An Operational Definition of Microbial Keystone Taxa for Advancing Sustainable Microbiome Solutions

Pamela Espíndola-Hernández[1], Samiran Banerjee[2], Abdulkabir O. Abdulmalik[3], Diana Andrade-Linares[1], Germana Baldi[4], Gabriele Berg[5,6], Francis Q. Brearley[7], Cecilia Flocco[8], Luisa Galgani[9], Lina Gegeckienė[10], Silvia Gschwendtner[1], Tim Hensen[11], Tanja Kostic[12], Rodrigo Ledesma-Amaro[13,14,15], Lisa Maier[16,17], Anna Marciniak[18], Juliette Ohan[1], Jörg Overmann[8], Teresa Rito[19,20], Matthew Ryan[21], Stefanie Schulz[1], Selma Vieira[8], **MICROBE consortium**, and Michael Schloter[22,1]

Affiliations:

1. Research Unit for Comparative Microbiome Analysis, Helmholtz Munich, German Research Center for Environmental Health, Neuherberg, Germany
2. Department of Microbiological Sciences, North Dakota State University, ND 58102, USA
3. Institute of Population Genetics, Heinrich-Heine-University Düsseldorf, Universitätsstrasse 1, D-40225 Düsseldorf, Germany; CEPLAS - Cluster of Excellence on Plant Sciences, Heinrich-Heine-University Düsseldorf, Universitätsstrasse 1, D-40225 Düsseldorf, Germany
4. European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, United Kingdom
5. Institute of Environmental Biotechnology, Graz University of Technology, Petersgasse 12, Graz, 8010 Austria
6. Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), Max-Eyth Allee 100, 14469 Potsdam, Germany
7. Department of Natural Sciences, Manchester Metropolitan University, Chester Street, Manchester, M1 5GD, UK
8. DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany
9. Department of Biotechnology, Chemistry and Pharmacy, University of Siena, Italy; National Biodiversity Future Center (NBFC), Italy
10. Life Sciences Center, Institute of Biosciences, Vilnius University, LT-10257 Vilnius, Lithuania
11. School of Medicine, Digital Metabolic Twin Centre, University of Galway, Galway, Ireland
12. AIT Austrian Institute of Technology GmbH, Tulln, Austria
13. Department of Bioengineering and Imperial College Centre for Synthetic Biology, Imperial College London, London SW7 2AZ, UK
14. Bezos Centre for Sustainable Protein, Imperial College London, London SW7 2AZ, UK
15. UKRI Engineering Biology Mission Hub on Microbial Food, Imperial College London, London SW7 2AZ, UK
16. Interfaculty Institute for Microbiology and Infection Medicine Tübingen, University of Tübingen, Tübingen, Germany
17. Cluster of Excellence EXC 2124 Controlling Microbes to Fight Infections, University of Tübingen, Tübingen, Germany
18. Alina Pienkowska Center for Cancer Prevention and Epidemiology, Poznan, Poland
19. Centre of Molecular and Environmental Biology (CBMA), University of Minho, Portugal
20. Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Portugal
21. Genetic Resources Collection, CABI, Egham, United Kingdom
22. Chair for Environmental Microbiology, TUM School of Life Sciences, Technical University Munich, Germany

Abstract

Developing synthetic communities (SynComs) and harnessing their functional potential to create microbe-based solutions across diverse applications, including agriculture, biotechnology, and medicine, holds promise for improving environmental and human health. However, defining the composition of SynComs is challenging due to the enormous microbial diversity in most ecosystems. Implementing the keystone taxa concept into microbiome research may help develop strategies for defining SynComs and prioritizing the isolating of microbes that catalyze critical functions in a particular environment. The idea was developed in the '60s of the last century for macroorganisms and describes organisms that are disproportionately important in an ecosystem compared to others. Despite this, a comprehensive definition of microbial keystone taxa has remained elusive. Here, we provide an 8-point updated definition that characterizes its ecological significance and could contribute to its identification. We further describe an operational 4-step approach to detect “keystone taxa candidates” from complex “omics” data and propose strategies for their isolation.

