The need for an integrated multi-OMICs approach in microbiome science in the food system

Ilario Ferrocino, Kalliopi Rantsiou, Ryan Mcclure, Tanja Kostic, Rafael Soares Correa de Souza, Lene Lange, Jamie Fitzgerald, Aicha Kriaa, Paul Cotter, Emmanuelle Maguin, et al.

Received: 10 July 2022 | Revised: 5 December 2022 | Accepted: 19 December 2022

DOI: 10.1111/1541-4337.13103

COMPREHENSIVE REVIEW

Comprehensive REVIEWS in Food Science and Food Safety WILEY

The need for an integrated multi-OMICs approach in microbiome science in the food system

Ilario Ferrocino[1] | **Kalliopi Rantsiou**[1] | **Ryan McClure**[2] | **Tanja Kostic**[3] | **Rafael Soares Correa de Souza**[4] | **Lene Lange**[5] | **Jamie FitzGerald**[6] | **Aicha Kriaa**[7] | **Paul Cotter**[6] | **Emmanuelle Maguin**[7] | **Bettina Schelkle**[8] | **Michael Schloter**[9] | **Gabriele Berg**[10] | **Angela Sessitsch**[3] | **Luca Cocolin**[1] | **The MicrobiomeSupport Consortium**

[1]Department of Agriculture, Forest and Food Science, University of Turin, Grugliasco, Italy

[2]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, USA

[3]AIT Austrian Institute of Technology GmbH, Bioresources Unit, Tulln, Austria

[4]Genomics for Climate Change Research Center (GCCRC), Universidade Estadual de Campinas (UNICAMP), Campinas, São Paulo, Brazil

[5]BioEconomy, Research & Advisory, Valby, Denmark

[6]Teagasc Food Research Centre, Moorepark, Fermoy, County Cork, Ireland

[7]MICALIS, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France

[8]European Food Information Council, Brussels, Belgium

[9]Helmholtz Zentrum München, Oberschleissheim, Germany

[10]Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria

Correspondence
Luca Cocolin, Department of Agriculture, Forest and Food Science, University of Turin, Largo Braccini 2, 10095 Grugliasco, Italy.
Email: lucasimone.cocolin@unito.it.

Funding information
European Union's Horizon 2020 coordinated support action, Grant/Award Number: 818116

Abstract

Microbiome science as an interdisciplinary research field has evolved rapidly over the past two decades, becoming a popular topic not only in the scientific community and among the general public, but also in the food industry due to the growing demand for microbiome-based technologies that provide added-value solutions. Microbiome research has expanded in the context of food systems, strongly driven by methodological advances in different -omics fields that leverage our understanding of microbial diversity and function. However, managing and integrating different complex -omics layers are still challenging. Within the Coordinated Support Action MicrobiomeSupport (https://www.microbiomesupport.eu/), a project supported by the European Commission, the workshop "Metagenomics, Metaproteomics and Metabolomics: the need for data integration in microbiome research" gathered 70 participants from different microbiome research fields relevant to food systems, to discuss challenges in microbiome research and to promote a switch

from microbiome-based descriptive studies to functional studies, elucidating the biology and interactive roles of microbiomes in food systems. A combination of technologies is proposed. This will reduce the biases resulting from each individual technology and result in a more comprehensive view of the biological system as a whole. Although combinations of different datasets are still rare, advanced bioinformatics tools and artificial intelligence approaches can contribute to understanding, prediction, and management of the microbiome, thereby providing the basis for the improvement of food quality and safety.

KEYWORDS
Omics integration, microbiome, food system, metagenomics, metatranscriptomics, metaproteomics, metabolomics

1 | INTRODUCTION

Microbiome science is an interdisciplinary research field connected to diverse areas, for example, agriculture, food science, biotechnology, bioeconomy, mathematics (informatics, statistics, modeling), plant pathology, and human medicine. The term microbiome refers not only to the microorganisms involved but also encompasses their theatre of activity, which results in the formation of specific ecological niches (Berg et al., 2020). The microbiome, which forms a dynamic and interactive micro-ecosystem prone to change in time and scale, is integrated with macro-ecosystems, including eukaryotic hosts, becoming crucial for their functioning and health.

The microbiome has attracted a lot of attention from both researchers and policy makers and was defined as a pathway for action within the Food 2030 strategy of the European Commission (https://research-and-innovation.ec.europa.eu/research-area/environment/bioeconomy/food-systems/food-2030_en). The microbiome can be one of the major game-changers in how we manage our planet's resources, allowing us to produce food in a more sustainable way, in line with improvements under the "One Health" approach. It has the potential to beneficially impact primary food production and advance sustainable agriculture, as well as food science, human health, and waste management (https://knowledge4policy.ec.europa.eu/publication/food-2030-pathways-action-%E2%80%93-microbiome-world_en). This is strongly aligned with our current understanding of the food system, defined as "a network that integrate[s] the food value chain up to consumption and goes beyond the farm-to-fork principle by including all activities, actors, drivers, boundaries as well as input factors and various dimensions and forms of outcomes" (https://etp.fooddrinkeurope.eu/news-and-publications/publications/30-etp-food-for-life-sria-2021.html). The detailed representation of food systems requires a comprehensive and holistic understanding of the interactions within it, as well as a high degree of interdisciplinarity, as the basis for the successful transition to sustainable food systems. Microbiome science can effectively help in reaching these objectives.

Microbiome research has been strongly driven by advances in DNA sequencing technologies (often referred to as next-generation sequencing, NGS). With the advent of DNA sequencing and high-throughput technologies applied in all fields of biological sciences, we are able to generate billions of data points, which can be used for an in-depth characterization of the structure, function, interaction, and complexity of microbial ecosystems. Indeed, we are now able to track and map shifts in microbial communities, discover new molecules, new metabolic pathways, new adaptation strategies, and also new strains (Ferrocino et al., 2022). However, the analytical methodology needed to model microbiome data and integrate it with genomics, transcriptomics, proteomics, metabolomics, or other "-omics" data, remains nascent (Ferrocino et al., 2022). This toolset frequently uses techniques developed in other multi-omics investigations, especially the growing array of statistical and computational techniques for integrating and representing data. Thus, to best implement these complementary techniques from parallel fields of analytical biology, minimum requirements are needed both for the experimental work and for the bioinformatical/biostatistical data analysis to meaningfully integrate -omics data and discern their functional linkages.

In this frame, the workshop was organized to discuss the advantages and disadvantages of various study designs, methodologies, and statistical tools and methods of integration in order to assess their applicability to microbiome data and discuss their biological interpretability in the context of food systems. We also highlighted ongoing statistical challenges and opportunities in integrating multi-omics

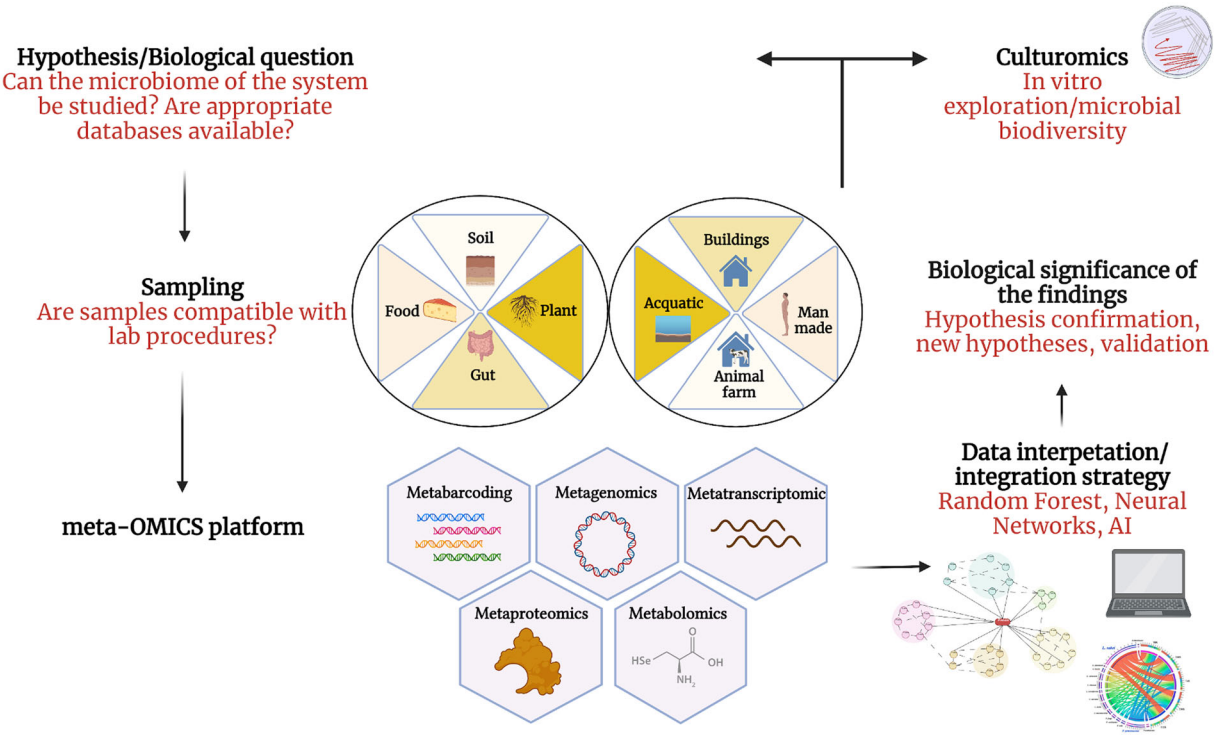

FIGURE 1 A scheme highlighting the most important points a researcher should consider when planning a microbiome experiment. Choosing the right sample size and sampling procedure to observe shift/modification in the microbiomes and to test the primary hypothesis; selection of target organisms and molecules and related -omics platform, sampling procedures, and storage; selection of meta-analysis interpretation/integration strategies to test the hypothesis; validation of the -omics output with an extensive single cell or multiple cells culture-based approach are some of them

data with prior knowledge in order to generate hypotheses with regard to biological function, role, and impact.

Microbiomes are widely well-characterized by simultaneously measuring thousands of molecules in biological samples (DNA, RNA, proteins, and metabolites) using various -omics technologies, such as metagenomics, metatranscriptomics, metaproteomics, and metabolomics. It is becoming increasingly common to apply two or more -omics technologies in parallel with techniques recently defined as "trans-omics analysis" (Dugourd et al., 2021). However, the choice of -omic tools usually depends on the technical/analytical skills and resources of a research group, previous experience with those technologies, availability of a data analyst/platform, as well as the availability of funds. Several aspects must be considered when an experiment is planned as shown in Figure 1. This review builds on these considerations while specifically focusing on the design of microbiome studies and on the development of downstream statistical analysis plans which answer complex questions in the food system in a translational setting.

2 | STRATEGIES THAT SHOULD BE IMPLEMENTED FOR A SUCCESSFUL OUTCOME OF A MULTI-OMICS STUDY

Due to their nature, -omics data are intrinsically both highly variable and noisy, leading to several issues when trying to compare or reproduce them. When planning an experiment, several steps should be considered: experimental design and challenges, individual -omics datasets, integration and data issues, and biological knowledge. Here, we present a summary of the workshop discussions in the context of these key themes.

2.1 | Experimental design and sampling procedures

Experiments are often planned for one individual -omics tool (e.g., genomics, transcriptomics, or proteomics), but in some cases, once the experiment is carried out, new analyses can be required or added to verify or validate

the hypothesis (O'Donnell et al., 2020). Sample preparation and collection, if not properly performed, may prevent accurate analysis with complementary techniques. Considerations that should be considered when an experiment is planned include: method of collection, quantity and selection of biological samples, number of technical replicates, preservation and storage techniques (Ryan, Schloter, Berg, Kostic, et al., 2021) this reference is the one published in Trends in Microbiology, use of standards for internal controls and standardized procedures, and identification of limitations of available -omics platforms. For a given study, the procedures chosen may be suitable for genomics studies but not for metabolomics, proteomics, or transcriptomics. As recently reviewed by Pinu et al. (2019), the sampling procedure is one of the most important steps a researcher should carefully plan.

2.2 | Metadata collection

Another important aspect that should be considered is to record the metadata associated with samples (Ryan, Schloter, Berg, Kinkel et al., 2021) this reference is the one published in Environmental microbiology. This is essential when an -omics protocol is applied, since metadata are not often standardized between biological systems or methods and are often mis-annotated or ambiguous, generating an overly complex environment for sample reanalysis in the event of novel methods and bioinformatic techniques (Kasmanas et al., 2021). Several metadata standards are already available from the NCBI and the Genomes OnLine Database (GOLD), while consortia such as the International Human Microbiome Standards promote standard operating procedures in human microbiome research. Several consortia have already proposed "gold standards" that should be employed. An example is given from the area of agricultural microbiome research (Dundore-Arias et al., 2020).

A lack of complete or comprehensive metadata attributable to an experiment or even the absence of entire experimental datasets (including nucleic acid and protein sequence, metabolites, spectra and images) from public repositories also makes it extremely difficult for the scientific community to reuse valuable data for meta-analyses or comparative studies ("Overcoming hurdles …," 2017). Some onus must also be placed upon the scientific journals, such as obliging authors to publish both their datasets and associated metadata in accordance with predefined standards, depending on the field. This effort will encourage researchers to preserve their data in repositories and by enabling its reuse, increase its value to microbiome research (Hu et al., 2022; Ryan, Schloter, Berg, Kinkel et al., 2021).

Throughout the workshop discussions, it emerged that measures to record and make available metadata, including detailed information of samples, sample collection and storage, microbiome detection, and analysis, are strongly needed. From the workshop brainstorming emerged the need to encourage scientists to include positive controls (e.g., bacterial/fungi/virus mock community) and negative controls (e.g., blank sample) and to apply one of the already available detailed standard procedures (De Souza et al., 2020; Dundore-Arias et al., 2020; Molina et al., 2021; Santiago-Rodriguez & Hollister, 2021; Vangay et al., 2021; Yilmaz et al., 2011) depending on the ecosystem investigated.

2.3 | Hypothesis/biological question and sample size

Microbiome research is enjoying immense growth, but study designs are not always properly implemented to test hypotheses. Although researchers usually subscribe to the hypothesis-driven approach, the broad nature of -omics tools often leads to a non-hypothesis-driven approach being applied, with multi-omics studies often being treated as "hypothesis-free" as a result (Read & Sharma, 2021). However, the technique, sample type/size, and standard procedure must be considered carefully to obtain valid results. Sometimes researchers are unable to find a "suitable story" in their data because they do not properly estimate sample sizes ahead of time. The number of samples or sampling points must be chosen in order to reduce the probability of agreeing that the groups are different when they are not (Type I Error) and to reduce the probability of deciding that the groups are not different when they are (Type II Error) (La Rosa et al., 2012). In addition, the number of samples should be determined based on the available sequencing depth, rather than on the cost per sample. Tripathi and colleagues (Tripathi et al., 2018) clearly showed that even with small sample coverage, but with a high number of samples, it is possible to observe significant differences if compared with an experiment where fewer samples and higher coverage are used. However, it should also be considered that different sample sizes are required depending on the -omics technologies (Pinu et al., 2019).

