OX, Datta MS. 2016. Microbial interactions and community assembly at microscales. *Curr Opin Microbiol* 31:227–234. <https://doi.org/10.1016/j.mib.2016.03.015>.

30. Armitage DW, Jones SE. 2019. How sample heterogeneity can obscure the signal of microbial interactions. *ISME J* 13:2639–2646. <https://doi.org/10.1038/s41396-019-0463-3>.

31. Berg G, Rybakova D, Fischer D, Cernava T, Vergès M-CC, Charles T, Chen X, Cocolin L, Eversole K, Corral GH, Kazou M, Kinkel L, Lange L, Lima N, Loy A, Macklin JA, Maguin E, Mauchline T, McClure R, Mitter B, Ryan M, Sarand I, Smidt H, Schelkde B, Roume H, Kiran GS, Selvin J, de Souza RSC, van Overbeek L, Singh BK, Wagner M, Walsh A, Sessitsch A, Schloter M. 2020. Microbiome definition re-visited: old concepts and new challenges. *Microbiome* 8:103. <https://doi.org/10.1186/s40168-020-00875-0>.

32. Coyte KZ, Schluter J, Foster KR. 2015. The ecology of the microbiome: networks, competition, and stability. *Science* 350:663–666. <https://doi.org/10.1126/science.aad2602>.

33. Palmer JD, Foster KR. 2022. Bacterial species rarely work together. *Science* 376:581–582. <https://doi.org/10.1126/science.abn5093>.

34. Jackson MA, Bonder MJ, Kuncheva Z, Zierer J, Fu J, Kurilshikov A, Wijmenga C, Zhernakova A, Bell JT, Spector TD, Steves CJ. 2018. Detection of stable community structures within gut microbiota co-occurrence networks from different human populations. *PeerJ* 6:e4303. <https://doi.org/10.7717/peerj.4303>.

35. Röttjers L, Faust K. 2022. Fast and flexible analysis of linked microbiome data with mako. *Nat Methods* 19:51–54. <https://doi.org/10.1038/s41592-021-01335-9>.

36. Vannier N, Agler M, Hacquard S. 2019. Microbiota-mediated disease resistance in plants. *PLoS Pathog* 15:e1007740. <https://doi.org/10.1371/journal.ppat.1007740>.

37. Lawson CE, Harcombe WR, Hatzenpichler R, Lindemann SR, Löffler FE, O'Malley MA, García Martín H, Pfleger BF, Raskin L, Venturelli OS, Weissbrodt DG, Noguera DR, McMahon KD. 2019. Common principles and best practices for engineering microbiomes. *Nat Rev Microbiol* 17: 725–741. <https://doi.org/10.1038/s41579-019-0255-9>.

38. Kumar N, Hitch TCA, Haller D, Lagkouvardos I, Clavel T. 2021. MiMiC: a bioinformatic approach for generation of synthetic communities from metagenomes. *Microb Biotechnol* 14:1757–1770. <https://doi.org/10.1111/1751-7915.13845>.

39. Faust K. 2021. Open challenges for microbial network construction and analysis. *ISME J* 15:3111–3118. <https://doi.org/10.1038/s41396-021-01027-4>.

40. Sokolovskaya OM, Shelton AN, Taga ME. 2020. Sharing vitamins: cobamides unveil microbial interactions. *Science* 369:eaba0165. <https://doi.org/10.1126/science.aba0165>.

41. Niccum BA, Kastman EK, Kfouri N, Robbat A, Jr, Wolfe BE. 2020. Strain-level diversity impacts cheese rind microbiome assembly and function. *mSystems* 5:e00149-20. <https://doi.org/10.1128/mSystems.00149-20>.

42. Lidicker WZ, Jr. 1979. A clarification of interactions in ecological systems. *Bioscience* 29:475–477. <https://doi.org/10.2307/1307540>.

43. Hu Z, Kishore D, Wang Y, Birzu G, DeLisi C, Korolev KS, Segre D. 2022. A resource for the comparison and integration of heterogeneous microbiome networks. *bioRxiv*. <https://doi.org/10.1101/2022.08.07.503059>.