Keywords:

Keystone Taxa; SynCom; Core Microbiome; Co-occurrence Analysis; Metagenomics; Metabarcoding;

Introduction

Nature provides ecosystem services that are essential for our life on Earth. In addition, our direct interaction with nature significantly impacts our health, influencing well-being and disease prevalence[1]. However, compared to pre-industrial times, our environment has changed dramatically due to human activities, resulting in habitat destruction, biodiversity loss, pollution, climate change, overexploitation of species, and the introduction of invasive organisms[2]. These changes have been considered potential causes for notable increases in environment-related diseases in recent years, including infections, cancer, allergies, and metabolic disorders[3]. The growing awareness of the connection between environmental and human health is reflected in the relatively recent development of holistic concepts, such as One Health, Planetary Health, or Global Health[4]. In these concepts, environmental microbes and their functional potential are recognized as architects of the ecosystem services provided by nature. Therefore, changes in the environmental microbiome or losses of diversity have been considered a reason for losses of ecosystem services and reduced resilience towards environmental stressors. Further, environmental microbes act as a source of microbes, which humans recruit and implement into their microbiome. Thus, changes in the environmental microbiome are mirrored in our microbiome and may induce functional changes, which have been linked to dysbiotic states and disease[5]. As a consequence, preventing or mitigating the loss of microbial diversity in nature is crucial for promoting both environmental and human health.

As the natural re-establishment of microbial biodiversity at an affected site might take decades or even centuries, in recent years, a growing field of research has focused on compensating for microbial losses through the targeted reintroduction of microbiota, either as single strains or synthetic communities (SynCom) into affected ecosystems[6]. This approach has gained significant interest, for example, in agriculture, where “biostimulants” (*i.e.*, environmental probiotics) have been proven to reduce the use of pesticides and the amount of inorganic fertilizers[7]. In human medicine, losses in microbial diversity, *e.g.*, after antibiotic therapy, are compensated by using beneficial microbiota (*i.e.*, probiotics) to reestablish a healthy gut microbiome[8]. Recently, industrial biotechnology has also started to implement SynComs in fermentation and bioprocesses, which can improve bioproduction yields[9].

The successful application of SynComs is context-dependent, often limited by the fact that the needed microbes often remain uncultured due to their specific physiological conditions or the need for a minimal microbial consortium for their successful establishment[9]. Thus, a vast part of current microbiome research aims to develop targeted isolation approaches, using new omics techniques to identify “the most important” microbes to be prioritized[10]. However, this

can be very challenging, considering the vast microbial diversity in most ecosystems. For example, one gram of soil may harbor more than 54,000 different microbial species, and the number of strains might be much higher[11]. Their unique eco-physiological needs hinder the successful isolation of most microbial diversity based on current technologies.

Implementing the microbial keystone taxa concept might help prioritize microbes for targeted isolation and implementation into SynComs. However, although the term has been frequently used in the microbiome field, it lacks a clear and unifying definition. Here, we provide an 8-point updated definition, considering its crucial features to understand its ecological significance and eco-physiological needs for isolation. We further describe an operational 4-step methodological approach to detect "keystone taxa candidates" from large-scale *omics* data and define strategies for its isolation.

How do we define microbial keystones?

Since Paine introduced the idea of "keystone species" in 1969[12], this concept has been a central point of interest in ecology and related fields. At that time, Paine was studying a shore ecosystem, and he realized that the presence of one particular species could mostly control the ecosystem structure and functioning. That species was the sea star, and he labeled this species a "keystone species" for that ecosystem[12]. In the following years, Paine and colleagues generalized this finding and proposed that "keystone species" disproportionately impact their community or ecosystem relative to their abundance[13] (Box 1.1 **High impact independent of abundance**). After its first appearance, the term "keystone species" has been broadly used in macroecology, especially in food-web studies[14].