Not only does the sample size affect the final results but also the power analysis remains underutilized in -omics studies and should be performed as a first step in order to determine what sorts of outcomes a study can feasibly generate (Ferdous et al., 2022). By way of definition, it informs the researcher on how large a study must be conducted in order to have sufficient statistical power for detecting differences at a given significance (e.g., $p < .05$)

between two groups of samples, for example, case versus controls (Mattiello et al., 2016). While many microbiome studies are underpowered (unable to detect significant effects even when present), combining data from several studies may find correlations or other associations that cannot be revealed by individual studies alone (Hu et al., 2022). This approach is commonly used in human microbiome studies but could also be adopted in other -omics fields for the detection of meaningful differences (Impactt, 2022).

The workshop participants also discussed about the needs of a strong biological knowledge before planning an -omics study. The first step is a good knowledge of the field where the research is applied. Further prerequisites are a solid understanding of the system under investigation, achieving the correct experimental design, and the consideration of existing data as a useful resource—for example, the identification of specific taxa or genes which boost the production of a metabolite or protein of interest (Daliri et al., 2021): this can be done by, for example, analyzing data from public repositories. Adequate information on the source (whether microbial or host) of metabolites/genes/species and knowledge of the biological sample are required before planning an experiment. Due to the limitations of automatic annotation pipelines, a relatively high number of errors occur during the structural annotation of genes coding for proteins (Armengaud et al., 2013). Frequent practice is to discard unannotated microbes or unannotated molecules, focusing on the subset of microbes or molecules that can be matched to an existing database. Alternate strategies do not lump unknowns, hypotheticals, and "others" together but treat them as yet-uncharacterized entities that can be identified by a single sequence, enabling future association of function with pure culture and genomic data from more and more microorganisms. Whenever possible, machine learning protocols should be applied to predicting the function of such uncharacterized sequences, with the stated goal of increasing understanding and enlarging the existing databases. Moreover, applying an extensive and systematic collection of microbe maps and molecules across different biological environments can improve our understanding of microbiome structure and function. This same approach also improves existing databases by reducing the proportion of unknown molecules/proteins/genomes. In vitro studies (protein expression at single cell or single-cell sequencing) (Pineda et al., 2020) and microbe interaction studies can help decipher the unknown function/potential of the microbes and may be a valuable tool to enlarge databases. The use of cell culture from microbial communities remains a challenge due to the interactions that often become too complex when communities grow beyond two or three members. However, the application of model-based identification of key microbial members that can co-grow to produce a specific outcome would be a valuable contribution to databases (Kessell et al., 2020). It should be mentioned that several repositories that are currently available contain hundreds of thousands of samples (just to cite few: QIITA https://qiita.ucsd.edu/, MGnify https://www.ebi.ac.uk/metagenomics/, FoodMicrobionet http://www.foodmicrobionet.org/), integrating the deposit/retrieval of sequence data with analysis tools for both amplicon-targeted or shotgun studies (Gonzalez et al., 2018; Mitchell et al., 2018; Parente et al., 2016). The continuous development of bioinformatic technologies enables the re-analysis of data and additional information compared to a previous analysis.

3 | OPTIMAL COMBINATIONS OF -OMICS THAT SHOULD BE USED AS A STANDARD IN DIFFERENT ECOSYSTEMS

To validate a hypothesis, at least two -omics approaches should be used simultaneously to obtain more reliable and valuable insights (Zapalska-Sozoniuk et al., 2019). Due to the greater popularity and accessibility of different -omics techniques, it is increasingly common to find descriptive papers based on a single omics platform. The choice of the technique used is mostly dependent on the samples that should be studied. The question of what combination of -omics tools should be used in each field was extensively discussed throughout the workshop, considering the scientific background of the participants as well as the future perspectives of microbiome-based studies. During the workshop, an online survey with multiple choice and open-end questions was conducted to establish a general overview of the importance that the participants of the workshop placed on -omics tools in microbiome studies. The results are summarized in Figure 2.

It is obvious that the choice of the -omics platform must consider the hypothesis and the experimental question. The survey highlighted that most of the participants prioritize transcriptomics, lipidomics, and genomics as the top three methods that should be used in a microbiome study. Metabolomics, proteomics, interaction secretome, and culturomics followed in the prioritization list, while at the end of the list, metataxonomics (also known as metabarcoding) was placed. As a further confirmation of this trend, a search on the Web of Science in the last 10 years yielded 217,798 articles and reviews on microbiome studies using transcriptomic (using the words transcriptomic*, meta-transcriptomic*, RNA-seq, transcriptome*), 9283 articles on lipidomics (using lipidomic*, meta-lipidomics*, lipidome*), and 379,929 articles on genomics (using genomic*, meta-genomic*, WGS, shotgun, DNA-seq). Based on the survey, Figure 3 shows the schematic overview of major combinations of -omics a researcher

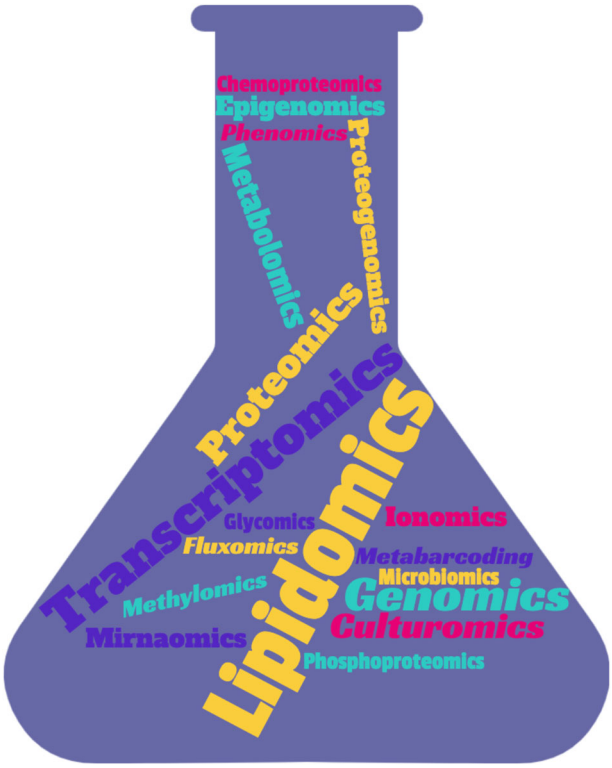

FIGURE 2 Preferred -omics tools in microbiome studies. Word cloud summarizes the preferences of 70 participants of an online survey during the workshop to have a general overview of the prioritization in -omics tools in microbiome studies applied in different ecosystems

should consider when planning an experiment. In detail, we observed that scientists investigating soil microbiome prioritized metagenomics plus culturomics, followed by metataxonomics plus metabolomics and then transcriptomics plus metabolomics. Plant microbiome studies utilized mostly metagenomics plus culturomics, followed by transcriptomics plus genomics. Food scientists prioritized metagenomics and metabolomics, then metabolomics and transcriptomics; and scientists in the field of gut microbiome prioritized transcriptomics (preferably annotated to function) and culturomics, followed by metagenomics and metabolomics. In the following sections, we summarize the multi-omics approach currently used in different segments of the food system (soil, plant, food, human, and animal), highlighting the main pitfalls and concrete actions that should be applied during experimental design to allow more meaningful research results.

3.1 | Overview of -omics integration in soil science

Soil is one of the most complex ecosystems and soil microbiomes have a crucial role not only in plant health and productivity (Trivedi et al., 2021), but also in animal and human health when considering the interconnectivity of the microbiomes throughout the ecosystems (Blum et al., 2019; D'Hondt et al., 2021; Hirt, 2020). Soil microbiomes have been extensively studied, and a wealth of data has been gathered through coordinated efforts such as The Earth Microbiome Project (Thompson et al., 2017) (https://earthmicrobiome.org/) and numerous research projects (Mishra et al., 2022). However, most of these efforts were DNA-based and collected metataxonomics and metagenomic information only. Nevertheless, these efforts also resulted in the development of the first set of standards and recommendations for soil microbiome analysis (https://earthmicrobiome.org/protocols-and-standards/) (Nannipieri et al., 2019) that could be extended to other -omics technologies to enable comparability and integrability of datasets from different studies. Notably, a range of different -omics technologies, including transcriptomics, proteomics, and metabolomics, were applied to study soil microbiomes (Biswas & Sarkar, 2018) in the past.

However, many studies applied one technology only, resulting in limited insight into the complexity of soil microbiomes. The need to go beyond taxonomic and one-dimensional functional studies was elaborated in several excellent reviews (Jansson & Hofmockel, 2018; Mishra et al., 2022; Nannipieri et al., 2019) and will thus not be discussed further here. Rather, we will describe some recent success stories that showcase how the integration of multiple-omics approaches contributed to a better understanding of the soil microbiomes and their function. A combination of metagenomics, metatranscriptomics, and metaproteomics was used to elucidate the phylogenetic composition, functional potential, and activity of the microbial communities in three soils representing different states of thaw (Hultman et al., 2015). The data revealed that some processes, that is, nitrogen, sulfur, and methane cycling, could be found in all soil types, whereas overall activity differed greatly between the three soil states (with permafrost exhibiting the lowest and thermokarst bog the highest functional potential). Furthermore, a novel survival strategy, based on the dissimilatory Fe(III) reduction, was proposed for potentially active microbes in permafrost. Metataxonomics and metaproteomics were used to decipher the key microbial players that are potentially involved in composts suppressive against soil-borne pathogen *Phytophthora nicotianae* (Ros et al., 2018). Indicators of suppression were found both at the taxonomic (Proteobacteria) and functional levels (proteins associated with carbohydrate processes, cell wall structure, and inorganic ion transport and metabolism). Yao and co-workers (Yao et al., 2018) investigated how soil microbial communities cope with growth-limiting phosphorus (P) deficiency

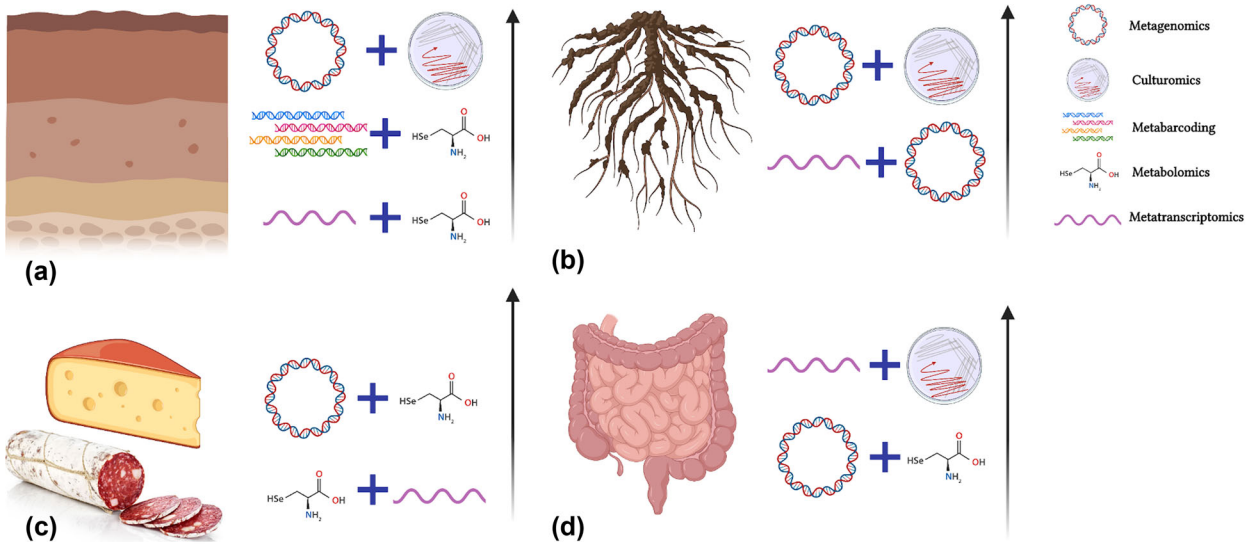

FIGURE 3 Overview of different biological samples/environment along with the major integrated -omics approaches that should be considered when planning an experiment in the food system: (a) soil; (b) plant; (c) food; (d) human. The arrow in each panel represents the prioritization of the -omics tools applied in each system

using a metagenomic and metaproteomic approach. The large-scale comparison of microbial communities in P-deficient and P-rich soils in a 17-year fertilization experiment in a tropical forest revealed the adaptive response of genes and proteins in soil microbial communities in response to shifting nutrient constraints. In P-deficient soils, a significant (greater than fourfold) increase in the gene abundance of 3-phytase that catalyzes the release of phosphate from phytate, the most recalcitrant phosphorus-containing compound in soil organic matter, was observed. Additionally, genes and proteins for the degradation of phosphorus-containing nucleic acids and phospholipids, as well as the decomposition of labile carbon and nitrogen, were also enhanced in the P-deficient soils. In the P-rich soils, increased abundances of genes involved in the degradation of recalcitrant aromatic compounds, the transformation of nitrogenous compounds, and assimilation of sulfur were observed.

Another comprehensive study of the agroecosystem under different management practices, using multi-omics analysis (Ichihashi et al., 2020), revealed complex interactions between soil metabolic, mineral, and microbial components. Within the integrated multi-omics data network, one node correlated with plant productivity (as measured through shoot dry weight), and within this node, soil organic nitrogen and thermophilic rhizosphere bacteria (including *Paenibacillaceae* and *Thermaceae*) were among the key components. These results indicated that the integrated multi-omics approach has predictive power to detect multilevel interactions between plants, microbes, and soils, and to identify key components in the agricultural ecosystem.

Applying a combination of metataxonomics, metagenomics, metatranscriptomics, and phenomics for the analysis of the soil samples from a long-term liming experiment improved the current understanding of soil denitrifier communities and how pH affects their activity, which organisms are involved, and their control and accumulation of denitrification intermediates (Frostegård et al., 2021). The data showed that the un-limed, low pH soil had severely delayed nitrous oxide (N_2O) reduction despite early transcription of the *nosZ* gene, encoding N_2O reductase, and the presence of the accessory genes from the nos cluster involved in the NosZ maturation. High *nir* transcript abundances (encoding nitrite (NO_2-) reductase) in un-limed, low pH suggested that low NO_2- concentration in acidic soils is the result of the biological activity and not, as commonly ascribed, of abiotic degradation. Analogously, high expression of the *nar* gene (encoding nitrate reductase) was correlated to the accumulation of NO_2- in the limed, pH neutral soil. Interestingly, the -omics results revealed dominance of *nirK* over *nirS* in both soils while qPCR showed the opposite, which led the authors to hypothesize that standard primer pairs only partially capture the *nirK* pool.

