Exploring sub-species variation in food microbiomes: a roadmap to reveal hidden diversity and functional potential

Lena Flörl,[1] Annina Meyer,[1] Nicholas A. Bokulich[1]

AUTHOR AFFILIATION See affiliation list on p. 11.

ABSTRACT Within-species diversity of microorganisms in food systems significantly shapes community function. While next-generation sequencing (NGS) methods have advanced our understanding of microbiomes at the community level, it is essential to recognize the importance of within-species variation for understanding and predicting the functional activities of these communities. This review highlights the substantial variation observed among microbial species in food systems and its implications for their functionality. We discuss a selection of key species in fermented foods and food systems, highlighting examples of strain-level variation and its influence on quality and safety. We present a comprehensive roadmap of methodologies aimed at uncovering this often overlooked underlying diversity. Technologies like long-read marker-gene or shotgun metagenome sequencing offer enhanced resolution of microbial communities and insights into the functional potential of individual strains and should be integrated with techniques such as metabolomics, metatranscriptomics, and metaproteomics to link strain-level microbial community structure to functional activities. Furthermore, the interactions between viruses and microbes that contribute to strain diversity and community stability are also critical to consider. This article highlights existing research and emphasizes the importance of incorporating within-species diversity in microbial community studies to harness their full potential, advance fundamental science, and foster innovation.

KEYWORDS microbiome, metagenomics, food systems, strain profiling

FOOD SYSTEMS ARE AN IDEAL MODEL TO SHOWCASE AND INVESTIGATE WITHIN-SPECIES VARIATION

Different strains belonging to the same species can vary drastically in their functional potential, resistances, and pathogenicity. In microbial communities, within-species variations can contribute to the distinct function and properties of entire microbiomes. In food systems, such within-species variation therefore conveys unique characteristics and influences the quality and safety of a product (1–5). The resulting properties can be unique to an extent that it allows for the authentication of protected designation of origin and protected geographical indication products by tracking distinct associated strains (6, 7). Furthermore, strain tracking plays a crucial role in monitoring and mitigating disease outbreaks and contamination in food production facilities (8). Hence, as these effects are particularly easily observable, literally tastable, fermented foods and food systems are an ideal model for studying within-species diversity. Additionally, their wide availability, greater number of culturable species, and relatively simpler microbial consortia enhance their suitability as study systems (9). But how can these complex microbial communities and their variation within be efficiently and comprehensively studied? Here, we give examples for the effect of within-species diversity in food systems, review state-of-the-art methodologies, and showcase their applications.

THE EMERGENCE OF WITHIN-SPECIES DIVERSITY AND ITS IMPLICATIONS

Microbial communities in food systems naturally assemble from a species pool originating from raw materials, production facilities, and likely even the food producers themselves (10, 11), with environmental conditions acting as a selective pressure (12). Genetic diversification selects for the emergence of novel strains with novel functions, whose growth is enhanced, reduced, or even inhibited by biotic and abiotic factors encountered in the surrounding environment (12). Additionally, bacteria and archaea have an open gene pool and exchange genetic elements through horizontal gene transfer (HGT), a process that occurs more frequently in related organisms but is even observed across domains (13). The magnitude of this gene flow is determined by the genetic similarity between organisms as well as their natural competence and recombination machinery. Generally, a higher degree of genetic similarity, or related-ness, increases DNA exchange, subsequent recombination, and genomic integration of novel genes or even entire gene clusters (14). Ecological overlap, i.e., the close physical proximity and shared ecological niche with exposure to the same selective conditions, increases the likelihood of exchanging genes (14). Such highly localized, micro-evolu-tionary changes can lead to a large genomic diversity for single species within the same sample (15). This genetic flexibility complicates the delineation of species and, even more so, strains. Hence, the terminology surrounding microbial classification is often inconsistent, leading to significant confusion (15). Species are typically delineated by having >95% average nucleotide identity (ANI) across the core genome, and further subdivided into distinct strains which are genetically (>99% ANI) and/or phenotypically similar, with the specific definition potentially varying by research field (14, 16) (see Fig. 1). Strains can highly vary in metabolic capabilities, phenotype, or stress response. This ostensible redundancy of strains co-existing in the same system is expected to produce unique characteristics and to increase community stability against viral invasions and environmental fluctuations (17).