More recently, the "keystone species" concept was extended to microbiology in line with sequencing technology and computational infrastructure developments. The current definition of *microbial keystone taxa* refers to specific microorganisms within a microbial community that play critical roles in maintaining the overall structure, function, and stability of a given microbial community or an ecosystem[15]. Thus, microbial keystone taxa may contribute to essential ecosystem functions such as nutrient cycling, organic matter degradation, energy flow, pathogen suppression, and regulation of the host immune system (Box 1.2 **Functional Importance**) or regulate the abundance and activity of other microorganisms within the community through various mechanisms. This regulation may involve competition, predation, symbiosis, or the production of signaling molecules that modulate microbial interactions and phenotypes (Box 1.3 **Ecological Interactions**). Microbial keystone taxa may have specialized metabolic capabilities that enable them to perform unique or critical roles within their

community, occupying specific ecological niches within their habitat and exploiting resources or environmental conditions less accessible to other organisms. Their ability to thrive in these niches may be facilitated by specialized adaptations or metabolic pathways (Box 1.4 **Niche Specialization**). Consequently, their presence and activity may create microhabitats, alter nutrient availability, or modify physicochemical conditions, shaping the broader ecosystem landscape (Box 1.5 **Ecological Engineering**). The presence of keystone taxa enhances the stability and resilience of microbial communities. The processes behind these phenomena are still not well understood, but might be linked to larger genome sizes or highly structured genomes (Box 1.6 **Stability and Resilience**). This feature might enhance their ecological roles and adaptability, not only to themselves but also to the community as other microbiota might incorporate genetic material from the keystone taxa (Box 1.7 **Genetic material reservoir**). *Vice versa*, genetic material could also be transferred to keystone taxa genomes; arguing that keystone taxa are active in their particular environment due to their central position in the community network, they might have an enhanced possibility to uptake genetic information from other microorganisms compared to those microbes which are less active or dormant — resulting in higher adaptability under changing environmental conditions (for resent cases see [16,17]; Box 1.8 **Uptake of genetic material**).

From species to strains

However, like many other concepts derived from macroecology, applying the keystone taxa concept to the microbial world has several potential constraints that must be considered. The limitations range from the difficulties in directly observing microbial interactions to their high spatial and temporal fluctuations and issues related to the exchange of genetic information via horizontal gene transfer, among others[18]. The latter –horizontal gene transfer and other factors – hinders species-level resolution in microbial classification, which raises the question of the suitable taxonomic level for identifying keystone microbes. Considering that horizontal gene transfer, rearrangements in the genome, and gene losses can only be detected at the strain level and that the current definition of a microbial species corresponds to all strains with a DNA similarity of > 70 % in their genome[19], we propose that strain-specific resolution should be considered as the suitable level to define keystone microbes. For example, strains from the same species can be health-promoting or detrimental to a particular host, depending on their genomes[20]. Additionally, bacteria and archaea frequently have an increased genetic change rate in relation to plants and animals, raising the question about the stability of a particular strain in its natural environment. Indeed, laboratory experiments have shown transfers of plasmids from recipients to donors within minutes[21] and a very fast rearrangement of genomes to scope with the applied stress[21,22]. As the process of horizontal gene transfer and genome rearrangement is energy-consuming, transfer rates might be significantly lower in nature, but

the dynamics of genomes of strains might still be high. Hence, when referring to microorganisms, the term “keystone taxa”[15] is preferred instead of “keystone species”, and the suitable taxonomic resolution depends on whether a function is essential for the life of a bacterial species and part of the corresponding core genome or only present in the pangenome and harbored only by some strains. Moreover, it needs to be discussed, taking the large spatiotemporal dynamics into account if the status of a microbial keystone taxa is permanent – inherent to the taxon itself and its environment – or whether its keystone status might change according to time/space variation such as nutrient availability, pH, temperature, or oxygen gradients, etc [23]. In other words, does a particular ecosystem always depend on the same keystone taxa? Or might the keystone status circulate across taxa? Taking recent findings about the different activity patterns of microbes from living to dormant into account[24], the keystone taxa status of an organism might also change in a given ecosystem depending on the environmental conditions present. However, a certain resilience towards natural fluctuations should be expected to allow for the stabilization of the ecosystem itself as well as its microbiome.