Multi-omics approaches also contributed to novel insights and developments in the bioremediation research field. In an early study, Mason and colleagues (Mason et al., 2012) investigated the functional role of the *Oceanospirillales* and other active members of the microbial community using metagenomics and metatranscriptomics, as well as single-cell genomics. The results showed that genes for motility, chemotaxis, and aliphatic hydrocarbon degradation were significantly enriched and expressed

in the hydrocarbon-contaminated plume samples compared with uncontaminated seawater. Contrarily, genes coding for degradation of more recalcitrant compounds, such as benzene, toluene, ethylbenzene, total xylenes, and polycyclic aromatic hydrocarbons, were identified in the metagenomes but expressed at low levels or not at all. Isolation and sequencing of two *Oceanospirillales* single cells enabled elucidation of the near-complete pathway for cyclohexane oxidation. Microorganisms responsible for aerobic biodegradation of biphenyl were investigated using biphenyl-degrading enrichment cultures and applying a combination of metataxonomics, metagenomics, stable isotope probing, and metaproteomics (Chen et al., 2021). Using ^{13}C-labeled biphenyl and tracing the flow of pollutant-derived carbon throughout the system resulted in the detection of the uncultured Alphaproteobacteria clade UBA11222, containing a distinctive biphenyl dioxygenase gene widely retrieved from contaminated environments. Furthermore, biphenyl oxidation potential was linked for the first time to *Azoarcus* and *Rugosibacter* genera.

Similarly, the application of combined metagenomics, metatranscriptomics, and metaproteomics enabled the identification of the microorganisms actively methylating arsenic in anoxic soil-derived microbial cultures (Viacava et al., 2022). Furthermore, based on the analysis of the metagenome-assembled genomes of microorganisms expressing arsenite S-adenosylmethionine methyltransferase (ArsM), a targeted cultivation strategy was developed, resulting in the isolation of the *Paraclostridium* sp. strain EML, which was confirmed to actively methylate arsenic under anaerobic conditions. These examples showcase the potential of multi-omics approaches to improve our understanding of microbial systems and their functioning, but also to support the development of novel beneficial management solutions and applications.

3.2 | Overview of -omics integration in plant science

It is well known that the microbiome directly affects plant health and survival (Compant et al., 2019; Trivedi et al., 2020). Through chemical exchange and supply of carbon and nutrients, plants selectively influence the composition and activity of their associated microbiota, which, in return, plays vital functions, including nutrient uptake, disease suppression, and abiotic stress tolerance (Compant et al., 2019; Lemanceau et al., 2017; Vandenkoornhuyse et al., 2015). Even though a large number of studies on plant-microbe interaction are solely based on amplicon sequencing, there is an increased number of cases that successfully applied multi-omics approaches to create a more complete view of the plant microbiome.

DNA-, RNA-seq, and metabolomics have successfully connected alterations in plant exudates caused by stresses to the enrichment of specific microbial groups. An increasing number of studies have shown that under a stressful abiotic condition, plants alter their exudation profile in a "cry for help" response to recruit a stress-relieving microbiome (Rizaludin et al., 2021; Trivedi et al., 2022; Yi et al., 2011). Water-limiting conditions, for instance, have been shown to shift many plant secondary metabolites and exudates (such as carbohydrates, mucilage, and osmolytes) and increase abundance and activity of certain monodermic bacteria and fungi associated with drought tolerance in plants (López-Ráez, 2016; Pang et al., 2021; Santos-Medellín et al., 2017; Xu et al., 2018; Xu & Coleman-Derr, 2019). Conversely, by using a combination of -omic approaches, it has been shown that the "cry for help" phenomenon also applies to plants under biotic stress, where root exudate metabolites shape a beneficial microbiota that protects plant against pathogens (Gómez Expósito et al., 2017; Hu et al., 2018; Huang et al., 2021).

As microbiome research suggests that plant-associated microbiota influences several host traits related to growth and health, studies have progressed beyond the microbiota to also include investigation of plant genetics and physiology. For instance, changes in microbial profile has been linked to transcriptional changes in plant host, with direct implications in host development, defense, and stress tolerance (Chaparro et al., 2014; Finkel et al., 2019; Hu et al., 2018; Teixeira et al., 2021). The integration of phenomics approaches has also allowed a deeper understanding of key alterations in plant physiology caused by association with microbes, such as water usage and nutrient status (Armanhi et al., 2021; Chai et al., 2021). Conversely, a number of GWAS studies has explored the microbiome as a quantitative plant trait and connected the recruitment of specific microbial taxa to plant genes related to different traits, such as carbon metabolism and plant defense (Deng et al., 2021; Horton et al., 2014; Wallace et al., 2018).

3.3 | Overview of -omics integration in food science

The food microbiome confers particular organoleptic characteristics to the final products (Bertuzzi et al., 2018; Ferrocino et al., 2018; Filippis et al., 2016), is responsible for spoilage and safety issues (Chaillou et al., 2015; Hultman et al., 2020; McHugh et al., 2018), and can be transferred from foods to humans (Milani et al., 2019; Pasolli et al., 2020).

From 2011 when -omics started to be used routinely in foods studies (De Filippis et al., 2018), an impressive number of papers reported the application of these technologies to allow thorough characterization of the microbiome in agro-food systems (Olmo et al., 2022). Amplicon sequencing remains the main approach for such studies (Parente et al., 2020) and is used mainly for observational studies with an ecological purpose. Several platforms offer the opportunity to predict with a certain accuracy the microbiota's composition from amplicon data (Caicedo et al., 2020); however, this technique does not detect important changes at species level, is blind to organisms not targeted by the specific DNA probes employed, and further is susceptible to PCR errors (Ferrocino et al., 2022).

To create both form and function, all biological systems rely on the interplay of information (genomics), function (proteomics), and environment (metabolomics) (Santiago-Rodriguez & Hollister, 2021). The optimal set of -omics tools that can be applied is to use nucleic acid sequencing (both DNA and RNA) combined with metabolomics or metaproteomics. Direct nucleic acid sequencing (shotgun metagenomics/metatranscriptomics) is often used in food-based studies with the aim of providing new insights into the components and functions of food ecosystems. In food-based studies, DNA sequencing is frequently combined with metabolomics due to the stability of DNA, the relatively affordable cost of DNA sequencing, and the availability of metabolomics platforms in most research centers. Several examples are already available that successfully integrated such an approach.

DNA-seq with metabolomics was successfully used to connect metagenomic clusters with the modification of color, variation of pH, and flavor development during cheese ripening (Bertuzzi et al., 2018); to detect genes and their correlation with flavor development in soy sauce (Sulaiman et al., 2014), fermented meat (Ferrocino et al., 2018; Franciosa et al., 2021), fermented cocoa (Mota-Gutierrez et al., 2021), fermented fish (Zhao & Eun, 2020), Daqu, Baijiu, and Xiaoqu jiu Chinese liquors (Huang et al., 2020; Yang et al., 2021; Zhao et al., 2021); or during water kefir fermentation (Verce et al., 2019).

It should be pointed out that DNA can originate also from dead cells and all the aforementioned studies showed a global view of the food microbiome, but gene expression dynamics cannot be assessed by the use of DNA as a target molecule. However the detection of one metabolite can, from one side, overcome this limitation even if it is very challenging to identify which specific species or gene produced it. Tests of association or correlation are often used to show the relationship between two or more microbes, gene functions, proteins, and son, as well as how they are impacted by the experimental design. DNA as a target molecule is easier to study compared to RNA, and scientific literature in food-omics is mostly oriented toward DNA-based approaches. The added value of DNA as the target molecule is the simultaneous detection of bacteria, fungi and other eukarya, some viruses (Beghini et al., 2021; Manni et al., 2021), genetic elements (ARGs, bacteriocins, etc.) (Raymond et al., 2019), as well as the ability to resolve differences at a taxonomic strain-level (Franciosa et al., 2021; Walsh et al., 2018).

However, we must consider the value that RNA gives to a study. Working with RNA implies that a researcher must take into account several factors like RNA instability, cost, and complexity of reverse transcription (Cottier et al., 2018). RNA-based studies in food systems ecology and function need to be coupled with other techniques such as metataxonomics, metaproteomics, or metabolomics. Several examples of data integration are available in the literature in this light.

An important example of a multi-omics approach based on transcriptomics applied in foods is the soy sauce fermentation (*ganjang*). The simultaneous application of DNA and RNA-seq coupled with metabolomics is helping to reconstruct the metabolic networks of *ganjang*. The study showed how the minor aerobic or facultative halophilic bacterial populations (*Chromohalobacter*, *Halomonas*, and *Marinobacter*) as well as minor yeasts (*Debaryomyces* and *Wickerhamomyces*) play an important role during fermentations influencing quality and taste and of ganjang (Chun et al., 2021). Using a combination of metatranscriptome and proteome data (in order to verify the transfer of information from nucleic acids to proteins), an upregulation of specific microbial genes was shown to boost the total polyphenols of Chinese yam (Guo et al., 2021). By coupling RNA-seq with metabolomics, it was also possible to see that perturbations during food processing modify the function of the microbiome. For example, process temperature during cheese ripening (De Filippis et al., 2016), fruit ripening (Li et al., 2021; Xu et al., 2019), plant-based fermentation (Kim et al., 2021), and vinegar production (Wu et al., 2021) regulates the metabolic activity and modifies the gene expression of the microbiome with important changes in volatilome profile of the final products. Apart from observational studies showing how process parameters can modify the microbiome with important changes in the final sensorial characteristic of the products, the integration of RNA-seq with metabolomics offers the possibility to study the mechanism of bacteria acid-and-ethanol tolerance directly connected with the quality characteristics and ammonia flavor formation in fermented fish (Zhao & Eun, 2020), the identification of microbes responsible for producing biogenic amines during soy sauce fermentation (Kim et al., 2021), or bacterial metabolism under high salinity-induced osmotic stress (Chun et al., 2019). Additionally, different

-omics tools were successfully integrated to identify the mechanisms behind the interconversion between hydroxytyrosol and oleuropein in order to define the optimal harvest period of olives for oil production (Rao et al., 2021).

Not only bacterial function can be monitored, but also fungi, especially in cheese (Dugat-Bony et al., 2015; Lessard et al., 2014). Integration of different -omics techniques showed that in cheese maturation, fungi promote the growth of motile over non-motile bacteria (Zhang et al., 2018). In addition, metagenomics and metatranscriptomics are helping identify and discover new gene groups that can cause antibiotic resistance (Nghiem et al., 2019).

These examples explore various adaptation strategies and biological phenomena of the microbiome and often report the use of unsupervised techniques like network or association/correlation analysis in order to identify quantitative relationships between -omic features (gene/taxa abundance and metabolites development) and samples. However, the quantitative relationships identified using correlation-based approaches (Pearson or Sperman) may not reflect the biological significance, nor do they specifically account for complex interactions (Santiago-Rodriguez & Hollister, 2021). Observed changes in metabolic features as a result of the production process, or due to external factors, will correspond to physicochemical changes in the substrate of a sample. A validation step which recapitulates the observed or imputed results (as in Koch's postulates) is often missing, since culture-based evaluations are absent in most of the examined studies. Only a few studies use this "complete" approach to validate the data obtained. It was recently showed that cheese discoloration defect was caused by an uncommon microbe (*Thermus thermophilus*) through a NGS approach. As a second step, the isolation of this causative bacterium by culture-based approaches allowed the cheese defect to be reproduced as a confirmation of the association retrieved from NGS data (Quigley et al., 2016). This exploits a unique advantage of food process technology, whereby it is relatively simple, in comparison with other larger biological systems, to isolate the main microbial component in order to verify the hypothesis.

3.4 | Overview of -omics integration in human-based study

The human microbiome is a complex ecosystem of bacteria, archaea, viruses (including bacteriophages), fungi, and, in some instances, a variety of microeukaryotes that vary between environments and populations and operate differently across the various ecological niches within the body (Pasolli et al., 2019). Gut microbial communities play an important role in synthesizing essential vitamins, out-competing pathogenic microbes, modulating immunity, impacting mucosal permeability, and regulating metabolic processes (Debnath et al., 2021; Wang et al., 2019). Indeed, alterations in the composition and function of the gut microbiota have been associated with several non-communicable diseases ranging from inflammatory to metabolic disorders and respiratory conditions (Rooks & Garrett, 2016; Sultan et al., 2021).

Multiomic studies have already provided a range of striking insights into microbiome–host interrelationships, many of which have been revealed through the contexts of diet. For example, the integration of metabolomics (SCFA, cytokines, bile acids, hormones), metagenomics, and amino-acid tagging has demonstrated the effect of fermentable fibers in improving satiety, through an increase in propionate-producing microbes over those who produce butyrate (Deehan et al., 2022). Another microbial metabolite, lithocholic acid, is one of a number of secondary bile acids produced by bacteria that act as powerful stimulators of thermogenesis through ligand binding (FXR, TGR5), leading to consumption of adipose material and subsequent weight loss (Pathak et al., 2018). Antigen cross-reactivity between the heat-shock protein, ClpB, and the feeding and emotional regulatory hormone, aMSH, induces fasting behavior in the host which can be recapitulated through consumption of ClpB-expressing *Escherichia coli* in rodents—as further validation, this cross reactivity can also be observed in humans who exhibit dysregulated feeding processes (Tennoune et al., 2014). In contrast with these desirable interactions, many mechanisms are characterized due to their roles in disease and, in particular, in metabolic dysfunction.

For instance, dietary studies have shown trimethylamine (TMA), a microbial byproduct of high protein-consumption, to accelerate a range of heart disease factors (Zhu, 2016), while direct chemical inhibition of microbial TMA synthesis reduces those same factors (Wang et al., 2015). Sufferers of type-2 diabetes have been shown to present higher microbial productivity and higher circulating concentrations of imidazole propionate, a microbial metabolite which reduces insulin sensitivity in the host by inhibiting the insulin signaling cascade post ligand-binding (Koh et al., 2016).

Nonalcoholic fatty liver disease (NAFLD) is associated with enrichment in microbial producers of hepatotoxic materials, including 2-butanone and 4-methyl-2-pentanone (Del Chierico et al., 2017). Ethanol, which can lead directly to inflammation and liver damage, is produced by *Klebsiella pneumoniae* and other proteobacteria, which have also been connected to the exacerbation of nonalcoholic steatosis and related conditions through the production of phenylacetate (Hoyles et al., 2018; Yuan et al., 2019). Branched-chain amino acids (L, I, V), which

are known agents of oxidative stress in cardiac disease and markers of T2D, are also generated by the gut microbiota (Wang et al., 2011). It seems likely that further research will uncover a panoply of roles in physiological function for the gut microbiome.

Although much is known about microbiome–host interrelationships through carefully conducted mechanistic studies, a comprehensive toolkit is required to further advance our understanding of the extensive range of microorganisms and their molecules in the gut. Fortunately, functional omics-based approaches have become more accessible and are used with increasing frequency, with the outputs from such analyses highlighting their potential to identify functional features of the microbiome related to health and disease (Han et al., 2021; Zheng et al., 2021).