EXAMPLES FOR WITHIN-SPECIES VARIATION IN FOOD SYSTEMS

The described genetic stratification leads to phenotypic and functional variation, which is evident across numerous key species in food systems. First and foremost, *Saccharomyces cerevisiae* is one of the most important species in fermented foods and contains a huge and highly specialized variety of strains (18). This within-species diversity arose from the millennia of co-domestication of the yeast with humans fermenting fruits and cereals. Hence, strains demonstrate significant adaptation to specific nutrient sources and demonstrate genotypic and phenotypic variation between different types of fermentations (e.g., wine, beer, sake, bread), as well as having highly variable stress responses (e.g., to sulfur or ethanol), nitrogen and vitamin acquisition (19). On the other hand, many strains of *S. cerevisiae* act as food spoilage organisms, exhibiting undesirable qualities, for example, in olives or cheese (20), and some strains are opportunistic human pathogens (21) or probiotics in the case of *S. cerevisiae* var. *boulardii,* the most common commercial human probiotic yeast (22).

Similarly, *Penicillium roqueforti* strains exhibit particularly high phenotypic variability (e.g., spore formation, growth rate, salt tolerance) and can strongly vary in the production of secondary metabolites (23). Certain compounds produced by some strains are considered beneficial to human health (e.g., mycophenolic acid and andrastin A) (24), while other strains can produce secondary mycotoxins or other secondary metabolites that can cause serious illness (23). Additionally, proteolytic activities as well as aroma formation vary strongly, and specific strains are therefore kept and propagated through back-slopping for centuries to produce desired characteristics (23). *P. roqueforti* strains form such distinct genetic clusters in different ecological niches, such as in cheese production and food spoilage, that they are no longer capable of genetic recombination (23).