Box 1. Components of a Microbial Keystone Taxa Definition

COMPONENT	EXPLANATION	ASSESSMENT METHOD
1 High impact independent regardless of abundance	Disproportionate influence on the ecosystem despite often being in low abundance.	High-resolution taxonomic identification
2 Functional Importance	Critical functions in ecosystem and organismal health.	Metagenomics Metabolic profiling
3 Community shape	Ability to drive ecological interactions and regulate microbial community dynamics, often preventing the overgrowth of other species populations.	Co-occurrence networks analysis Experimental interaction assays Perturbation experiments Long-term field experiments
4 Niche Specialization	Unique ecological roles often thrive under specific environmental conditions.	Environmental gradients Metabolic profiling
5 Stability and Resilience	Contribution to the stability and recovery of microbial communities following environmental disturbances.	Perturbation experiments Long-term field experiments

6	**Ecological Engineering**	Ability to modify its environment can shape the ecosystem by, for example, altering nutrient availability or oxygen levels,	Controlled experiments
7	**Genetic material reservoir**	Genomes can contain key genetic material, including plasmids or resistance genes that enhance community survival.	Genome-wide analyses Identification of transposable elements
8	**Uptake of genetic material**	Capacity to acquire genetic material from its environment drives evolutionary adaptability.	Genome-wide analyses Identification of transposable elements

From definition to implementation: A toolbox to define microbial keystone taxa

Although a number of in vivo approaches based on microevolution have been used successfully to identify SynComs with certain functional properties D, still in silico tools are more often used to identify microbial keystone taxa [25]. This requires a synergistic and integrative approach, combining the power of high-throughput sequencing technologies and bioinformatics with a comprehensive understanding of microbial ecology for a particular ecosystem (Figure 1).

Step I includes the definition of a “candidate microbial keystone taxa”. **Network or co-occurrence analysis** using metabarcoding data based on marker genes like the 16S rRNA gene for bacteria and archaea, the ITS region of the ribosomal operon for fungi or the 18S rRNA gene for protists is often used for this purpose (Figure 1a). In the original study on the role of sea stars, Paine intuitively represented the shore community as a net of species connected by lines. However, formal network thinking in ecology was just introduced in 2005, when Proulx and colleagues, taking advantage of the vast theoretical background from the graph theory, applied the concept of co-occurrence networks to the study of ecological interaction[26]. Even though network analysis has proved to be a valuable tool for representing community structure, some aspects hinder the direct application to studying environmental microbiomes[27–29]. One relevant issue is that network analyses – based on co-occurrence or its derived associations – allow the identification of hubs with a high influence on the network morphology. These hubs can be interpreted as “keystone taxa candidates” as their removal/loss significantly alters network morphology. However, the network represents putative ecological associations that require further assessment to gain mechanistic insights into the particular community. The cooccurrence networks also may also give information

about the stability of certain taxa under changing conditions of an ecosystem if samples from different time points are available. These microbes might be considered as part of the **Core microbiome** of an ecosystem (Fig. 1b) [30,31] and are typically considered important for maintaining the stability, functionality, and health of an ecosystem facing natural dynamics in chemical and physical properties, an essential feature of a keystone taxa (see above). However, in contrast to macro-organisms, where these interactions are visible and functional dependencies between organisms can be easily assessed, the microbial world network analysis is a purely statistical analysis, and links to functions are missing. Therefore, a functional annotation of the "candidate microbial keystone taxa" is needed. **Functional analysis** (Fig. 1c) could include high throughput metabolic *in vitro* profiling[32], single gene annotation, stable isotope probing (SIP), and, more recently, metagenomics, meta-transcriptomics, metabolomics, genome-scale metabolic modeling[33], among others[34]. These approaches provide insights into microbiome biosynthetic pathways and metabolic activities, contributing to understanding the functional roles of specific microorganisms within a community or environment. "Candidate keystone taxa" inferred from metabarcoding data can then be mapped with functional data, for example, by mapping to Metagenome-assembled genomes (MAGs). MAGs can be used to characterize the functional potential of the identified keystone taxa and to define specific strategies for targeted isolation of the microbiota of interest.