44. Bjørbaekmo MFM, Evenstad A, Rosæg LL, Krabberød AK, Logares R. 2020. The planktonic protist interactome: where do we stand after a century of research? *ISME J* 14:544–559. <https://doi.org/10.1038/s41396-019-0542-5>.

45. Zeng X, Yang X, Fan J, Tan Y, Ju L, Shen W, Wang Y, Wang X, Chen W, Ju D, Chen YZ. 2021. MASi: microbiota-active substance interactions database. *Nucleic Acids Res* 49:D776–D782. <https://doi.org/10.1093/nar/gkaa924>.

46. Kosina SM, Greiner AM, Lau RK, Jenkins S, Baran R, Bowen BP, Northen TR. 2018. Web of Microbes (WoM): a curated microbial exometabolomics database for linking chemistry and microbes. *BMC Microbiol* 18:115. <https://doi.org/10.1186/s12866-018-1256-y>.

47. Poelen JH, Simons JD, Mungall CJ. 2014. Global biotic interactions: an open infrastructure to share and analyze species-interaction datasets. *Ecol Inform* 24:148–159. <https://doi.org/10.1016/j.ecoinf.2014.08.005>.

48. McLaughlin JA, Beal J, Misirlı G, Grünberg R, Bartley BA, Scott-Brown J, Vaidyanathan P, Fontanarrosa P, Oberortner E, Wipat A, Gorochowski TE, Myers CJ. 2020. The Synthetic Biology Open Language (SBOL) version 3: simplified data exchange for bioengineering. *Front Bioeng Biotechnol* 8: 1009. <https://doi.org/10.3389/fbioe.2020.01009>.

49. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr J-H, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novère N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J, SBML Forum. 2003. The Systems Biology Markup Language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19: 524–531. <https://doi.org/10.1093/bioinformatics/btg015>.

50. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matei JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M. 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 29:365–371. <https://doi.org/10.1038/ng1201-365>.

51. Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mirzachi I, Johnston A, Cochrane G, Vaughan R, Hunter C, Park J, Morrison N, Rocca-Serra P, Sterk P, Arumugam M, Bailey M, Baumgartner L, Birren BW, Blaser MJ, Bonazzi V, Booth T, Bork P, Bushman FD, Buttigieg PL, Chain PSG, Charlson E, Costello EK, Huot-Creasy H, Dawyndt P, DeSantis T, Fierer N, Fuhrman JA, Gallery RE, Gevers D, Gibbs RA, San Gil I, Gonzalez A, Gordon JI, Guralnick R, Hankeln W, Highlander S, Hugenholtz P, Jansson J, Kau AL, Kelley ST, Kennedy J, Knights D, Koren O, et al. 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MixS) specifications. *Nat Biotechnol* 29:415–420. <https://doi.org/10.1038/nbt.1823>.

52. Nature. 2002. Microarray standards at last. *Nature* 419:323.

53. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3:160018. <https://doi.org/10.1038/sdata.2016.18>.

54. Brunk E, Sahoo S, Zielinski DC, Altunkaya A, Dräger A, Mih N, Gatto F, Nilsson A, Preciat Gonzalez GA, Aurich MK, Prlić A, Sastry A, Danielsdottir AD, Heinken A, Noronha A, Rose PW, Burley SK, Fleming RMT, Nielsen J, Thiele I, Palsson BO. 2018. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat Biotechnol* 36:272–281. <https://doi.org/10.1038/nbt.4072>.

55. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, et al. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* 36:566–569. <https://doi.org/10.1038/nbt.4163>.

56. Wood-Charlson EM, Auberry D, Blanco H, Borkum MI, Corilo YE, Davenport KW, Deshpande S, Devarakonda R, Drake M, Duncan WD, Flynn MC, Hays

D, Hu B, Huntemann M, Li P-E, Lipton M, Lo C-C, Millard D, Miller K, Piehowski PD, Purvine S, Reddy TBK, Shakya M, Sundaramurthi JC, Vangay P, Wei Y, Wilson BE, Canon S, Chain PSG, Fagnan K, Martin S, McCue LA, Mungall CJ, Mouncey NJ, Maxon ME, Eloe-Fadrosh EA. 2020. The National Microbiome Data Collaborative: enabling microbiome science. *Nat Rev Microbiol* 18:313–314. <https://doi.org/10.1038/s41579-020-0377-0>.