Through the application of diverse -omics techniques such as metagenomic and metatranscriptomic profiling, changes in microbiome diversity, dysbiosis, and gene expression have been implicated in several different metabolic and inflammatory disorders, in particular diets, or in treatments with antibiotics or other drugs, thereby unveiling a more dynamic picture of the microbiome (Lavelle & Sokol, 2018; Schirmer et al., 2018). While metatranscriptomic profiles offer valuable insight into the functional expression of genes within microbiomes, metaproteomic screening offers even greater potential to accurately reflect the actual phenotype expressed. However, as of yet, metaproteomic analysis is still unable to achieve the information content or sampling depth of sequencing-based technologies. Indeed, highly abundant and stably expressed conserved proteins typically dominate in metaproteomic datasets, resulting in artificially stable profiles being generated (Ferrer et al., 2013). Despite this challenge, it should be noted that current limitations can be addressed by refining protein preparation protocols (Xiong et al., 2015) and improved identification (Tanca et al., 2016).

Another -omic approach, metabolomics, involving the direct measurement of metabolic outcomes, is likely to offer greater sensitivity in resolving the functional microbiome. Indeed, quantitative methods, such as proton nuclear magnetic resonance (NMR) analyses, are able to capture some of the most important, high-abundance microbial metabolites, such as short-chain fatty acids. While advanced metabolomic methods allow the resolution of thousands of metabolic features from human microbiome samples, current limitations relate to the very large fraction of unknown metabolites that can be generated (more than 90% of measured features may be unknown) even when comparing metabolomic data to comprehensive databases, as well as the challenges of relating specific metabolic features to their microbial provenance. Advances in computational mass spectrometry, *de novo* reconstruction, and modeling of metabolic networks have the potential to address some of these limitations and help understand the functional microbiome. A final -omic approach of relevance is culturomics. Notably, in the past, when traditional culture methods were employed, it was estimated that only 10%−25% of the gut microbiota could be isolated and identified (Jandhyala et al., 2015). However, armed with a greater understanding of the growth requirements of many gut microbes, more recent studies have significantly increased the success with which the dominant human gut species can be cultured and thus studied (Forster et al., 2019; Liu et al., 2021).

End-points of these studies should involve the monitoring of health-related physiological markers as well as following, in detail, the induced changes in the microbiome over time using -omics measurements to understand the role of the microbiome in managing human health and disease.

3.5 | Overview of -omics integration in animal based study

Sequence-based studies of animal gut microbiome started later in development than plant, soil, and the human microbiome. However, the role and importance of gut-health attracted attention in farmed animals prior to the sequencing era. Recently, sequence-based studies of animal gut microbiome have been intensified and building up in momentum and complexity. Microbiome studies in animals have developed rapidly in methodology in the last few years, initially with focus on 16S rRNA gene amplicon-based studies, later combined with culturomics (Zehavi et al., 2018). In recent years, studies have broadened to focus also on compounds for possible immunomodulating/anti-inflammatory effects. The strongest effect so far achieved by combining probiotics, prebiotics, and anti-inflammatory compounds (e.g., via lactic acid bacteria fermentation of feed biomass components, which after fermentation have high prebiotic and anti-inflammatory effects) (Grela et al., 2019) was revealed when focusing on the microbiome interactive (digestive) secretome and was enabled by combining metagenomic and metaproteomics with a strong functional annotation methodology (Barrett & Lange, 2019). Such studies represent state-of-the-art attempts to integrate use of many types of data deriving from DNA, RNA, proteins, and metabolites.

In addition, prioritizing animal microbiome studies holds the potential to provide significant new insights into broader roles of the microbiome in human health. Studies have progressed beyond the animal, to include

the effects of physical surroundings—the indoor vertical and horizontal surfaces of the stable (Grela et al., 2019). The inclusion of microbiome sampling from slaughtered animals in study design provides more detailed information about the different roles and functions of the microbiome in different parts of the gut system. There are also substantial opportunities to apply multi-omic approaches in understanding the function and role of the microbiome in the animal rumen, an evolutionarily unique, highly-specialized, and efficiency-optimized habitat for breaking recalcitrant plant structures. Such efforts are made even more valuable through their direct relevance to animal nutrition, production of red meat, and the resultant greenhouse gas emissions.

Carbon emissions from animal husbandry represent a significant threat for climate change, methane being 23 times more potent a greenhouse gas than CO_2. Intensive studies of the highly unique digestive microorganisms of the cow rumen were carried out decades before the genome sequencing era (Borneman & Akin, 1994). Attention was specifically drawn to studies of the large group of specialized rumen bacteria and not the least studying the unique rumen fungi (early lineage, zoosporic, anaerobic *Chytridiomycota spp*.). Notably, conceptual understanding of many facets of the rumen activities at the molecular level was achieved already in the 1980s. Most stunning was the discovery and characterization of the cellulosome, a common denominator between rumen bacteria and rumen fungi, being composed of a dockerin structure with a portfolio of structurally integrated digestive enzymes (Haitjema et al., 2017). Such insightful results were built on a combination of several sets and types of data, for example, rumen culturomics, excelled for both fungi, bacteria, and archaea; and rumen protein (enzyme) and metabolite and emission studies (e.g., effect and role of protozoa on pH and for starch metabolism), all with integrated use of taxonomic data, distinguishing a wide spectrum of taxonomically well-described archaeal, bacterial, and fungal species. At the start of the microbiome and metagenome sequencing era, many cow rumen research studies were based on 16S amplicon sequencing only, aiming at elucidating the taxonomic composition but methodologically including only prokaryotes.

The basis for choosing this simpler approach, often including only bacteria in the rumen microbiome studies, was analysis of total DNA from the rumen, estimating the ratio of bacteria versus fungi to hold only a smaller fraction of fungi (<5%–8%). Notably, Elekwachi and co-workers developed a new rumen microbiome sampling method (through inclusion of the solid fraction of the rumen) and a new RNA preparation method (Elekwachi et al., 2017), by which it was demonstrated that fungi constitute a larger part of the microbiome. Such underestimation of the fungal role in the rumen microbiome metabolism can be explained by a widespread sampling bias: Making the microbiome studies from the rumen fluid inherently underrepresented the fungal DNA, as the rumen fungi are attached by rhizoids to the feed biomass, thus found primarily in the (small particle) solid fraction of the rumen content.

Another parameter of complexity impacting the studies of the rumen microbiome is the intensive genetic research for optimized cattle breeding, suggesting at least to some extent that host genetics are controlling the rumen microbiome. A pertinent question is how we can optimize multi-omics approaches to include also the impact of host genetics and microbiome–host interaction. The rumen microbiome represents an optimal model for developing truly multi-omic microbiome research studies. Sampling hosts can be selected to have a highly similar, almost identical genetic background, rumen microbiome sampling methods can be fully standardized, and it can be carried out with high reproducibility from fistulated cows.

Extraction and preparation of samples for sequencing and analysis can be both comprehensive and standardized for quality and reproducibility. DNA, RNA, proteins, and metabolites can be analyzed and functionally annotated, with high specificity and sensitivity, by using a spectrum of -omics technologies—(meta)genomics, (meta)transcriptomics, metaproteomics, and metabolomics. For all steps, the analysis should be supplemented by functional analysis of the microbiome and microbiome–host interaction secretome. Last but most importantly, meticulous measurements of CO_2, hydrogen, and methane emissions can be made. In such a standardized system, hypotheses can be tested reliably by changing only one dimension: either the genetic background of the host or the composition of the feed intake (Noel et al., 2019). The objective, significantly reducing the methane emission from milk and meat production, could be within reach by simply changing the feeding regime and choosing the genetic background where such changes can take place with neither yield nor animal welfare being damaged.

4 | DATA INTEGRATION: CURRENT APPROACHES AND REMAINING HURDLES

Complex cellular functions depend on the interplay between genes, transcripts, proteins, metabolites, and other biological molecules. Because of this complexity, if a biological system is to be understood with the greatest accuracy and detail, the methods used must try, as much as is possible, to carry out a simultaneous and integrated

analysis of these molecules. This type of multi-omic data integration offers a powerful way to view a cellular system but can be fraught with challenges (Jiang et al., 2019). However, new approaches are addressing these challenges and making multi-omic data integration easier (McClure et al., 2019). In this section, we address some of the approaches currently used to integrate multi-omic data as well as the hurdles associated with these methods and the steps taken to mitigate them.

4.1 | Network analysis of multi-omic data

One method of integrating multi-omic data is using a feature co-expression or co-abundance network approach. In this approach, multiple kinds of -omics data (e.g., proteomics and transcriptomics) can be integrated through identifying proteins and transcripts that shift their abundance in a coordinated manner across a range of experimental conditions or samples.

There are a number of mathematical methods for inferring such networks that have been applied including Pearson and Spearman correlation (Afshari et al., 2020) and mutual information methods such as context likelihood of relatedness (CLR) (Faith et al., 2007), overall random forest based methods such as GENIE3 (Huynh-Thu et al., 2010), and several others (Margolin et al., 2006). While most of these network inference tools have been applied to networks of a single -omics type (mainly amplicon analysis or transcriptomics), their use as multi-omic network generation tools has also been explored. In these networks, each feature (either a species, transcript, protein, metabolite, or lipid) represents a node in the network, and instances of high coordination of expression or abundance between two features represent edges (connections) within the network.

The key difference with single -omic versus multi-omic networks is that in multi-omic networks, edges can link features of the same type as well as features of two different types (e.g., a transcript linked to a metabolite). Some studies have set out specifically with the goal of comparing the strength of various network inference tools as it relates to multi-omic networks. One study compared the ability of 10 network inference methods and ranked them based on their ability to generate highly integrated multi-omic networks using transcriptomic and proteomic data (McClure et al., 2019). Integration was measured both by accuracy of the resulting network and by the ratio of edges connecting features of different types (protein/transcript) versus features of the same type (protein/protein or transcript/transcript). This study found that the random forest method, GENIE3, was by far the best method for inferring associations within multi-omic data. Another recent review compared a number of different methods and also found that GENIE3 was a promising candidate for multi-omic network integration (Hawe et al., 2019). It should also be noted that GENIE was a top contender in the DREAM challenge which sought to compare network inference tools regarding their accuracy in linking regulator-target pairs of *E. coli* (Marbach et al., 2012). Though GENIE3 has been found to be useful for multi-omic networks, other methods have also been used with success. These include PALM, a Bayesian network that is designed to work with multi-omic longitudinal time scale data (Ruiz-Perez et al., 2021), and TIGRESS, a regression approach that uses the behavior of some features to predict the behavior of others (Yan et al., 2017).

Further analysis of both of these methods and others is warranted due to the rather specific input data on which GENIE3 ranked so highly. Most of these data were from tightly controlled in vivo studies, and it is likely that in a natural microbiome setting (the gut, soil, marine, or other sites), natural variability may lead to other network inference methods being able to draw more edges between features, lowering the impact of GENIE3.

While networks can provide a great deal of data, the possibility exists for spurious links or edges to be inferred in a network, reducing the accuracy and use of the network. With any network analysis approach, some edges that do not reflect true biological associations but rather random correlations of data are inevitable. The key to proper network analysis is to: (1) identify to what degree such spurious edges exist in a network and (2) to confirm edges of interest as much as possible with additional data types or experiments. To gain a view of how abundant spurious edges are, accuracy of networks can be calculated by looking at the ratio of edges linking genes in the same functional category (likely true links) compared to edges linking genes in different functional categories (possible spurious links). This method was shown to work well in networks of transcriptomic or proteomic data (McClure et al., 2019), where the ratio of true links to spurious links dropped when networks were made with sparse data or liberal edge cutoffs (wherein spurious edges would be more prominent) and rose when networks were made with rich data and more conserved edge cutoffs. Using this metric, networks with a large number of spurious edges can at least be identified and set aside or modified before conclusions are made.

Confirming edges of interest is a second way that spurious edges can be dealt with. This can be done by either using multiple network methods on the same data set, if an edge of interest is found regardless of the specific mathematical method used to infer the network than it can be assumed that the edge represents a true biological link and not a random mathematical correlation. This can

also be done by incorporating other analyses. For example, in a network of genes linked to regulatory DNA-binding proteins, networks were initially used to link regulators and targets (McClure et al., 2016). These links were then partially confirmed by looking for DNA-binding sites for the identified regulator in the promoter region of target genes linked in the network. Genes that were found to be linked to a regulator in a network and were found to harbor the binding site of that regulatory in their 5′ untranslated region were strong candidates for being new targets of the regulator. While such pieces of data in isolation would not allow any strong conclusion to be drawn, integration of these different datasets provides significant insight.

4.2 | Overlaying multi-omic data onto existing metabolic pathways

Much of the network analysis above can be done with minimal knowledge of gene classification of the system under analysis. In fact, networks have even been used to expand knowledge of gene function in a process termed "guilt by association," though other studies have emphasized the limits of this approach (Gillis & Pavlidis, 2012). However, while not universally true, there does often exist some genomic annotation data for the system under analysis. In addition, many pathways and processes in microbial systems are universal, offering the possibility of using general KEGG pathway information to guide analysis of specific systems. Because of this, new methods have been emerging that combine existing knowledge with multi-omic data from the system to infer which processes are being activated under certain conditions.

One such approach is the Metabolite-Expression-Metabolic Network Integration for Pathway Identification and Selection (MEMPIS) program. This program collects transcriptomic and metabolic data from a system and integrates it in combination with annotated metabolic pathways to identify which processes are being expressed by a biological system such as a microbiome. Interestingly, the current applications of MEMPIS are not on well-annotated model systems but on complex soil microbiome samples showing that this approach can be applied to sites where little specific knowledge of genome annotation has yet been collected (McClure et al., 2020). Another tool, XCMS Online, also includes metabolomics data but is able to integrate proteomic data or genomic data as well (Forsberg et al., 2018). However, unlike MEMPIS, this tool has not yet been tested extensively in complex microbiome systems. Other methods that take as input multi-omic data and overlay it onto existing pathways using KEGG, KBase, or Reactome include GraphOmics and InterTADs (Tsagiopoulou et al., 2022). Other approaches have taken a broader view and instead of developing wholly new methods of multi-omic integration have instead developed pipelines that guide users in applying existing multi-omic integration tools. This is the case with STATegra that has been demonstrated to be successful in multi-omic analysis of human systems, though its application to microbial data has not yet been evaluated (Planell et al., 2021).

5 | CONCLUSIONS

In this paper, we have presented the outcomes of the workshop "Metagenomics, Metaproteomics, and Metabolomics: The need for data integration in microbiome research" organized to discuss the pros and cons of study designs, methodologies, and statistical/integration tools to advance microbiome applications in the context of the food system. In doing this, we have taken a systems thinking approach, and to this end, a number of domains associated with food (e.g., soil, plant, human, and animal sciences) were taken into consideration to address the complexity of the food system.