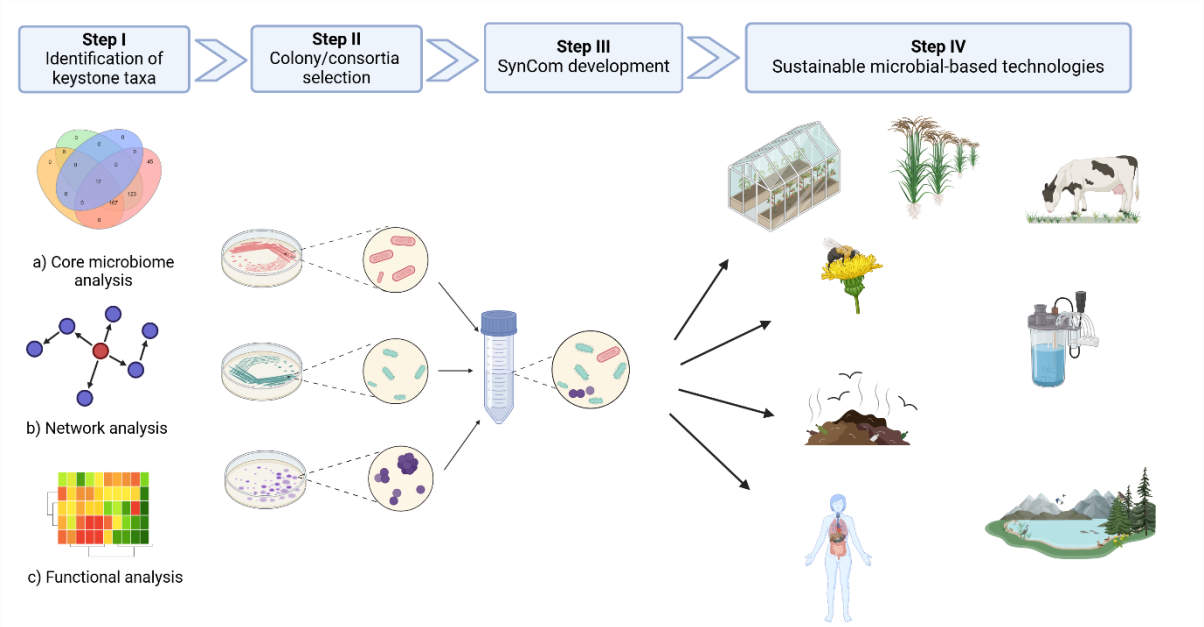

Fig. 1. Identification of microbial keystone taxa and SynComs assembly for sustainable microbiome solutions. Step I (a, b, and c) include data-analysis techniques for identifying the environment's keystone taxa: a) Core microbiome analysis involves metabarcoding taxonomic identification to determine shared taxa among two or more microbial communities in a given host species or environment. b) Network analysis is also usually based on

metabarcoding data to determine co-occurrence associations between taxa of the microbial community, allowing the identification of hubs that strongly influence network morphology. c) Functional analysis could include high throughput metabolic *in vitro* profiling[32], single gene annotation, stable isotope probing (SIP), and, more recently, meta-transcriptomics and metabolomics, among others[34]. **Step II** includes a combination of isolation techniques that target the selected taxa or consortia from step I. **Step III** includes developing the SynComs, including development, testing, and preservation. **Step IV** to the implementation of the designed SynCom into an affected environment. Created in https://BioRender.com.

Step II includes the subsequent targeted isolation of the "candidate keystone taxa". The development of media and the use of subsequent helper microbes for the isolation play a key role here, which can be facilitated by the functional mapping described above. After successful isolation, "classical microbiology" starts by confirming the functional traits defined *in silico* with *in vitro* and *in vivo* tests, taking Koch's postulates from 1893 into account, stating that any postulated trait of a microorganism must be confirmed with the isolates of the respective strain at *in vitro* and *in vivo* models. Koch referred to pathogenic microorganisms, but the statement also refers to microbes with potential beneficial environmental and human health traits. This may also include the construction of knockout mutants and their supplementation, which may become challenging if the protocols for targeted manipulations of desired microbes are still missing.

Step III is SynCom development, which combines the selected strains at specific proportions, generally *in vitro*, under controlled conditions. The design of the SynCom should consider not only the strains' requirements, *i.e.*, nutrient availability, physicochemical conditions, etc. but also the interactions between strains, which determine the functionality of a SynCom to a large extent. Although synergies between strains may be essential to obtain the proposed activity pattern of a SynCom, and some genes might be expressed only if a certain combination of microbes is present, at the same time, negative interactions need to be considered as microbes may also interfere with each other, in a way that growth and/or activity is reduced or even completely blocked, *e.g.*, by the formation of substances with antimicrobial properties by single members of the SynCom. This third step might imply a loop of sub-steps of testing and tuning the culture conditions, namely the design-build-test-learn cycle (DBTL)[35]. This step should include testing and preservation sub-steps.