The within-species variability of many other foodborne yeasts and bacteria diverges between beneficial and detrimental. For example, various strains of the yeast *Pichia*
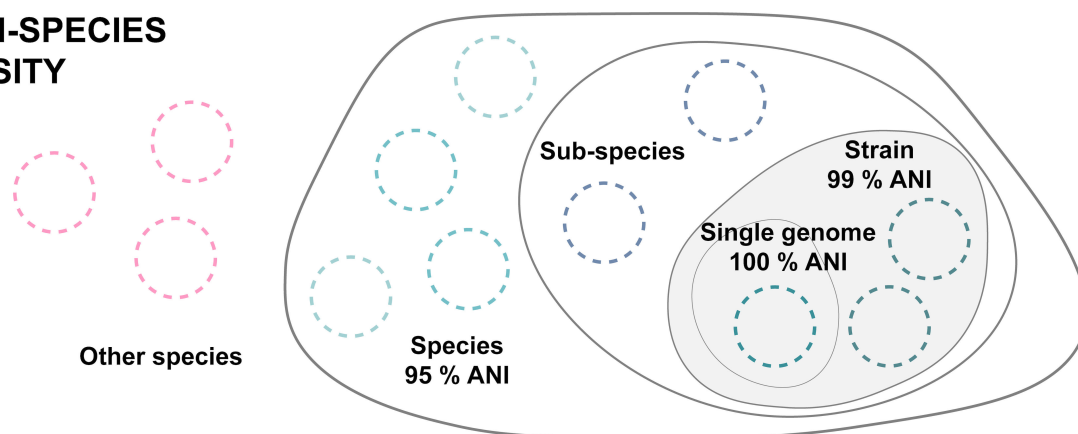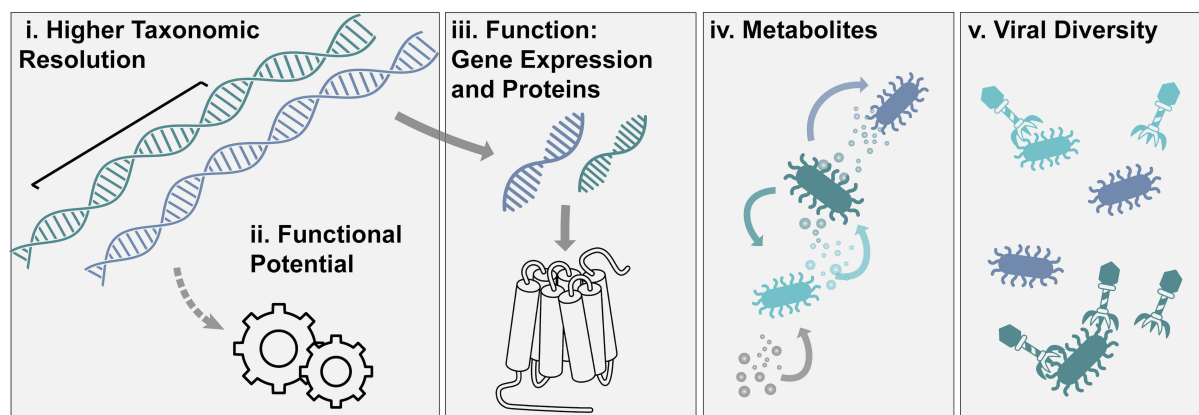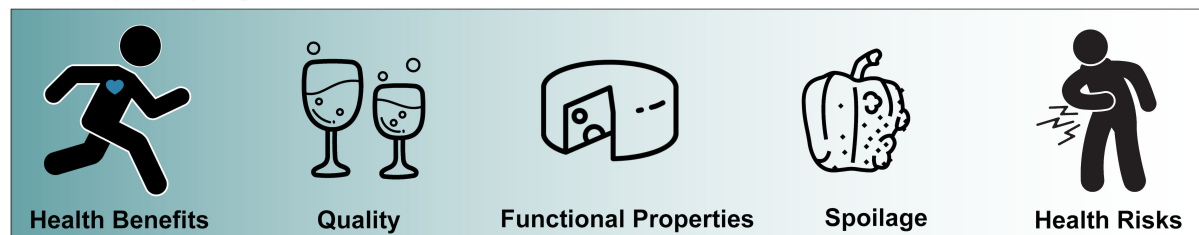
FIG 1 Summarizes the contents of this article. WITHIN-SPECIES DIVERSITY: The term "strain" lacks a definitive definition and varies by field, but is commonly understood as a phenotypically and/or genotypically similar subset of a species (ANI in the core genome >99%). A species can comprise multiple sub-species, each sub-species may contain several strains, and a strain may consist of multiple distinct genomes (for a more detailed review on this complex issue, see Van Rossum et al. [14]). STUDYING DIFFERENCES: Different strains can have distinct properties, and here, we outline different technologies to study these. (i) Covering larger regions of the genome with long-read sequencing or whole metagenome sequencing provides higher resolution of microbial communities. (ii) This metagenomic data can also reveal the functional potential of individual strains or communities. (iii) Methods such as metatranscriptomics and metaproteomics allow for a deeper exploration of microbial functions and the distinct effects of different strains. (iv) Metabolomics, the analysis of the chemical environment surrounding and modulated by microbes, offers insights into strain development and functionality within ecological systems. (v) Additionally, understanding the interactions between viruses and microbes is crucial, as these relationships play a key role in shaping strain diversity and maintaining community stability. IMPLICATIONS: These diverse functionalities and phenotypes have distinct implications for food production.

kudriavzevii exhibit human pathogenic or probiotic properties (25). The species is highly relevant for the food industry, used to convey functional properties (e.g., shelf-life, texture) or to improve sensory and nutritional qualities, while some strains also exhibit food spoilage potential (26). This within-species variation, which can determine whether *P. kudriavzevii* acts as a beneficial strain in food fermentations or as a spoilage agent, is

also evident in species such as *Debaryomyces hansenii* (27), *Hanseniaspora uvarum* (28), *Levilactobacillus brevis* (formerly *Lactobacillus brevis*) (29), and *Penicillium roqueforti* (23).