57. Reimer LC, Sardà Carbasse J, Koblitz J, Ebeling C, Podstawk A, Overmann J. 2022. BacDive in 2022: the knowledge base for standardized bacterial and archaeal data. *Nucleic Acids Res* 50:D741–D746. <https://doi.org/10.1093/nar/gkab961>.

58. Mirzayi C, Renson A, Zohra F, Elsafoury S, Geistlinger L, Kasselman LJ, Eckenrode K, van de Wijgert J, Loughman A, Marques FZ, MacIntyre DA, Arumugam M, Azhar R, Beghini F, Bergstrom K, Bhatt A, Bisanz JE, Braun J, Bravo HC, Buck GA, Bushman F, Casero D, Clarke G, Collado MC, Cotter PD, Cryan JF, Demmer RT, Devkota S, Elinay E, Escobar JS, Fettweis J, Finn RD, Fodor AA, Forslund S, Franke A, Furlanello C, Gilbert J, Grice E, Haibe-Kains B, Handley S, Herd P, Holmes S, Jacobs JP, Karstens L, Knight R, Knights D, Koren O, Kwon DS, Langille M, Lindsay B, et al. 2021. Reporting guidelines for human microbiome research: the STORMS checklist. *Nat Med* 27:1885–1892. <https://doi.org/10.1038/s41591-021-01552-x>.

59. Van Rossum T, Ferretti P, Maistrenko OM, Bork P. 2020. Diversity within species: interpreting strains in microbiomes. *Nat Rev Microbiol* 18: 491–506. <https://doi.org/10.1038/s41579-020-0368-1>.

60. Nadendla S, Jackson R, Munro J, Quaglia F, Mészáros B, Olley D, Hobbs ET, Goralski SM, Chibucus M, Mungall CJ, Tosatto SCE, Erill I, Giglio MG. 2022. ECO: the Evidence and Conclusion Ontology, an update for 2022. *Nucleic Acids Res* 50:D1515–D1521. <https://doi.org/10.1093/nar/gkab1025>.

61. Buttigieg PL, Morrison N, Smith B, Mungall CJ, Lewis SE, ENVO Consortium. 2013. The environment ontology: contextualising biological and biomedical entities. *J Biomed Semantics* 4:43. <https://doi.org/10.1186/2041-1480-4-43>.

62. Jackson R, Matentzoglu N, Overton JA, Vita R, Balhoff JP, Buttigieg PL, Carbon S, Courtot M, Diehl AD, Dooley DM, Duncan WD, Harris NL, Haendel MA, Lewis SE, Natale DA, Osumi-Sutherland D, Ruttenberg A, Schriml LM, Smith B, Stoeckert CJ, Jr, Vasilevsky NA, Walls RL, Zheng J, Mungall CJ, Peters B. 2021. OBO Foundry in 2021: operationalizing open data principles to evaluate ontologies. *Database (Oxford)* 2021:baab069. <https://doi.org/10.1093/database/baab069>.

63. Walls RL, Deck J, Guralnick R, Baskauf S, Beaman R, Blum S, Bowers S, Buttigieg PL, Davies N, Endresen D, Gandalfo MA, Hanner R, Janning A, Krishtalka L, Matsunaga A, Midford P, Morrison N, Ó Tuama É, Schildhauer M, Smith B, Stucky BJ, Thomer A, Wieczorek J, Whitacre J, Wooley J. 2014. Semantics in support of biodiversity knowledge discovery: an introduction to the Biological Collections Ontology and related ontologies. *PLoS One* 9:e89606. <https://doi.org/10.1371/journal.pone.0089606>.

64. Hur J, Özgür A, Xiang Z, He Y. 2015. Development and application of an interaction network ontology for literature mining of vaccine-associated gene-gene interactions. *J Biomed Semantics* 6:2. <https://doi.org/10.1186/2041-1480-6-2>.