Despite the success over the last several years, challenges still remain with multi-omic data integration and their effective exploitation in the food system. First, technologies applied to study these factors can generate continuous, discrete, and categorical data which contain batch effects and are often noisy, sparse (i.e., contain many zeros), and high-dimensional (Tsagiopoulou et al., 2022). Apart from normalization, transformation, or scaling steps used before integration, another challenge is the use of reproducible pipelines and a better community effort for data analysis strategies based on the ecological niche. The common approach now used is based on network-based methods, multi-variate random forests, and Bayesian approaches used to find probabilistic causal relationships between variables, and to identify the most probabilistic network that is predictive of the observed data (Chong & Xia, 2017). An extensive, but not exhaustive, list of integration tools is accessible via GitHub (https://github.com/mikelove/awesome-multi-omics) along with commercial platforms that may be more user friendly. Most methods utilize randomizations to generate data without regard to biochemical structure; thus, biological significance is sometimes overestimated. Univariate correlations are relatively straightforward but lack context for interpretation in terms of biological plausibility and mechanistic insight (Chong et al., 2020). Various tools have been developed to integrate -omics datasets, but there are limited strategies to systematically extract mechanistic hypotheses from them. To give an example, the COSMOS platform (Dugourd et al., 2021) offers the opportunity to find potential mechanisms by linking deregulated protein

activities and metabolite concentrations with a known network by coupling signaling, transcription, and metabolites. The advantage of this tool is of interest, since apart from proposed mechanisms between pairs of molecules (metabolites and proteins) it can also take into account the interaction with other inputs (molecules or genes) that can be included in the model. Throughout the workshop discussion, several propositions emerged including: (i) the need for classical microbiology to produce knockouts to extract mechanistic hypotheses; (ii) advertising how important the mechanistic hypothesis is, since at present it is still often simply ignored; (iii) obtain more knowledge about the biological system as well as updating and integrating this information into popular databases; (iv) drive scientists to apply mechanistic, rather than descriptive studies; and (v) focusing on the combination of extensive prior knowledge and (new) computational methods to obtain new mechanistic insight. Only in this way will microbiome science efficiently contribute to the food system, allowing for solutions to be designed which can help to address the challenges our planet currently faces.

AUTHOR CONTRIBUTIONS

Ilario Ferrocino: Writing – original draft; Conceptualization; Methodology; Writing – review & editing. Kalliopi Rantsiou: Conceptualization; Writing – original draft; Validation; Writing – review & editing; Supervision. Ryan McClure: Writing – original draft; Writing – review & editing; Conceptualization. Tanja Kostic: Conceptualization; Writing – original draft; Writing – review & editing. Rafael Soares Correa de Souza: Conceptualization; Writing – original draft; Writing – review & editing. Lene Lange: Conceptualization; Writing – original draft; Writing – review & editing. Jamie Fitzgerald: Writing – original draft; Writing – review & editing. Aicha Kriaa: Writing – original draft; Writing – review & editing. Paul Cotter: Conceptualization; Writing – original draft; Writing – review & editing. Emmanuelle Maguin: Conceptualization; Writing – original draft; Writing – review & editing. Bettina Schelkle: Conceptualization; Writing – review & editing. Michael Schloter: Writing – original draft; Writing – review & editing; Conceptualization. Gabriele Berg: Conceptualization; Writing – original draft; Writing – review & editing. Angela Sessitsch: Conceptualization; Funding acquisition; Writing – original draft; Writing – review & editing. Luca Cocolin: Conceptualization; Funding acquisition; Writing – original draft; Writing – review & editing; Supervision.

ACKNOWLEDGMENTS

We would like to thank all attendees of the online workshop titled "Metagenomics, Metaproteomics, and Metabolomics: The need for data integration in microbiome research" held on April 13, 2021, for the interesting discussions.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ORCID

Luca Cocolin https://orcid.org/0000-0003-4799-7074

REFERENCES

Afshari, R., Pillidge, C. J., Read, E., Rochfort, S., Dias, D. A., Osborn, A. M., & Gill, H. (2020). New insights into cheddar cheese microbiota-metabolome relationships revealed by integrative analysis of multi-omics data. *Scientific Reports*, *10*(1), 1–13. https://doi.org/10.1038/s41598-020-59617-9

Armanhi, J. S. L., de Souza, R. S. C., Biazotti, B. B., Yassitepe, J. E., de, C. T., & Arruda, P. (2021). Modulating drought stress response of maize by a synthetic bacterial community. *Frontiers in Microbiology*, *12*(10). https://doi.org/10.3389/fmicb.2021.747541

Armengaud, J., Marie Hartmann, E., & Bland, C. (2013). Proteogenomics for environmental microbiology. *Proteomics*, *13*(18–19), 2731–2742. https://doi.org/10.1002/pmic.201200576

Barrett, K., & Lange, L. (2019). Peptide-based functional annotation of carbohydrate-active enzymes by conserved unique peptide patterns (CUPP). *Biotechnology for Biofuels*, *12*(1), 1–21. https://doi.org/10.1186/s13068-019-1436-5

Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Thomas, A. M., Valles-Colomer, M., Weingart, G., Zhang, Y., Zolfo, M., Huttenhower, C., Franzosa, E. A., & Segata, N. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *ELife*, *10*, 1–42. https://doi.org/10.7554/eLife.65088

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C. C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G. H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J. A., Maguin, E., Mauchline, T., McClure, R., … Schloter, M. (2020). Microbiome definition re-visited: Old concepts and new challenges. *Microbiome*, *8*(1), 103. https://doi.org/10.1186/s40168-020-00875-0

Bertuzzi, A. S., Walsh, A. M., Sheehan, J. J., Cotter, P. D., Crispie, F., McSweeney, P. L. H., Kilcawley, K. N., & Rea, M. C. (2018). Omics-based insights into flavor development and microbial succession within surface-ripened cheese. *MSystems*, *3*(1), 1–15. https://doi.org/10.1128/msystems.00211-17

Biswas, R., & Sarkar, A. (2018). *'Omics' tools in soil microbiology: The state of the art*. In T. Adhya, B. Lal, B. Mohapatra, D. Paul, S. Das (Eds.), Advances in soil microbiology: Recent trends and future prospects (pp. 35–64). Springer. https://doi.org/10.1007/978-981-10-6178-3_3

Blum, W. E. H., Zechmeister-Boltenstern, S., & Keiblinger, K. M. (2019). Does soil contribute to the human gut microbiome? *Microorganisms*, *7*(9), 287. https://doi.org/10.3390/MICROORGANISMS7090287

Borneman, S., & Akin, D. E. (1994). The nature of anaerobic fungi and their polysaccharide degrading enzymes. *Mycoscience*, *35*(2), 199–211. https://doi.org/10.1007/BF02318501

Caicedo, H. H., Hashimoto, D. A., Caicedo, J. C., Pentland, A., & Pisano, G. P. (2020). PICRUSt2 for prediction of metagenome functions. *Nature Biotechnology*, *38*(6), 669–673. https://doi.org/10.1038/s41587-020-0550-z

Chai, Y. N., Ge, Y., Stoerger, V., & Schachtman, D. P. (2021). High-resolution phenotyping of sorghum genotypic and phenotypic responses to low nitrogen and synthetic microbial communities. *Plant, Cell & Environment*, *44*(5), 1611–1626. https://doi.org/10.1111/pce.14004

Chaillou, S., Chaulot-Talmon, A., Caekebeke, H., Cardinal, M., Christieans, S., Denis, C., Hélène Desmonts, M., Dousset, X., Feurer, C., Hamon, E., Joffraud, J.-J. J., La Carbona, S., Leroi, F., Leroy, S., Lorre, S., Macé, S., Pilet, M.-F. F., Prévost, H., Rivollier, M., … Champomier-Vergès, M.-C. C. (2015). Origin and ecological selection of core and food-specific bacterial communities associated with meat and seafood spoilage. *The ISME Journal*, *9*(5), 1105–1118. https://doi.org/10.1038/ismej.2014.202

Chaparro, J. M., Badri, D. V., & Vivanco, J. M. (2014). Rhizosphere microbiome assemblage is affected by plant development. *The ISME Journal*, *8*(4), 790–803. https://doi.org/10.1038/ismej.2013.196

Chen, S. C., Budhraja, R., Adrian, L., Calabrese, F., Stryhanyuk, H., Musat, N., Richnow, H. H., Duan, G. L., Zhu, Y. G., & Musat, F. (2021). Novel clades of soil biphenyl degraders revealed by integrating isotope probing, multi-omics, and single-cell analyses. *The ISME Journal*, *15*(12), 3508–3521. https://doi.org/10.1038/s41396-021-01022-9

Chong, J., Liu, P., Zhou, G., & Xia, J. (2020). Using microbiomeanalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nature Protocols*, *15*(3), 799–821. https://doi.org/10.1038/s41596-019-0264-1

Chong, J., & Xia, J. (2017). Computational approaches for integrative analysis of the metabolome and microbiome. *Metabolites*, *7*(4), 62. https://doi.org/10.3390/metabo7040062

Chun, B. H., Han, D. M., Kim, H. M., Park, D., Jeong, D. M., Kang, H. A., & Jeon, C. O. (2021). Metabolic features of ganjang (a Korean traditional soy sauce) fermentation revealed by genome-centered metatranscriptomics. *MSystems*, *6*(4). https://doi.org/10.1128/msystems.00441-21

Chun, B. H., Han, D. M., Kim, K. H., Jeong, S. E., Park, D., & Jeon, C. O. (2019). Genomic and metabolic features of *Tetragenococcus halophilus* as revealed by pan-genome and transcriptome analyses. *Food Microbiology*, *83*, 36–47. https://doi.org/10.1016/j.fm.2019.04.009

Compant, S., Samad, A., Faist, H., & Sessitsch, A. (2019). A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *Journal of Advanced Research*, *19*, 29–37. https://doi.org/10.1016/j.jare.2019.03.004

Cottier, F., Srinivasan, K. G., Yurieva, M., Liao, W., Poidinger, M., Zolezzi, F., & Pavelka, N. (2018). Advantages of meta-total RNA sequencing (MeTRS) over shotgun metagenomics and amplicon-based sequencing in the profiling of complex microbial communities. *Npj Biofilms and Microbiomes*, *4*(1). https://doi.org/10.1038/s41522-017-0046-x

Daliri, E. B. M., Ofosu, F. K., Chelliah, R., Lee, B. H., & Oh, D. H. (2021). Challenges and perspective in integrated multi-omics in gut microbiota studies. *Biomolecules*, *11*(2), 1–10. https://doi.org/10.3390/biom11020300

Debnath, N., Kumar, R., Kumar, A., Mehta, P. K., & Yadav, A. K. (2021). Gut-microbiota derived bioactive metabolites and their functions in host physiology. *Biotechnology and Genetic Engineering Reviews*, *37*(2), 105–153. https://doi.org/10.1080/02648725.2021.1989847

Deehan, E. C., Zhang, Z., Riva, A., Armet, A. M., Perez-Muñoz, M. E., Nguyen, N. K., Krysa, J. A., Seethaler, B., Zhao, Y. Y., Cole, J., Li, F., Hausmann, B., Spittler, A., Nazare, J. A., Delzenne, N. M., Curtis, J. M., Wismer, W. V., Proctor, S. D., Bakal, J. A., … Walter, J. (2022). Elucidating the role of the gut microbiota in the physiological effects of dietary fiber. *Microbiome*, *10*(1), 1–22. https://doi.org/10.1186/s40168-022-01248-5

De Filippis, F., Genovese, A., Ferranti, P., Gilbert, J. A., & Ercolini, D. (2016). Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. *Scientific Reports*, *6*, 1–12. https://doi.org/10.1038/srep21871

De Filippis, F., Parente, E., & Ercolini, D. (2018). Recent past, present, and future of the food microbiome. *Annual Review of Food Science and Technology*, *9*(25), 1–20.

Del Chierico, F., Nobili, V., Vernocchi, P., Russo, A., Stefanis, C. D. e., Gnani, D., Furlanello, C., Zandonà, A., Paci, P., Capuani, G., Dallapiccola, B., Miccheli, A., Alisi, A., & Putignani, L. (2017). Gut microbiota profiling of pediatric nonalcoholic fatty liver disease and obese patients unveiled by an integrated meta-omics-based approach. *Hepatology*, *65*(2), 451–464. https://doi.org/10.1002/hep.28572

Deng, S., Caddell, D. F., Xu, G., Dahlen, L., Washington, L., Yang, J., & Coleman-Derr, D. (2021). Genome wide association study reveals plant loci controlling heritability of the rhizosphere microbiome. *The ISME Journal*, *15*(11), 3181–3194. https://doi.org/10.1038/s41396-021-00993-z

De Souza, R. S. C., Armanhi, J. S. L., & Arruda, P. (2020). From microbiome to traits: Designing synthetic microbial communities for improved crop resiliency. *Frontiers in Plant Science*, *11*, 1–7. https://doi.org/10.3389/fpls.2020.01179

D'Hondt, K., Kostic, T., McDowell, R., Eudes, F., Singh, B. K., Sarkar, S., Markakis, M., Schelkle, B., Maguin, E., & Sessitsch, A. (2021). Microbiome innovations for a sustainable future. *Nature Microbiology*, *6*(2), 138–142. https://doi.org/10.1038/s41564-020-00857-w

Dugat-Bony, E., Straub, C., Teissandier, A., Onésime, D., Loux, V., Monnet, C., Irlinger, F., Landaud, S., Leclercq-Perlat, M.-N., Bento, P., Fraud, S., Gibrat, J.-F., Aubert, J., Fer, F., Guédon, E., Pons, N., Kennedy, S., Beckerich, J.-M., Swennen, D., & Bonnarme, P. (2015). Overview of a surface-ripened cheese community functioning by meta-omics analyses. *PLoS ONE*, *10*(4), e0124360. https://doi.org/10.1371/journal.pone.0124360

Dugourd, A., Kuppe, C., Sciacovelli, M., Gjerga, E., Gabor, A., Emdal, K. B., Vieira, V., Bekker-Jensen, D. B., Kranz, J., Bindels, E. M. J., Costa, A. S. H., Sousa, A., Beltrao, P., Rocha, M., Olsen, J. V., Frezza, C., Kramann, R., & Saez-Rodriguez, J. (2021). Causal integration of multi-omics data with prior knowledge to generate mechanistic hypotheses. *Molecular Systems Biology*, *17*(1), 1–17. https://doi.org/10.15252/msb.20209730

Dundore-Arias, J. P., Eloe-Fadrosh, E. A., Schriml, L. M., Beattie, G. A., Brennan, F. P., Busby, P. E., Calderon, R. B., Castle, S. C., Emerson, J. B., Everhart, S. E., Eversole, K., Frost, K. E., Herr, J. R., Huerta, A. I., Iyer-Pascuzzi, A. S., Kalil, A. K., Leach, J. E., Leonard, J., Maul, J. E., … Kinkel, L. L. (2020). Community-driven metadata standards for agricultural microbiome research. *Phytobiomes Journal*, *4*(2), 115–121. https://doi.org/10.1094/PBIOMES-09-19-0051-P

Elekwachi, C. O., Wang, Z., Wu, X., Rabee, A., & Forster, R. J. (2017). Total rRNA-seq analysis gives insight into bacterial, fungal, protozoal and archaeal communities in the rumen using an optimized RNA isolation method. *Frontiers in Microbiology*, *8*, 1–14. https://doi.org/10.3389/fmicb.2017.01814