This list of the implications of within-species variation in food systems could be extended to include *Lactiplantibacillus plantarum* (formerly *Lactobacillus plantarum*), known for its notably large genome, exceptional metabolic adaptability, and extensive ecological range (30). Similarly, *Hanseniaspora uvarum* demonstrates overall phenotypic uniformity yet harbors substantial genetic diversity that clusters distinctly across spatial and temporal scales (28). Such within-species variations, with critical functional relevance in food systems, have also been documented for various other fungi and bacteria (31–41).

The relative importance of strain-level diversity likely varies across fermented foods and food systems. In some cases, functional variation is primarily species-driven, as with *Saccharomyces cerevisiae* ensuring alcohol production in wine fermentation. In other contexts, any strain variations are critical, such as wild *S. cerevisiae* strains contaminating beer fermentations and causing flaws like over-attenuation, haze, or off-flavors (42). We speculate that systems prone to overgrowth by cross-contaminants or wild strains, like non-starter lactic acid bacteria (LAB) in cheese (43), may depend more heavily on strain-level variation and therefore would require stricter strain monitoring and management to maintain desired fermentation outcomes. While within-species variation undoubtedly contributes to the unique properties of a fermented product and food safety, the extent to which it interacts with environmental factors and interspecies dynamics remains difficult to disentangle. Comparative studies evaluating the relative importance of strain-level variation across food systems are still notably lacking.

CURRENT APPROACHES TO STUDY MICROBIAL DIVERSITY: DIVERGENCE BETWEEN CULTIVATION-BASED AND SEQUENCING-BASED METHODS

Techniques to isolate and grow microorganisms in solid or liquid media laid the foundation of the field of microbiology in the mid-19th century, and to this day are key routine applications for enumerating viable organisms or studying their phenotype and behavior, e.g., metabolism (44). Cultivation therefore remains highly relevant to discover new strains (4) and investigate their functionality in response to abiotic factors or community interactions (2, 9). Notably, the growth of microorganisms depends on chosen cultivation conditions and is inherently biased by the media type, oxygen availability, incubation time, and temperature, etc. Due to this selectivity, the bottom-up approach of cultivation-based methods lacks information on the composition and interactions within the actual microbial community (45). Axenic cultivation often poorly represents the natural environments in which microorganisms are found, and hence skews assessments of microbial phenotypes, e.g., growth or metabolism (46). Studies using both culture-dependent and -independent approaches have shown that cultivation does not recover the actual microbial diversity and often overlooks even dominant taxa or fails to detect significant differences between production sites (47, 48). Culture-independent approaches utilizing next-generation sequencing (NGS) technology enable a top-down approach to profile microbial communities. NGS is based on massive parallelization of millions of sequencing reactions occurring simultaneously (49). This facilitates rapid and deep sequencing of whole genomes of isolated strains, but also allows for the sequencing of all present DNA in a sample (shotgun metagenomics) or profiling microbial communities with representative marker genes, e.g., 16S SSU rRNA for bacteria and archaea, or the internal transcribed spacer domain for fungi (49). The relative inexpensiveness, availability of robust protocols, as well as broad establishment in the field makes marker-gene NGS using Illumina sequencing-by-synthesis technology the most commonly used tool to profile microbial communities in food systems (50). However, the actual taxonomic resolution depends on the marker gene and primer choice, among other factors. The most commonly used bacterial primers cover only one or two hypervariable regions of the 16S rRNA gene and are not sufficient to reliably resolve genetic differences within a species, and even the full-length 16S rRNA gene

is insufficient to fully resolve some species (4, 51). Hence, focusing exclusively on the broader community level risks overlooking the importance of within-species variation. To prevent a divergence between these two approaches that may create isolated subfields, we suggest bridging this emerging gap between cultivation-dependent and -independent approaches for studying microbial diversity in fermented foods.