65. Vangay P, Burgin J, Johnston A, Beck KL, Berrios DC, Blumberg K, Canon S, Chain P, Chandonia J-M, Christianson D, Costes SV, Damerow J, Duncan WD, Dundore-Arias JP, Fagnan K, Galazka JM, Gibbons SM, Hays D, Hervey J, Hu B, Hurwitz BL, Jaiswal P, Joachimiak MP, Kinkel L, Ladau J, Martin SL, McCue LA, Miller K, Mouncey N, Mungall C, Pafilis E, Reddy TBK, Richardson L, Roux S, Schriml LM, Shaffer JP, Sundaramurthi JC, Thompson LR, Timmer RE, Zheng J, Wood-Charlson EM, Eloe-Fadrosh EA. 2021. Microbiome Metadata Standards: Report of the National Microbiome Data Collaborative's Workshop and Follow-On Activities. *mSystems* 6. <https://doi.org/10.1128/mSystems.01194-20>.

66. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, Ashburner M, Axelrod N, Baldauf S, Ballard S, Boore J, Cochrane G, Cole J, Dawyndt P, De Vos P, DePamphilis C, Edwards R, Faruque N, Feldman R, Gilbert J, Gilna P, Glöckner FO, Goldstein P, Guralnick R, Haft D, Hancock D, Hermjakob H, Hertz-Fowler C, Hugenholtz P, Joint I, Kagan L, Kane M, Kennedy J, Kowalchuk G, Kottmann R, Kolker E, Kravitz S, Kyripides N, Leebens-Mack J, Lewis SE, Li K, Lister AL, Lord P, Maltsev N, Markowitz V, Martiny J, et al. 2008. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 26:541–547. <https://doi.org/10.1038/nbt1360>.

67. Touré V, Vercruyse S, Acencio ML, Lovering RC, Orchard S, Bradley G, Casals-Casas C, Chaouiya C, Del-Toro N, Flobak Å, Gaudet P, Hermjakob H, Hoyt CT, Licata L, Lægreid A, Mungall CJ, Niknejad A, Panni S, Perfetto L, Porras P, Pratt D, Saez-Rodriguez J, Thieffry D, Thomas PD, Türei D, Kuiper M. 2021. The Minimum Information about a Molecular Interaction CAusal Statement (MI2CAST). *Bioinformatics* 36:5712–5718. <https://doi.org/10.1093/bioinformatics/btaa622>.

68. Lieven C, Beber ME, Olivier BG, Bergmann FT, Ataman M, Babaei P, Bartell JA, Blank LM, Chauhan S, Correia K, Diener C, Dräger A, Ebert BE, Edirisinghe JN, Faria JP, Feist AM, Fengos G, Fleming RMT, García-Jiménez B, Hatzimanikatis V, van Helvoirt W, Henry CS, Hermjakob H, Herrgård MJ, Kaafarani A, Kim HU, King Z, Klamt S, Klipp E, Koehorst JJ, König M, Lakshmanan M, Lee D-Y, Lee SY, Lee S, Lewis NE, Liu F, Ma H, Machado D, Mahadevan R, Maia P, Mardinoglu A, Medlock GL, Monk JM, Nielsen J, Nielsen LK, Nogales J, Nookaei I, Palsson BO, Papin JA, et al. 2020. MEMOTE for standardized genome-scale metabolic model testing. *Nat Biotechnol* 38:272–276. <https://doi.org/10.1038/s41587-020-0446-y>.

69. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, Haraldsdóttir HS, Wachowiak J, Keating SM, Vlasov V, Magnusdóttir S, Ng CY, Preciat G, Žagare A, Chan SHJ, Aurich MK, Clancy CM, Modamio J, Sauls JT, Noronha A, Bordbar A, Cousins B, El Assal DC, Valcarcel LV, Apaolaza I, Ghaderi S, Ahookhosh M, Ben Guebila M, Kostromins A, Sompairac N, Le HM, Ma D, Sun Y, Wang L, Yurkovich JT, Oliveira MAP, Vuong PT, El Assal LP, Kuperstein I, Zinovyev A, Hinton HS, Bryant WA, Aragón Artacho FJ, Planes FJ, Stalidzans E, Maass A, Vempala S, Hucka M, Saunders MA, Maranas CD, et al. 2019. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat Protoc* 14:639–702. <https://doi.org/10.1038/s41596-018-0098-2>.

70. McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D, Wilke A, Huse S, Hufnagle J, Meyer F, Knight R, Caporaso JG. 2012. The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *Gigascience* 1:7. <https://doi.org/10.1186/2047-217X-1-7>.