Faith, J. J., Hayete, B., Thaden, J. T., Mogno, I., Wierzbowski, J., Cottarel, G., Kasif, S., Collins, J. J., & Gardner, T. S. (2007). Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLOS Biology*, *5*(1), e8. https://doi.org/10.1371/JOURNAL.PBIO.0050008

Ferdous, T., Jiang, L., Dinu, I., Groizeleau, J., Kozyrskyj, A. L., Greenwood, C. M. T., & Arrieta, M. C. (2022). The rise to power of the microbiome: Power and sample size calculation for microbiome studies. *Mucosal Immunology*, *15*(6), 1060–1070. https://doi.org/10.1038/s41385-022-00548-1

Ferrer, M., Ruiz, A., Lanza, F., Haange, S. B., Oberbach, A., Till, H., Bargiela, R., Campoy, C., Segura, M. T., Richter, M., von Bergen, M., Seifert, J., & Suarez, A. (2013). Microbiota from the distal guts of lean and obese adolescents exhibit partial functional redundancy besides clear differences in community structure. *Environmental Microbiology*, *15*(1), 211–226. https://doi.org/10.1111/J.1462-2920.2012.02845.X

Ferrocino, I., Bellio, A., Giordano, M., Macori, G., Romano, A., Rantsiou, K., Decastelli, L., & Cocolin, L. (2018). Shotgun metagenomics and volatilome profile of the microbiota of fermented sausages. *Applied and Environmental Microbiology*, *84*(3), 1–15. https://doi.org/10.1128/AEM.02120-17CE

Ferrocino, I., Rantsiou, K., & Cocolin, L. (2022). Investigating dairy microbiome: An opportunity to ensure quality, safety and typicity. *Current Opinion in Biotechnology*, *73*, 164–170. https://doi.org/10.1016/j.copbio.2021.08.009

Finkel, O. M., Salas-González, I., Castrillo, G., Spaepen, S., Law, T. F., Teixeira, P. J. P. L., Jones, C. D., & Dangl, J. L. (2019). The effects of soil phosphorus content on plant microbiota are driven by the plant phosphate starvation response. *PLoS Biology*, *17*(11), e3000534. https://doi.org/10.1371/journal.pbio.3000534

Forsberg, E. M., Huan, T., Rinehart, D., Benton, H. P., Warth, B., Hilmers, B., & Siuzdak, G. (2018). Data processing, multi-omic pathway mapping, and metabolite activity analysis using XCMS online. *Nature Protocols*, *13*(4), 633–651. https://doi.org/10.1038/nprot.2017.151

Forster, S. C., Kumar, N., Anonye, B. O., Almeida, A., Viciani, E., Stares, M. D., Dunn, M., Mkandawire, T. T., Zhu, A., Shao, Y., Pike, L. J., Louie, T., Browne, H. P., Mitchell, A. L., Neville, B. A., Finn, R. D., & Lawley, T. D. (2019). A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nature Biotechnology*, *37*(2), 186–192. https://doi.org/10.1038/S41587-018-0009-7

Franciosa, I., Ferrocino, I., Giordano, M., Mounier, J., Rantsiou, K., & Cocolin, L. (2021). Specific metagenomic asset drives the spontaneous fermentation of Italian sausages. *Food Research International*, *144*, 110379. https://doi.org/10.1016/j.foodres.2021.110379

Frostegård, Å., Vick, S. H. W., Lim, N. Y. N., Bakken, L. R., & Shapleigh, J. P. (2021). Linking meta-omics to the kinetics of denitrification intermediates reveals pH-dependent causes of N2O emissions and nitrite accumulation in soil. *The ISME Journal*, *16*(1), 26–37. https://doi.org/10.1038/s41396-021-01045-2

Gillis, J., & Pavlidis, P. (2012). "Guilt by association" is the exception rather than the rule in gene networks. *PLOS Computational Biology*, *8*(3), e1002444. https://doi.org/10.1371/JOURNAL.PCBI.1002444

Gómez Expósito, R., de Bruijn, I., Postma, J., & Raaijmakers, J. M. (2017). Current insights into the role of rhizosphere bacteria in disease suppressive soils. *Frontiers in Microbiology*, *8*(3), 610–621. https://doi.org/10.3389/fmicb.2017.02529

Gonzalez, A., Navas-Molina, J. A., Kosciolek, T., McDonald, D., Vázquez-Baeza, Y., Ackermann, G., DeReus, J., Janssen, S., Swafford, A. D., Orchanian, S. B., Sanders, J. G., Shorenstein, J., Holste, H., Petrus, S., Robbins-Pianka, A., Brislawn, C. J., Wang, M., Rideout, J. R., Bolyen, E., … Knight, R. (2018). Qiita: Rapid, web-enabled microbiome meta-analysis. *Nature Methods*, *15*(10), 796–798. https://doi.org/10.1038/s41592-018-0141-9

Grela, E. R., Czech, A., Kiesz, M., Wlazło, Ł., & Nowakowicz-Dębek, B. (2019). A fermented rapeseed meal additive: Effects on production performance, nutrient digestibility, colostrum immunoglobulin content and microbial flora in sows. *Animal Nutrition*, *5*(4), 373–379. https://doi.org/10.1016/j.aninu.2019.05.004

Guo, S., Wang, D., Ma, Y., Zhang, Y., & Zhao, X. (2021). Combination of RNA-Seq transcriptomics and iTRAQ proteomics reveal the mechanism involved in fresh-cut yam yellowing. *Scientific Reports*, *11*(1), 1–16. https://doi.org/10.1038/s41598-021-87423-4

Haitjema, C. H., Gilmore, S. P., Henske, J. K., Solomon, K. V., De Groot, R., Kuo, A., Mondo, S. J., Salamov, A. A., LaButti, K., Zhao, Z., Chiniquy, J., Barry, K., Brewer, H. M., Purvine, S. O., Wright, A. T., Hainaut, M., Boxma, B., Van Alen, T., Hackstein, J. H. P., … O'Malley, M. A. (2017). A parts list for fungal cellulosomes revealed by comparative genomics. *Nature Microbiology*, *2*(5), 1–8. https://doi.org/10.1038/nmicrobiol.2017.87

Han, S., Van Treuren, W., Fischer, C. R., Merrill, B. D., DeFelice, B. C., Sanchez, J. M., Higginbottom, S. K., Guthrie, L., Fall, L. A., Dodd, D., Fischbach, M. A., Sonnenburg, J. L., Treuren, W. V., Fischer, C. R., Merrill, B. D., DeFelice, B. C., Sanchez, J. M., Higginbottom, S. K., Guthrie, L., … Sonnenburg, J. L. (2021). A metabolomics pipeline for the mechanistic interrogation of the gut microbiome. *Nature*, *595*(7), 415–420. https://doi.org/10.1038/s41586-021-03707-9

Hawe, J. S., Theis, F. J., & Heinig, M. (2019). Inferring interaction networks from multi-omics data. *Frontiers in Genetics*, *10*(6), 1–13. https://doi.org/10.3389/fgene.2019.00535

Hirt, H. (2020). Healthy soils for healthy plants for healthy humans. *EMBO Reports*, *21*(8), e51069. https://doi.org/10.15252/EMBR.202051069

Horton, M. W., Bodenhausen, N., Beilsmith, K., Meng, D., Muegge, B. D., Subramanian, S., Vetter, M. M., Vilhjálmsson, B. J., Nordborg, M., Gordon, J. I., & Bergelson, J. (2014). Genome-wide association study of arabidopsis thaliana leaf microbial community. *Nature Communications*, *5*(1), 5320. https://doi.org/10.1038/ncomms6320

Hoyles, L., Fernández-Real, J. M., Federici, M., Serino, M., Abbott, J., Charpentier, J., Heymes, C., Luque, J. L., Anthony, E., Barton, R. H., Chilloux, J., Myridakis, A., Martinez-Gili, L., Moreno-Navarrete, J. M., Benhamed, F., Azalbert, V., Blasco-Baque, V., Puig, J., Xifra, G., … & Dumas, M. E. (2018). Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. *Nature Medicine*, *24*(7), 1070–1080. https://doi.org/10.1038/s41591-018-0061-3

Hu, B., Canon, S., Eloe-Fadrosh, E. A., Anubhav Babinski, M., Corilo, Y., Davenport, K., Duncan, W. D., Fagnan, K., Flynn, M., Foster, B., Hays, D., Huntemann, M., Jackson, E. K. P., Kelliher, J., Li, P.-E., Lo, C.-C., Mans, D., McCue, L. A., … Chain, P. S. G. (2022). Challenges in bioinformatics workflows for processing microbiome omics data at scale. *Frontiers in Bioinformatics*, *1*, 1–11. https://doi.org/10.3389/fbinf.2021.826370

Hu, L., Robert, C. A. M., Cadot, S., Zhang, X., Ye, M., Li, B., Manzo, D., Chervet, N., Steinger, T., van der Heijden, M. G. A., Schlaeppi, K., & Erb, M. (2018). Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nature Communications*, *9*(1), 2738. https://doi.org/10.1038/s41467-018-05122-7

Huang, W., Sun, D., Chen, L., & An, Y. (2021). Integrative analysis of the microbiome and metabolome in understanding the causes of sugarcane bitterness. *Scientific Reports*, *11*(1), 6024. https://doi.org/10.1038/s41598-021-85433-w

Huang, X., Fan, Y., Lu, T., Kang, J., Pang, X., Han, B., & Chen, J. (2020). Composition and metabolic functions of the microbiome in fermented grain during light-flavor baijiu fermentation. *Microorganisms*, *8*(9), 1–15. https://doi.org/10.3390/microorganisms8091281

Hultman, J., Johansson, P., & Bjorkroth, J. (2020). Longitudinal metatranscriptomic analysis of a meat spoilage microbiome detects abundant continued fermentation and environmental stress responses during shelf life and beyond. *Applied and Environmental Microbiology*, *86*(24), e01575–20. https://doi.org/10.1128/AEM.01575-20

Hultman, J., Waldrop, M. P., Mackelprang, R., David, M. M., McFarland, J., Blazewicz, S. J., Harden, J., Turetsky, M. R., McGuire, A. D., Shah, M. B., VerBerkmoes, N. C., Lee, L. H., Mavrommatis, K., & Jansson, J. K. (2015). Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature*, *521*(7551), 208–212. https://doi.org/10.1038/nature14238

Huynh-Thu, V. A., Irrthum, A., Wehenkel, L., & Geurts, P. (2010). Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE*, *5*(9), e12776. https://doi.org/10.1371/JOURNAL.PONE.0012776

Ichihashi, Y., Ichihashi, Y., Date, Y., Date, Y., Shino, A., Shimizu, T., Shibata, A., Kumaishi, K., Funahashi, F., Wakayama, K., Yamazaki, K., Umezawa, A., Sato, T., Kobayashi, M., Kamimura, M., Kusano, M., Kusano, M., Che, F. S., O'Brien, M., … Nihei, N. (2020). Multi-omics analysis on an agroecosystem reveals the significant role of organic nitrogen to increase agricultural crop yield. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(25), 14552–14560. https://doi.org/10.1073/PNAS.1917259117/

Impactt, B. (2022). Beta-diversity distance matrices for microbiome sample size and power calculations—How to obtain good estimates. *Computational and Structural Biotechnology Journal*, *20*, 2259–2267. https://doi.org/10.1016/j.csbj.2022.04.032

Jandhyala, S. M., Talukdar, R., Subramanyam, C., Vuyyuru, H., Sasikala, M., & Reddy, D. N. (2015). Role of the normal gut microbiota. *World Journal of Gastroenterology*, *21*(29), 8836–8847. https://doi.org/10.3748/wjg.v21.i29.8787

Jansson, J. K., & Hofmockel, K. S. (2018). The soil microbiome — from metagenomics to metaphenomics. *Current Opinion in Microbiology*, *43*, 162–168. https://doi.org/10.1016/J.MIB.2018.01.013

Jiang, D., Armour, C. R., Hu, C., Mei, M., Tian, C., Sharpton, T. J., & Jiang, Y. (2019). Microbiome multi-omics network analysis: Statistical considerations, limitations, and opportunities. *Frontiers in Genetics*, *10*, 995. https://doi.org/10.3389/FGENE.2019.00995

Kasmanas, J. C., Bartholomäus, A., Corrêa, F. B., Tal, T., Jehmlich, N., Herberth, G., von Bergen, M., Stadler, P. F., de Leon Ferreira de Carvalho, A. C. P., & da Rocha, U. N. (2021). HumanMetagenomeDB: A public repository of curated and standardized metadata for human metagenomes. *Nucleic Acids Research*, *49*(D1), D743–D750. https://doi.org/10.1093/nar/gkaa1031

Kessell, A. K., McCullough, H. C., Auchtung, J. M., Bernstein, H. C., & Song, H. S. (2020). Predictive interactome modeling for precision microbiome engineering. *Current Opinion in Chemical Engineering*, *30*, 77–85. https://doi.org/10.1016/j.coche.2020.08.003

Kim, K. H., Chun, B. H., Kim, J., & Jeon, C. O. (2021). Identification of biogenic amine-producing microbes during fermentation of ganjang, a Korean traditional soy sauce, through metagenomic and metatranscriptomic analyses. *Food Control*, *121*, 107681. https://doi.org/10.1016/j.foodcont.2020.107681

Koh, A., De Vadder, F., Kovatcheva-Datchary, P., & Bäckhed, F. (2016). From dietary fiber to host physiology: Short-chain fatty acids as key bacterial metabolites. *Cell*, *165*(6), 1332–1345. https://doi.org/10.1016/J.CELL.2016.05.041

La Rosa, S, P., Brooks, J. P., Deych, E., Boone, E. L., Edwards, D. J., Wang, Q., Sodergren, E., Weinstock, G., & Shannon, W. D. (2012). Hypothesis testing and power calculations for taxonomic-based human microbiome data. *PLoS ONE*, *7*(12), 1–13. https://doi.org/10.1371/journal.pone.0052078

Lavelle, A., & Sokol, H. (2018). Beyond metagenomics, metatranscriptomics illuminates microbiome functionality in IBD. *Nature Reviews Gastroenterology & Hepatology*, *15*(4), 193–194. https://doi.org/10.1038/nrgastro.2018.15

Lemanceau, P., Blouin, M., Muller, D., & Moënne-Loccoz, Y. (2017). Let the core microbiota be functional. *Trends in Plant Science*, *22*(7), 583–595. https://doi.org/10.1016/j.tplants.2017.04.008

Lessard, M.-H., Viel, C., Boyle, B., St-Gelais, D., & Labrie, S. (2014). Metatranscriptome analysis of fungal strains *Penicillium camemberti* and *Geotrichum candidum* reveal cheese matrix breakdown and potential development of sensory properties of ripened Camembert-type cheese. *BMC Genomics [Electronic Resource]*, *15*, 235. https://doi.org/10.1186/1471-2164-15-235