A ROADMAP TO STUDY MICROBIAL DIVERSITY CONSIDERING WITHIN-SPECIES VARIATION IN FOOD SYSTEMS

Technologies that provide higher taxonomic resolution are crucial for comparing microbial communities in similar ecological niches or tracking microbial dispersal

The genetic differences between sub-species variants within similar ecological niches, which have been exposed to the same selective pressures, can be minimal. Thus, relying solely on Illumina marker-gene sequencing, which at the time of writing provides a maximum read length of 2×300 bases (52), may be insufficient to detect subtle differences when comparing microbiomes in closely related ecological niches (Table 1). Technologies able to retrieve longer-read lengths, like Oxford Nanopore sequencing or single-molecule real-time sequencing used in PacBio devices, can cover a larger section of a marker gene and accordingly increase resolution, provided that appropriate markers are targeted. For comparison, long-read sequencing technologies offer read lengths up to 20 kilobases for PacBio (53) and up to megabase-scale long reads for Nanopore (54). While long-read marker-gene sequencing can thereby reliably resolve intragenomic variation, e.g., of 16S rRNA alleles within a species, current taxonomic assignment methods often lack the resolution to distinguish sub-species variants due to the absence of strain-level annotated reference databases (55). Furthermore, the issue of strains potentially containing multiple, distinct rRNA operon copies remains unresolved by existing approaches (55) (see Table 1).

Therefore, a further step up in distinguishing within-species variants is metagenomic-assembled genomes (MAGs), which are reconstructed from shotgun metagenome sequencing data. Since shotgun sequencing gives insight into the entire genomic information content of a sample, it also allows for entirely new strain discovery, provided that sequencing was performed at an adequate depth to enable assembly of MAGs or use of read-based methods for strain detection (62). This is further aided by using long-read sequencing in metagenomics, which improves the identification of structural variations and thereby drastically enhances taxonomic resolution (63). Yet, as long-read marker-gene sequencing and shotgun metagenome sequencing (in sufficient depth to assemble MAGs) are still significantly more expensive than Illumina-based marker-gene sequencing (i.e., roughly an order-of-magnitude difference at the time of writing), these technologies are sometimes applied only to a subset of samples. Notably, shotgun metagenome data additionally require significantly more computational power and resources for analysis, compared to marker-gene amplicon sequence data. Sequencing quality and depth also limit how many organisms can be distinguished down to a sub-species level, which is commonly only possible for dominant species (59) (see Table 1). Despite these barriers, metagenomics analysis is considered to have huge potential in monitoring and mitigating contaminations and outbreaks in food production facilities by enabling tracing and investigating the persistence of distinct strains (8).

Ultimately, single-cell DNA sequencing (scDNA-seq) enables the resolution of individual genomes from single organisms in complex microbial communities through single-amplified genomes (61). Despite remaining challenges and barriers such as variable lysis efficiency, uneven genome amplification, high computational demands, elevated costs, as well as relatively low throughput, scDNA-seq methods hold significant potential for advancing our understanding of within-species microbial diversity (64, 65) (see Table 1). However, these techniques have yet to be broadly applied in microbiome

TABLE 1 Different DNA sequencing techniques provide varying degrees of resolution, with higher resolution typically associated with increased costs[a]

Method	Resolution	Advantages[b]	Caveats[c]	Application examples	Cost
Illumina marker-gene sequencing	Species, sub-species variants (14)	+ Robust + Well-established protocols + Extensive databases + Robust bioinformatics tools	Resolution strongly depends on – choice of marker gene – primers – bioinformatic tools (56)	• Microbial community profiling • Monitoring population shifts • Initial screening and basic quality control	$
Long-read marker-gene sequencing	Species, strain resolution possible (55)	+ Increased resolution + Improved phylogenetics (57)	Still difficulties in distinguishing strains: – marker genes too conserved – primer/target selection – multiple marker-gene copies – no databases with annotated variants (58)	• Pathogen detection and tracking (58) • Microbial fingerprinting: authentication (7)	$$
Metagenome sequencing	Strain level (59)	+ Discovery of novel strains + Assessing functional potential	Resolution strongly depends on – sequencing depth – run quality – computational resources (60)	• Identifying novel probiotics • Revealing strains and associated mechanisms of food spoilage (5, 8)	$$$
Single-cell DNA sequencing	Single amplified genomes (61)	+ Highest resolution of single-microbe genomic information (61)	– not widely applied in microbiome research yet	• Spread of antibiotic resistances • Studying genetic variations in fermented foods	$$$$

[a]It is crucial to distinguish between technical resolution, which relates to the sequencing method itself, and taxonomic resolution, which depends heavily on the quality and comprehensiveness of reference databases. To further clarify the hierarchy of microbial classification, a species can include multiple sub-species, each sub-species may comprise several strains, and a single strain can consist of multiple distinct genomes (14).
[b]+, advantages.
[c]–, drawbacks.

research or specifically to food systems data sets, most likely due to significant barriers imposed by cost, accessibility, and technical challenges.