Step IV, the final step, corresponds to implementing the established SynCom into applications. Any SynCom application must be considered as an invasion of microbes into an ecosystem. Thus, invasion theories derived from macroecology may help determine the inoculation's success in advance. In plant seeds, for example, due to the priority effects, SynCom strains

have shown low environmental invasion capacity but they can modify the seed microbiota composition, the community size, and the overall seedling[36]. In general, implementations of SynComs are more likely to be successful when applied after a significant disturbance or when the environment is naturally low in microbial diversity. Although not real SynComs but more quite undefined microbiota, this principle has already been implemented into clinical practice when using fecal transplants[37]. For complex ecosystems, using carrier materials to protect the inoculated microbes from grazing and allow for a stepwise adaptation to the new conditions has been proven to increase the success rate for colonization of an ecosystem after introducing single microbial strains or SynComs[38]. Carriers used successfully for soils can be of organic (*e.g.*, compost, biogas slurry, crushed corn cob, biochar, peat, etc.) or inorganic origin (*e.g.*, zeolite, perlite, lignite, talc, etc.)[39]. The final phenotype of the SynComs in vivo also depends on the environmental conditions present and can only be predicted for housekeeping genes or genes with constitutive expression. However, for most genes linked to the synthesis of beneficial substances or the degradation of harmful substances, the expression is highly regulated, and phenotypes must be assessed depending on the chemical and physical conditions present in the ecosystem.

In summary, identifying and understanding microbial keystone taxa is crucial for developing SynComs and unlocking their potential applications across diverse fields, including agriculture, bioremediation, and human health. We provide an 8-point updated definition of microbial keystone taxa, linking each characteristic to its ecological significance and highlighting its unique roles in microbial ecosystems. To prioritize microbes for isolation and SynComs development, we suggest a 4-step operational approach based on complementing cultivation-based approaches with methodologies that bring information about the structure and functionality of the microbial community. Successfully implementing a SynCom designed to evaluate the role of a "keystone taxa candidate" could confirm its status as a keystone taxon.

ACKNOWLEDGMENTS

This work was supported by the European Union's Horizon Europe research and innovation programme under grant agreement No. 101094353. We thank all the speakers and participants of the "MICROBE workshop – Keystone Taxa and Core Microbiome" on 15th January 2024 for their valuable contributions and insightful discussions, which served as a foundation for this article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Seymour, V. (2016). The human-nature relationship and its impact on health: A critical review. Front Public Health *4*, 1–12. https://doi.org/10.3389/FPUBH.2016.00260.

2. Corbett, S., Courtiol, A., Lummaa, V., Moorad, J., and Stearns, S. (2018). The transition to modernity and chronic disease: Mismatch and natural selection. Nat Rev Genet *19*, 419–430. https://doi.org/10.1038/s41576-018-0012-3.

3. Prüss-Ustün, A., Wolf, J., Corvalán, C., Bos, R., and Neira, M. (2016). Preventing disease through healthy environments: A global assessment of the burden of disease from environmental risks. (World Health Organization).

4. Cunningham, A.A., Daszak, P., and Wood, J.L.N. (2017). One health, emerging infectious diseases and wildlife: Two decades of progress? Philosophical Transactions of the Royal Society B: Biological Sciences *372*. https://doi.org/10.1098/rstb.2016.0167.

5. Banerjee, S., and van der Heijden, M.G.A. (2023). Soil microbiomes and one health. Nat Rev Microbiol *21*, 6–20. https://doi.org/10.1038/s41579-022-00779-w.

6. Großkopf, T., and Soyer, O.S. (2014). Synthetic microbial communities. Curr Opin Microbiol *18*, 72–77. https://doi.org/10.1016/j.mib.2014.02.002.

7. Jansson, J.K., and Hofmockel, K.S. (2020). Soil microbiomes and climate change. Nat Rev Microbiol *18*, 35–46. https://doi.org/10.1038/s41579-019-0265-7.

8. Goldin, B.R., and Gorbach, S.L. (2008). Clinical Indications for Probiotics: An Overview. Clinical Infectious Diseases *46*, S96–S100.

9. McCarty, N.S., and Ledesma-Amaro, R. (2019). Synthetic Biology Tools to Engineer Microbial Communities for Biotechnology. Trends Biotechnol *37*, 181–197. https://doi.org/10.1016/j.tibtech.2018.11.002.