Li, R., Zheng, M., Zheng, M., Cai, R., Cui, X., Wang, Y., Jiang, X., & Xu, C. (2021). Metagenomic analysis reveals the linkages between bacteria and the functional enzymes responsible for potential ammonia and biogenic amine production in alfalfa silage. *Journal of Applied Microbiology*, *132*(4), 2594–2604. https://doi.org/10.1111/jam.15411

Liu, C., Du, M. X., Abuduaini, R., Yu, H. Y., Li, D. H., Wang, Y. J., Zhou, N., Jiang, M. Z., Niu, P. X., Han, S. S., Chen, H. H., Shi, W. Y., Wu, L., Xin, Y. H., Ma, J., Zhou, Y., Jiang, C. Y., Liu, H. W., & Liu, S. J. (2021). Enlightening the taxonomy darkness of human gut microbiomes with a cultured biobank. *Microbiome*, *9*(1), 1–29. https://doi.org/10.1186/S40168-021-01064-3/

López-Ráez, J. A. (2016). How drought and salinity affect arbuscular mycorrhizal symbiosis and strigolactone biosynthesis? *Planta*, *243*(6), 1375–1385. https://doi.org/10.1007/s00425-015-2435-9

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring

of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution*, *38*(10), 4647–4654. https://doi.org/10.1093/molbev/msab199

Marbach, D., Costello, J. C., Küffner, R., Vega, N. M., Prill, R. J., Camacho, D. M., Allison, K. R., Kellis, M., Collins, J. J., Aderhold, A., Stolovitzky, G., Bonneau, R., Chen, Y., Cordero, F., Crane, M., Dondelinger, F., Drton, M., Esposito, R., Foygel, R., … Zimmer, R. (2012). Wisdom of crowds for robust gene network inference. *Nature Methods*, *9*(8), 796–804. https://doi.org/10.1038/nmeth.2016

Margolin, A. A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Favera, R. D., & Califano, A. (2006). ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics [Electronic Resource]*, *7*, 1–15. https://doi.org/10.1186/1471-2105-7-S1-S7/

Mason, O. U., Hazen, T. C., Borglin, S., Chain, P. S. G., Dubinsky, E. A., Fortney, J. L., Han, J., Holman, H. Y. N., Hultman, J., Lamendella, R., MacKelprang, R., Malfatti, S., Tom, L. M., Tringe, S. G., Woyke, T., Zhou, J., Rubin, E. M., & Jansson, J. K. (2012). Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to deepwater horizon oil spill. *The ISME Journal*, *6*(9), 1715–1727. https://doi.org/10.1038/ismej.2012.59

Mattiello, F., Verbist, B., Faust, K., Raes, J., Shannon, W. D., Bijnens, L., & Thas, O. (2016). A web application for sample size and power calculation in case-control microbiome studies. *Bioinformatics*, *32*(13), 2038–2040. https://doi.org/10.1093/bioinformatics/btw099

McClure, R. S., Lee, J. Y., Chowdhury, T. R., Bottos, E. M., White, R. A., Kim, Y. M., Nicora, C. D., Metz, T. O., Hofmockel, K. S., Jansson, J. K., & Song, H. S. (2020). Integrated network modeling approach defines key metabolic responses of soil microbiomes to perturbations. *Scientific Reports*, *10*(1), 1–9. https://doi.org/10.1038/s41598-020-67878-7

McClure, R. S., Overall, C. C., Mcdermott, J. E., Hill, E. A., Markillie, L. M., Mccue, L. A., Taylor, R. C., Ludwig, M., Bryant, D. A., & Beliaev, A. S. (2016). Network analysis of transcriptomics expands regulatory landscapes in *synechococcus* sp. PCC 7002. *Nucleic Acids Research*, *44*(18), 8810. https://doi.org/10.1093/NAR/GKW737

McClure, R. S., Wendler, J. P., Adkins, J. N., Swanstrom, J., Baric, R., Kaiser, B. L. D., Oxford, K. L., Waters, K. M., & McDermott, J. E. (2019). Unified feature association networks through integration of transcriptomic and proteomic data. *PLOS Computational Biology*, *15*(9), e1007241. https://doi.org/10.1371/JOURNAL.PCBI.1007241

McHugh, A. J., Feehily, C., Tobin, J. T., Fenelon, M. A., Hill, C., & Cotter, P. D. (2018). Mesophilic sporeformers identified in whey powder by using shotgun metagenomic sequencing. *Applied and Environmental Microbiology*, *84*(20), 1–14. https://doi.org/10.1128/AEM.01305-18

Milani, C., Duranti, S., Napoli, S., Alessandri, G., Mancabelli, L., Anzalone, R., Longhi, G., Viappiani, A., Mangifesta, M., Lugli, G. A., Bernasconi, S., Ossiprandi, M. C., Sinderen, D. V., Ventura, M., & Turroni, F. (2019). Colonization of the human gut by bovine bacteria present in parmesan cheese. *Nature Communications*, *10*(1286), 1–12. https://doi.org/10.1038/s41467-019-09303-w

Mishra, A., Singh, L., & Singh, · D. (2022). Unboxing the black box—One step forward to understand the soil microbiome: A systematic review. *Microbial Ecology*, *2022*(1), 1–15. https://doi.org/10.1007/S00248-022-01962-5

Mitchell, A. L., Scheremetjew, M., Denise, H., Potter, S., Tarkowska, A., Qureshi, M., Salazar, G. A., Pesseat, S., Boland, M. A., Hunter, F. M. I., Ten Hoopen, P., Alako, B., Amid, C., Wilkinson, D. J., Curtis, T. P., Cochrane, G., & Finn, R. D. (2018). EBI metagenomics in 2017: Enriching the analysis of microbial communities, from sequence reads to assemblies. *Nucleic Acids Research*, *46*(D1), D726–D735. https://doi.org/10.1093/nar/gkx967

Molina, N. M., Sola-Leyva, A., Haahr, T., Aghajanova, L., Laudanski, P., Castilla, J. A., & Altmäe, S. (2021). Analysing endometrial microbiome: Methodological considerations and recommendations for good practice. *Human Reproduction (Oxford, England)*, *36*(4), 859–879. https://doi.org/10.1093/humrep/deab009

Mota-Gutierrez, J., Ferrocino, I., Giordano, M., Suarez-Quiroz, M. L., Gonzalez-Ríos, O., & Cocolin, L. (2021). Influence of taxonomic and functional content of microbial communities on the quality of fermented cocoa pulp-bean mass. *Applied and Environmental Microbiology*, *87*(14), e0042521. https://doi.org/10.1128/AEM.00425-21

Nannipieri, P., Penton, C. R., Purahong, W., Schloter, M., & van Elsas, J. D. (2019). Recommendations for soil microbiome analyses. *Biology and Fertility of Soils*, *55*(8), 765–766. https://doi.org/10.1007/S00374-019-01409-Z

Nghiem, M. N., Nguyen, V. T., Jeung, E. B., & Vo, T. T. B. (2019). Alternate antimicrobial resistance genes in multidrug resistant *Salmonella* spp. isolated from retail meats in Vietnam using RNA-sequencing analysis. *Journal of Food Safety*, *39*(6), 1–8. https://doi.org/10.1111/jfs.12707

Noel, S. J., Olijhoek, D. W., McLean, F., Løvendahl, P., Lund, P., & Højberg, O. (2019). Rumen and fecal microbial community structure of Holstein and Jersey dairy cows as affected by breed, diet, and residual feed intake. *Animals*, *9*(8). https://doi.org/10.3390/ani9080498

O'Donnell, S. T., Ross, R. P., & Stanton, C. (2020). The progress of multi-omics technologies: Determining function in lactic acid bacteria using a systems level approach. *Frontiers in Microbiology*, *10*(January), 1–17. https://doi.org/10.3389/fmicb.2019.03084

Olmo, R., Wetzels, S. U., Armanhi, J. S. L., Arruda, P., Berg, G., Cernava, T., Cotter, P. D., Araujo, S. C., de Souza, R. S. C., Ferrocino, I., Frisvad, J. C., Georgalaki, M., Hansen, H. H., Kazou, M., Kiran, G. S., Kostic, T., Krauss-Etschmann, S., Kriaa, A., Lange, L., … Wagner, M. (2022). Microbiome research as an effective driver of success stories in agrifood systems—A selection of case studies. *Frontiers in Microbiology*, *13*, https://doi.org/10.3389/fmicb.2022.834622

Overcoming hurdles in sharing microbiome data [Editorial]. (2017). *Nature Microbiology*, *2*(12), 1573. https://doi.org/10.1038/s41564-017-0077-3

Pang, Z., Chen, J., Wang, T., Gao, C., Li, Z., Guo, L., Xu, J., & Cheng, Y. (2021). Linking plant secondary metabolites and plant microbiomes: A review. *Frontiers in Plant Science*, *12*, 621276. https://doi.org/10.3389/fpls.2021.621276

Parente, E., Cocolin, L., De Filippis, F., Zotta, T., Ferrocino, I., O'Sullivan, O., Neviani, E., De Angelis, M., Cotter, P. D., & Ercolini, D. (2016). FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis. *International Journal of Food Microbiology*, *219*, 28–37. https://doi.org/10.1016/j.ijfoodmicro.2015.12.001

Parente, E., Ricciardi, A., & Zotta, T. (2020). The microbiota of dairy milk : A review. *International Dairy Journal*, *107*, 104714. https://doi.org/10.1016/j.idairyj.2020.104714

Pasolli, E., Asnicar, F., Manara, S., Zolfo, M., Karcher, N., Armanini, F., Beghini, F., Manghi, P., Tett, A., Ghensi, P., Collado, M. C., Rice, B. L., DuLong, C., Morgan, X. C., Golden, C. D., Quince, C., Huttenhower, C., & Segata, N. (2019). Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell*, *176*(3), 649–662.e20. https://doi.org/10.1016/j.cell.2019.01.001

Pasolli, E., De Filippis, F., Mauriello, I. E., Cumbo, F., Walsh, A. M., Leech, J., Cotter, P. D., Segata, N., & Ercolini, D. (2020). Large-scale genome-wide analysis links lactic acid bacteria from food with the gut microbiome. *Nature Communications*, *11*(2610), 1–12. https://doi.org/10.1038/s41467-020-16438-8

Pathak, P., Xie, C., Nichols, R. G., Ferrell, J. M., Boehme, S., Krausz, K. W., Patterson, A. D., Gonzalez, F. J., & Chiang, J. Y. L. (2018). Intestine farnesoid x receptor agonist and the gut microbiota activate G-protein bile acid receptor-1 signaling to improve metabolism. *Hepatology*, *68*(4), 1574–1588. https://doi.org/10.1002/hep.29857

Pineda, S., Bunis, D. G., Kosti, I., & Sirota, M. (2020). Data integration for immunology. *Annual Review of Biomedical Data Science*, *3*(1), 113–136. https://doi.org/10.1146/annurev-biodatasci-012420-122454

Pinu, F. R., Beale, D. J., Paten, A. M., Kouremenos, K., Swarup, S., Schirra, H. J., & Wishart, D. (2019). Systems biology and multi-omics integration: Viewpoints from the metabolomics research community. *Metabolites*, *9*(4), 1–31. https://doi.org/10.3390/metabo9040076

Planell, N., Lagani, V., Sebastian-Leon, P., van der Kloet, F., Ewing, E., Karathanasis, N., Urdangarin, A., Arozarena, I., Jagodic, M., Tsamardinos, I., Tarazona, S., Conesa, A., Tegner, J., & Gomez-Cabrero, D. (2021). STATegra: Multi-omics data integration – a conceptual scheme with a bioinformatics pipeline. *Frontiers in Genetics*, *12*, 620453. https://doi.org/10.3389/fgene.2021.620453

Quigley, L., O'Sullivan, D. J., Daly, D., O'Sullivan, O., Burdikova, Z., Vana, R., Beresford, T. P., Ross, R. P., Fitzgerald, G. F., McSweeney, P. L. H., Giblin, L., Sheehan, J. J., & Cotter, P. D. (2016). Thermus and the pink discoloration defect in cheese. *MSystems*, *1*(3), 1–12. https://doi.org/10.1128/msystems.00023-16

Rao, G., Zhang, J., Liu, X., Li, X., & Wang, C. (2021). Combined metabolome and transcriptome profiling reveal optimal harvest strategy model based on different production purposes in olive. *Foods*, *10*, 360. https://doi.org/10.3390/foods10020360

Raymond, F., Boissinot, M., Ouameur, A. A., Déraspe, M., Plante, P. L., Kpanou, S. R., Bérubé, È., Huletsky, A., Roy, P. H., Ouellette, M., Bergeron, M. G., & Corbeil, J. (2019). Culture-enriched human gut microbiomes reveal core and accessory resistance genes. *Microbiome*, *7*(1), 1–13. https://doi.org/10.1186/s40168-019-0669-7

Read, J., & Sharma, S. (2021). Hypothesis-driven science in large-scale studies: The case of GWAS. *Biology and Philosophy*, *36*(5), 1–21. https://doi.org/10.1007/s10539-021-09823-0

Rizaludin, M. S., Stopnisek, N., Raaijmakers, J. M., & Garbeva, P. (2021). The chemistry of stress: Understanding the 'cry for help' of plant roots. *Metabolites*, *11*(6), 357. https://doi.org/10.3390/metabo11060357

Rooks, M. G., & Garrett, W. S. (2016). Gut microbiota, metabolites and host immunity. *Nature Reviews Immunology*, *16*(6), 341–352. https://doi.org/10.1038/nri.2016.42

Ros, M., Blaya, J., Baldrian, P., Bastida, F., Richnow, H. H., Jehmlich, N., & Pascual, J. A. (2018). In vitro elucidation of suppression effects of composts to soil-borne pathogen *Phytophthora nicotianae* on pepper plants using 16S amplicon sequencing and metaproteomics. *Renewable Agriculture and Food Systems*, *35*(2), 206–214. https://doi.org/10.1017/S1742170518000467

Ruiz-Perez, D., Lugo-Martinez, J., Bourguignon, N., Mathee, K., Lerner, B., Bar-Joseph, Z., & Narasimhan, G. (2021). Dynamic bayesian networks for integrating multi-omics time series microbiome data. *MSystems*, *6*(2). https://doi.org/10.1128/msystems.01105-20

Ryan, M. J., Schloter, M., Berg, G., Kinkel, L. L., Eversole, K., Macklin, J. A., Rybakova, D., & Sessitsch, A. (2021). Towards a unified data infrastructure to support european and global microbiome research: A call to action. *Environmental Microbiology*, *23*(1), 372–375. https://doi.org/10.1111/1462-2920.15323

Ryan, M. J., Schloter, M., Berg, G., Kostic, T., Kinkel, L. L., Eversole, K., Macklin, J. A., Schelkle, B., Kazou, M., Sarand, I., Singh, B. K., Fischer, D., Maguin, E., Ferrocino, I., Lima, N., McClure, R. S., Charles, T. C., de Souza, R. S. C., Kiran, G. S., … Sessitsch, A. (2021). Development of microbiome biobanks—Challenges and opportunities. *Trends in Microbiology*, *29*(2), 89–92. https://doi.org/10.1016/j.tim.2020.06.009