Overall, higher resolution sequencing technologies can more effectively uncover the within-species variability in similar ecosystems and are particularly valuable for applications such as tracking strain dispersal in food processing environments (66) or identifying pathogenic strains in food safety (67).

Shotgun metagenomics can elucidate the functional potential of strains or communities, guide isolation, reveal associations between metabolic pathways, and find relevant gene clusters

Assembled MAGs from shotgun metagenomics data also offer valuable insights into the functional potential of individual strains and entire microbial communities. From this functional potential, certain phenotypes, such as estimated growth conditions including optimal temperature, oxygen tolerance, salinity, and pH, can be predicted (68). The use of NGS data to guide subsequent cultivation is termed culturomics (69) and has enabled the isolation of previously uncultured microorganisms. For instance, Landis et al. (70) utilized shotgun metagenomic sequencing data to guide the isolation of key strains, revealing that within-species variation among Kombucha strains, despite over 99% genomic similarity, significantly influenced biofilm formation and metabolite production (70).

Additionally, the functional annotations can elucidate metabolic pathways, antimicrobial resistance genes, potentially health-associated gene clusters (PHAGCs), or other biosynthetic gene clusters. For example, Leech et al. (71) applied shotgun sequencing and MAG assembly to 58 different fermented foods, identifying 10 putatively novel strains and assessing antimicrobial potential, such as bacteriocin production and other

PHAGCs (71). Similarly, shotgun sequencing has been applied to assess how packaging conditions of beef select for distinct strains with spoilage potential (72) or to expand our knowledge of how food microbiomes link to the human microbiome and potentially impact health (73).

Correlating characteristics, e.g., volatile organic compound (VOC) production, with the predicted functional potential of bacterial and yeast strains has further been employed to screen dynamics and impact on cheese quality (74). Additionally, it has been used to reconstruct metabolic pathways of flavor compounds in distinct lactic acid bacteria strains found in traditional Mexican Cotija cheese (75). Moreover, improved accuracy of functional annotations also permits the construction of genome-scale metabolic models (GEMs). While currently limited to a few species, multiple GEMs can be integrated to simulate intermicrobial interactions that would be difficult to investigate *in vivo*, as strain-specific differences may confer differential community behaviors (76).

Exploring strain-specific functional differences and responses to specific conditions using metatranscriptomics and metaproteomics

While metagenomics reveals a community's functional potential, metatranscriptomics examines active gene expression and functional response to changed conditions (77). This approach allows us to understand microbial community responses to particular conditions and gain deeper insights into the functional differences between species and strains. Studies on kimchi (78) and cheese rind microbiomes (79) have shown species- and strain-specific gene expression impacting key metabolic pathways, e.g., carbohydrate metabolism and iron utilization. In both studies, assembled MAGs from shotgun metagenomics were used as reference genomes for mRNA transcript mapping and identification of upregulated or downregulated genes as well as entire metabolic pathways. Apart from comparing gene expression to reference genomes, some studies integrate RNA-seq analyses with targeted metabolomics to provide deeper insights into functional dynamics. For example, Bodinaku et al. (80) examined shifts in the metabolic processes of *Penicillium* strains on cheese rinds during domestication and revealed strain-specific metabolic down-regulations at the transcriptional level (80). Similarly, Vicente et al. (81) combined metabolomics and transcriptomics to show how within-species transcriptomic diversity in *Lachancea thermotolerans* drives acidification activity in wine-making (82).