10. Zhang, Z.F., Liu, F., Liu, L.R., Li, M., Cai, L., Liu, S., and Mao, J. (2024). Culturing the uncultured marine fungi in the omics age: Opportunities and challenges. Fungal Biol Rev *48*, 100353. https://doi.org/10.1016/J.FBR.2023.100353.

11. Roesch, L.F.W., Fulthorpe, R.R., Riva, A., Casella, G., Hadwin, A.K.M., Kent, A.D., Daroub, S.H., Camargo, F.A.O., Farmerie, W.G., and Triplett, E.W. (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. ISME J *1*, 283–290. https://doi.org/10.1038/ISMEJ.2007.53.

12. Paine, R.T. (1969). The Pisaster-Tegula Interaction : Prey Patches , Predator Food Preference , and Intertidal Community Structure. Ecology *50*, 950–961.

13. Power, M.E., Tilman, D., Estes, J.A., Menge, B.A., Bond, W.J., Mills, L.S., Daily, G., Castilla, J.C., Lubchenco, J., and Paine, R.T. (1996). Challenges in the quest for keystones: Identifying keystone species is difficult-but essential to understanding how loss of species will affect ecosystems. Bioscience *46*, 609–620. https://doi.org/10.2307/1312990.

14. Jordán, F. (2009). Keystone species and food webs. Philosophical Transactions of the Royal Society B: Biological Sciences *364*, 1733–1741. https://doi.org/10.1098/RSTB.2008.0335.

15. Banerjee, S., Schlaeppi, K., and van der Heijden, M.G.A. (2018). Keystone taxa as drivers of microbiome structure and functioning. Nat Rev Microbiol *16*, 567–576. https://doi.org/10.1038/s41579-018-0024-1.

16. Zhang, Z., Zhang, L., Zhang, L., Chu, H., Zhou, J., and Ju, F. (2024). Diversity and distribution of biosynthetic gene clusters in agricultural soil microbiomes. mSystems *9*, e01263-23. https://doi.org/10.1128/msystems.01263-23.

17. Li, L., Meng, D., Yin, H., Zhang, T., and Liu, Y. (2023). Genome-resolved metagenomics provides insights into the ecological roles of the keystone taxa in heavy-metal-contaminated soils. Front Microbiol *14*, 1–13. https://doi.org/10.3389/fmicb.2023.1203164.

18. Cottee-Jones, H.E.W., and Whittaker, R.J. (2012). The keystone species concept: a critical appraisal. Front Biogeogr *4*. https://doi.org/10.21425/f5fbg12533.

19. Riesco, R., and Trujillo, M.E. (2024). Update on the proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol *74*, 1–12. https://doi.org/10.1099/ijsem.0.006300.

20. Siani, R., Stabl, G., Gutjahr, C., Schloter, M., and Radl, V. (2021). Acidovorax pan-genome reveals specific functional traits for plant beneficial and pathogenic plant-associations. Microb Genom *7*. https://doi.org/10.1099/MGEN.0.000666.

21. Darling, A.E., Miklos, I., and Ragan, M.A. (2008). Dynamics of Genome Rearrangement in Bacterial Populations. PLoS Genet *4*, e1000128. https://doi.org/10.1371/journal.pgen.1000128.

22. Peck, L.D., Id, T.L., Bennetot, B., O', S., Id, D., Nowellid, R.W., Ryan, M.J., Flood, J., Rodríguez De La Vegaid, R.C., Roparsid, J., et al. (2024). Horizontal transfers between fungal Fusarium species contributed to successive outbreaks of coffee wilt disease. PLoS Biol *22*, e3002480. https://doi.org/10.1371/JOURNAL.PBIO.3002480.

23. Baveye, P.C., Otten, W., Kravchenko, A., Balseiro-Romero, M., Beckers, É., Chalhoub, M., Darnault, C., Eickhorst, T., Garnier, P., Hapca, S., et al. (2018). Emergent properties of microbial activity in heterogeneous soil microenvironments: Different research approaches are slowly converging, yet major challenges remain. Front Microbiol *9*, 1–48. https://doi.org/10.3389/fmicb.2018.01929.