Santiago-Rodriguez, T. M., & Hollister, E. B. (2021). Multiomic data integration: A review of concepts, considerations, and approaches. *Seminars in Perinatology*, *45*(6), 151456. https://doi.org/10.1016/j.semperi.2021.151456

Santos-Medellín, C., Edwards, J., Liechty, Z., Nguyen, B., & Sundaresan, V. (2017). Drought stress results in a compartment-specific restructuring of the rice root-associated microbiomes. *MBio*, *8*(4), 1–15. https://doi.org/10.1128/mBio.00764-17

Schirmer, M., Franzosa, E. A., Lloyd-Price, J., McIver, L. J., Schwager, R., Poon, T. W., Ananthakrishnan, A. N., Andrews, E., Barron, G., Lake, K., Prasad, M., Sauk, J., Stevens, B., Wilson, R. G., Braun, J., Denson, L. A., Kugathasan, S., McGovern, D. P. B., Vlamakis, H., …, & Huttenhower, C. (2018). Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. *Nature Microbiology*, *3*(3), 337–346. https://doi.org/10.1038/s41564-017-0089-z

Sulaiman, J., Gan, H. M., Yin, W. F., & Chan, K. G. (2014). Microbial succession and the functional potential during the fermentation of Chinese soy sauce brine. *Frontiers in Microbiology*, *5*, 556. https://doi.org/10.3389/fmicb.2014.00556

Sultan, S., El-Mowafy, M., Elgaml, A., Ahmed, T. A. E., Hassan, H., & Mottawea, W. (2021). Metabolic influences of gut microbiota dysbiosis on inflammatory bowel disease. *Frontiers in Physiology*, *12*, 715506. https://doi.org/10.3389/FPHYS.2021.715506

Tanca, A., Palomba, A., Fraumene, C., Pagnozzi, D., Manghina, V., Deligios, M., Muth, T., Rapp, E., Martens, L., Addis, M. F., & Uzzau, S. (2016). The impact of sequence database choice on metaproteomic results in gut microbiota studies. *Microbiome*, *4*, 51. https://doi.org/10.1186/s40168-016-0196-8

Teixeira, P. J. P. L., Colaianni, N. R., Law, T. F., Conway, J. M., Gilbert, S., Li, H., Salas-González, I., Panda, D., Del Risco, N. M., Finkel, O. M., Castrillo, G., Mieczkowski, P., Jones, C. D., & Dangl, J. L. (2021). Specific modulation of the root immune system by a community of commensal bacteria. *Proceedings of the National Academy of Sciences*, *118*(16), e2100678118. https://doi.org/10.1073/pnas.2100678118

Tennoune, N., Chan, P., Breton, J., Legrand, R., Chabane, Y. N., Akkermann, K., Järv, A., Ouelaa, W., Takagi, K., Ghouzali, I., Francois, M., Lucas, N., Bole-Feysot, C., Pestel-Caron, M., do Rego, J. C., Vaudry, D., Harro, J., Dé, E., Déchelotte, P., & Fetissov, S. O. (2014). Bacterial ClpB heat-shock protein, an antigen-mimetic of the anorexigenic peptide α-MSH, at the origin of eating disorders. *Translational Psychiatry*, *4*, e458. https://doi.org/10.1038/tp.2014.98

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., Navas-Molina, J. A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González, A., Morton, J. T., Mirarab, S., Xu, Z. Z., Jiang, L., … Zhao, H. (2017). A communal catalogue reveals earth's multiscale microbial diversity. *Nature*, *551*(7681), 457–463. https://doi.org/10.1038/nature24621

Tripathi, A., Marotz, C., Gonzalez, A., Vázquez-Baeza, Y., Song, S. J., Bouslimani, A., McDonald, D., Zhu, Q., Sanders, J. G., Smarr, L., Dorrestein, P. C., & Knight, R. (2018). Are microbiome studies ready for hypothesis-driven research? *Current Opinion in Microbiology*, *44*, 61–69. https://doi.org/10.1016/j.mib.2018.07.002

Trivedi, P., Batista, B. D., Bazany, K. E., & Singh, B. K. (2022). Plant–microbiome interactions under a changing world: Responses, consequences and perspectives. *New Phytologist*, 1–9. https://doi.org/10.1111/nph.18016

Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., & Singh, B. K. (2020). Plant–microbiome interactions: From community assembly to plant health. *Nature Reviews Microbiology*, *18*(11), 607–621. https://doi.org/10.1038/s41579-020-0412-1

Trivedi, P., Mattupalli, C., Eversole, K., & Leach, J. E. (2021). Enabling sustainable agriculture through understanding and enhancement of microbiomes. *New Phytologist*, *230*(6), 2129–2147. https://doi.org/10.1111/NPH.17319

Tsagiopoulou, M., Pechlivanis, N., Maniou, M. C., & Psomopoulos, F. (2022). InterTADs: Integration of multi-omics data on topologically associated domains, application to chronic lymphocytic leukemia. *NAR Genomics and Bioinformatics*, *4*(1), 1–13. https://doi.org/10.1093/nargab/lqab121

Vandenkoornhuyse, P., Quaiser, a., Duhamel, M., Le Van, a., & Dufresne, A. (2015). The importance of the microbiome of the plant holobiont. *New Phytologist*, *206*, 1196–1206.

Vangay, P., Burgin, J., Johnston, A., Beck, K. L., Berrios, D. C., Blumberg, K., Canon, S., Chain, P., Chandonia, J.-M., Christianson, D., Costes, S. V., Damerow, J., Duncan, W. D., Dundore-Arias, J. P., Fagnan, K., Galazka, J. M., Gibbons, S. M., Hays, D., Hervey, J., … Eloe-Fadrosh, E. A. (2021). Microbiome metadata standards: Report of the national microbiome data collaborative's workshop and follow-on activities. *MSystems*, *6*(1). https://doi.org/10.1128/msystems.01194-20

Verce, M., De Vuyst, L., & Weckx, S. (2019). Shotgun metagenomics of a water kefir fermentation ecosystem reveals a novel *Oenococcus* species. *Frontiers in Microbiology*, *10*, 1–16. https://doi.org/10.3389/fmicb.2019.00479

Viacava, K., Qiao, J., Janowczyk, A., Poudel, S., Jacquemin, N., Meibom, K. L., Shrestha, H. K., Reid, M. C., Hettich, R. L., & Bernier-Latmani, R. (2022). Meta-omics-aided isolation of an elusive anaerobic arsenic-methylating soil bacterium. *The ISME Journal*, 1–10. https://doi.org/10.1038/s41396-022-01220-z

Wallace, J. G., Kremling, K. A., Kovar, L. L., & Buckler, E. S. (2018). Quantitative genetics of the maize leaf microbiome. *Phytobiomes Journal*, *2*(4), 208–224. https://doi.org/10.1094/PBIOMES-02-18-0008-R

Walsh, A. M., Crispie, F., O'Sullivan, O., Finnegan, L., Claesson, M. J., & Cotter, P. D. (2018). Species classifier choice is a key consideration when analysing low-complexity food microbiome data. *Microbiome*, *6*(50). https://doi.org/10.1186/s40168-018-0437-0

Wang, L., Zhu, L., & Qin, S. (2019). Gut microbiota modulation on intestinal mucosal adaptive immunity. *Journal of Immunology Research*, 2019, 4735040. https://doi.org/10.1155/2019/4735040

Wang, T. J., Larson, M. G., Vasan, R. S., Cheng, S., Rhee, E. P., McCabe, E., Lewis, G. D., Fox, C. S., Jacques, P. F., Fernandez, C., O'Donnell, C. J., Carr, S. A., Mootha, V. K., Florez, J. C., Souza, A., Melander, O., Clish, C. B., & Gerszten, R. E. (2011). Metabolite profiles and the risk of developing diabetes. *Nature Medicine*, *17*(4), 448–453. https://doi.org/10.1038/nm.2307

Wang, Z., Roberts, A. B., Buffa, J. A., Levison, B. S., Zhu, W., Org, E., Gu, X., Huang, Y., Zamanian-Daryoush, M., Culley, M. K., DiDonato, A. J., Fu, X., Hazen, J. E., Krajcik, D., DiDonato, J. A., Lusis, A. J., & Hazen, S. L. (2015). Non-lethal inhibition of gut microbial trimethylamine production for the treatment of atherosclerosis. *Cell*, *163*(7), 1585–1595. https://doi.org/10.1016/j.cell.2015.11.055

Wu, Y., Xia, M., Zhao, N., Tu, L., Xue, D., Zhang, X., Zhao, C., Cheng, Y., Zheng, Y., & Wang, M. (2021). Metabolic profile of main organic acids and its regulatory mechanism in solid-state fermentation of Chinese cereal vinegar. *Food Research International*, *145*, 110400. https://doi.org/10.1016/j.foodres.2021.110400

Xiong, W., Abraham, P. E., Li, Z., Pan, C., & Hettich, R. L. (2015). Microbial metaproteomics for characterizing the range of metabolic functions and activities of human gut microbiota. *Proteomics*, *15*(20), 3424–3438. https://doi.org/10.1002/pmic.201400571

Xu, L., & Coleman-Derr, D. (2019). Causes and consequences of a conserved bacterial root microbiome response to drought stress. *Current Opinion in Microbiology*, *49*, 1–6. https://doi.org/10.1016/j.mib.2019.07.003

Xu, L., Naylor, D., Dong, Z., Simmons, T., Pierroz, G., Hixson, K. K., Kim, Y.-M., Zink, E. M., Engbrecht, K. M., Wang, Y., Gao, C., DeGraaf, S., Madera, M. A., Sievert, J. A., Hollingsworth, J., Birdseye, D., Scheller, H. V., Hutmacher, R., Dahlberg, J., … Coleman-Derr, D. (2018). Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. *Proceedings of the National Academy of Sciences*, *115*(18), E4284–E4293. https://doi.org/10.1073/pnas.1717308115

Xu, Y., Tong, Z., Zhang, X., Wang, Y., Fang, W., Li, L., & Luo, Z. (2019). Unveiling the mechanisms for the plant volatile organic compound linalool to control gray mold on strawberry fruits. *Journal of Agricultural and Food Chemistry*, *67*(33), 9265–9276. https://doi.org/10.1021/acs.jafc.9b03103

Yan, J., Risacher, S. L., Shen, L., & Saykin, A. J. (2017). Network approaches to systems biology analysis of complex disease: Integrative methods for multi-omics data. *Briefings in Bioinformatics*, *19*(6), 1370–1381. https://doi.org/10.1093/bib/bbx066

Yang, C., Li, M., & Peng, B. (2021). Transcriptomic analysis reveals the metabolic mechanism of patulin by *Saccharomyces cerevisiae* during fermentation. *Lwt*, *149*, 111808. https://doi.org/10.1016/j.lwt.2021.111808

Yao, Q., Li, Z., Song, Y., Wright, S. J., Guo, X., Tringe, S. G., Tfaily, M. M., Paša-Tolić, L., Hazen, T. C., Turner, B. L., Mayes, M. A., & Pan,

C. (2018). Community proteogenomics reveals the systemic impact of phosphorus availability on microbial functions in tropical soil. *Nature Ecology & Evolution*, *2*(3), 499–509. https://doi.org/10.1038/s41559-017-0463-5

Yi, H.-S., Yang, J. W., Ghim, S.-Y., & Ryu, C.-M. (2011). A cry for help from leaf to root. *Plant Signaling & Behavior*, *6*(8), 1192–1194. https://doi.org/10.4161/psb.6.8.15780

Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J. R., Amaral-Zettler, L., Gilbert, J. A., Karsch-Mizrachi, I., Johnston, A., Cochrane, G., Vaughan, R., Hunter, C., Park, J., Morrison, N., Rocca-Serra, P., Sterk, P., Arumugam, M., Bailey, M., Baumgartner, L., … Glöckner, F. O. (2011). Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nature Biotechnology*, *29*(5), 415–420. https://doi.org/10.1038/nbt.1823

Yuan, J., Chen, C., Cui, J., Lu, J., Yan, C., Wei, X., Zhao, X., Li, N. N., Li, S., Xue, G., Cheng, W., Li, B., Li, H., Lin, W., Tian, C., Zhao, J., Han, J., An, D., Zhang, Q., …, & Liu, D. (2019). Fatty liver disease caused by high-alcohol-producing *Klebsiella pneumoniae*. *Cell Metabolism*, *30*(4), 675–688.e7. https://doi.org/10.1016/j.cmet.2019.08.018

Zapalska-Sozoniuk, M., Chrobak, L., Kowalczyk, K., & Kankofer, M. (2019). Is it useful to use several "omics" for obtaining valuable results? *Molecular Biology Reports*, *46*(3), 3597–3606. https://doi.org/10.1007/s11033-019-04793-9

Zehavi, T., Probst, M., & Mizrahi, I. (2018). Insights into culturomics of the rumen microbiome. *Frontiers in Microbiology*, *9*, 1–10. https://doi.org/10.3389/fmicb.2018.01999

Zhang, Y., Kastman, E. K., Guasto, J. S., & Wolfe, B. E. (2018). Fungal networks shape dynamics of bacterial dispersal and community assembly in cheese rind microbiomes. *Nature Communications*, *9*(1), 1–12. https://doi.org/10.1038/s41467-017-02522-z

Zhao, C., Su, W., Mu, Y., Mu, Y., & Jiang, L. (2021). Integrative metagenomics–metabolomics for analyzing the relationship between microorganisms and non-volatile profiles of traditional Xiaoqu. *Frontiers in Microbiology*, *11*, 1–17. https://doi.org/10.3389/fmicb.2020.617030

Zhao, C. C., & Eun, J. B. (2020). Shotgun metagenomics approach reveals the bacterial community and metabolic pathways in commercial hongeo product, a traditional Korean fermented skate product. *Food Research International*, *131*, 109030. https://doi.org/10.1016/j.foodres.2020.109030

Zheng, Y., Ran, Y., Zhang, H., Wang, B., & Zhou, L. (2021). The microbiome in autoimmune liver diseases: Metagenomic and metabolomic changes. *Frontiers in Physiology*, *12*, 1546. https://doi.org/10.3389/FPHYS.2021.715852

Zhu, F. (2016). Staling of Chinese steamed bread: Quantification and control. *Trends in Food Science and Technology*, *55*, 118–127. https://doi.org/10.1016/j.tifs.2016.07.009

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ferrocino, I., Rantsiou, K., McClure, R., Kostic, T., de Souza, R. S. C., Lange, L., FitzGerald, J., Kriaa, A., Cotter, P., Maguin, E., Schelkle, B., Schloter, M., Berg, G., Sessitsch, A., Cocolin, L., & The MicrobiomeSupport Consortium. (2023). The need for an integrated multi-OMICs approach in microbiome science in the food system. *Comprehensive Reviews in Food Science and Food Safety*, *22,* 1082–1103. https://doi.org/10.1111/1541-4337.13103