24. Blagodatskaya, E., and Kuzyakov, Y. (2013). Active microorganisms in soil: Critical review of estimation criteria and approaches. Soil Biol Biochem *67*, 192–211. https://doi.org/10.1016/j.soilbio.2013.08.024.

25. Dubey, S., and Sharma, S. (2021). Rhizospheric Engineering by Plant-Mediated Indirect Selection of Microbiome for Agricultural Sustainability. CRC Crit Rev Plant Sci *40*, 379–397. https://doi.org/10.1080/07352689.2021.1959137.

26. Proulx, S.R., Promislow, D.E.L., and Phillips, P.C. (2005). Network thinking in ecology and evolution. Trends Ecol. Evol. *20*, 345–353. https://doi.org/10.1016/j.tree.2005.04.004.

27. Goberna, M., and Verdú, M. (2022). Cautionary notes on the use of co-occurrence networks in soil ecology. Soil Biol Biochem *166*, 108534. https://doi.org/10.1016/j.soilbio.2021.108534.

28. Faust, K. (2021). Open challenges for microbial network construction and analysis. ISME Journal *15*, 3111–3118. https://doi.org/10.1038/s41396-021-01027-4.

29. Freilich, M.A., Wieters, E., Broitman, B.R., Marquet, P.A., and Navarrete, S.A. (2018). Species co-occurrence networks: Can they reveal trophic and non-trophic interactions in ecological communities? Ecology *99*, 690–699. https://doi.org/10.1002/ECY.2142.

30. Neu, A.T., Allen, E.E., and Roy, K. (2021). Defining and quantifying the core microbiome: Challenges and prospects. Proc Natl Acad Sci U S A *118*. https://doi.org/10.1073/pnas.2104429118.

31. Risely, A. (2020). Applying the core microbiome to understand host–microbe systems. Journal of Animal Ecology *89*, 1549–1558. https://doi.org/10.1111/1365-2656.13229.

32. Garland, J.L., and Mills, A.L. (1991). Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole-carbon-source utilization. Appl Environ Microbiol *57*, 2351–2359. https://doi.org/10.1128/aem.57.8.2351-2359.1991.

33. Heinken, A., Basile, A., and Thiele, I. (2021). Advances in constraint-based modelling of microbial communities. Curr Opin Syst Biol *27*, 100346. https://doi.org/10.1016/j.coisb.2021.05.007.

34. Arıkan, M., and Muth, T. (2023). Integrated multi-omics analyses of microbial communities: a review of the current state and future directions. Mol Omics *19*, 607–623. https://doi.org/10.1039/d3mo00089c.

35. Lawson, C.E., Harcombe, W.R., Hatzenpichler, R., Lindemann, S.R., Löffler, F.E., O'Malley, M.A., García Martín, H., Pfleger, B.F., Raskin, L., Venturelli, O.S., et al. (2019). Common principles and best practices for engineering microbiomes. Nat Rev Microbiol *17*, 725–741. https://doi.org/10.1038/s41579-019-0255-9.

36. Arnault, G., Marais, C., Préveaux, A., Briand, M., Poisson, A.S., Sarniguet, A., Barret, M., and Simonin, M. (2024). Seedling microbiota engineering using bacterial synthetic community inoculation on seeds. FEMS Microbiol Ecol *100*. https://doi.org/10.1093/femsec/fiae027.

37. de Groot, P.F., Frissen, M.N., de Clercq, N.C., and Nieuwdorp, M. (2017). Fecal microbiota transplantation in metabolic syndrome: History, present and future. Gut Microbes *8*, 253–267. https://doi.org/10.1080/19490976.2017.1293224.

38. Aksani, D., Surono, Ginting, R.C.B., and Purwani, J. (2021). The assay of carrier material and bacteria isolate formula as a biofertilizer on soybean in Inceptisols from West Java. IOP Conf Ser Earth Environ Sci *648*, 012193. https://doi.org/10.1088/1755-1315/648/1/012193.

39. Sohaib, M., Zahir, Z.A., Khan, M.Y., Ans, M., Asghar, H.N., Yasin, S., and Al-Barakah, F.N.I. (2020). Comparative evaluation of different carrier-based multi-strain bacterial formulations to mitigate the salt stress in wheat. Saudi J Biol Sci *27*, 777–787. https://doi.org/10.1016/J.SJBS.2019.12.034.