However, metatranscriptomics still faces several challenges, starting with sample collection and RNA preservation, which are critical due to the short half-life of mRNA transcripts (77). The presence of host RNA contamination (e.g., from food substrate matrices) and the requirement for efficient rRNA depletion add further complexity (83). Additionally, the use of short-read sequencing methods for mRNA libraries limits the resolution of complex features such as RNA isoforms, operon structures, and full-length transcripts, which may be critical for achieving strain-level insights (77, 84) and can distort the representation of certain microorganisms and their strain-specific, transient transcriptional patterns. These technical hurdles, coupled with higher costs, lower scalability compared to DNA sequencing protocols, and the immaturity of currently available statistical analysis models (85), still restrict the widespread adoption of metatranscriptomics in large-scale food microbiome surveys.

While transcript levels can be used as a proxy to understand functional responses of communities, they do not always correlate with actual functional changes (86). A more robust indicator is metaproteomics, which confirms the presence of functional proteins, hence providing a direct link to microbial activity, enzyme-related traits, protein-protein interactions, and the discovery of relevant proteins (87, 88). Currently, the lower database coverage compared to genomic databases and the complexity of non-microbial proteins in substrate matrices limit the resolution of metaproteomics in foods (59). Nonetheless, future advances in high-throughput metaproteomics methods may enhance the study of strain-specific enzyme abundance and functional traits, directly impacting the diverse flavors and nutritional values (89–91).

Metabolomics is a key tool for understanding microbes in their ecosystem context

Microbes are influenced by the chemical environment surrounding them, including molecules, salts, and other compounds that affect their behavior, metabolism, and interactions. The chemical milieu shapes within-species diversity and is in turn influenced by the strain-specific functional capabilities of microbes. While metagenomics and metatranscriptomics provide insights into the potential metabolic capabilities of a strain, metabolomics—studying small molecules produced by metabolic processes—offers a direct representation of the phenotype. It is therefore an important tool to understand microbial behavior within the respective ecosystems (92).

Metabolomics is useful to explore the dynamics within microbial communities and interspecies interactions. For example, Blasche et al. (93) show how metabolic cooperation leads to stable coexistence of multiple strains within microbial consortia in milk kefir (93). Moreover, metabolomics can be employed to elucidate the specific function of individual strains in communities. This was shown by Rappaport et al. (4), which combined metabolomics, specifically analysis of VOCs, and genomic characterization of synthetic communities with varying strain composition to demonstrate that the function, i.e. aroma formation, of the entire community can vary depending on within-species genetic differences (4). Inversely to focusing on desired functions, metabolomics is applied to reveal which community members contribute to food defects or spoilage to increase food safety. An example thereof is the study by Liu et al. (94) which applied metabolomics coupled with NGS-guided screening of strains to reduce the formation of biogenic amines in rice wine (94). Alternatively, metabolomics is used to reveal how the chemical composition shapes microbial communities. This was done by Yap et al. (95), which connected environmental and climatic conditions with chemical profiles and metagenomic microbiome analysis in raw milk to reveal the importance of environmental context in shaping the strain composition (95). These studies collectively illustrate the power of metabolomics in providing insights into microbial community dynamics and function, as well as ecosystem interactions. In particular, the integration of metabolomics with genetic and ecological data is central in food systems for revealing strain-specific metabolite profiles, their roles in community metabolism, and adaptations to environmental changes leading to strain stratification (90).

Understanding how viruses influence within-species diversity and the implications for microbiome function and stability

Microbial systems are a rich reservoir for the development and propagation of diverse microorganism-infecting viruses (96). In bacteria, the susceptibility to viruses (bacteriophages) can vary dramatically within a species, and the presence thereof shapes communities by reducing the abundance of certain members or fostering mutualistic relationships through prophage domestication, where a bacteriophage integrates into the host genome (97, 98). Bacteriophages can also protect bacteria via superinfection exclusion, by reducing the reinfection rates of the same or related bacteriophages (99). The selective pressure imposed by constant bacteriophage predation can lead to diversification of and within species through HGT and induce drastic community shifts. This poses a risk to the reliability of industrial starter cultures, e.g., the viral infection of dairy lactic acid bacteria starter cultures is a common cause for failed fermentations (98). Yet, viral selective pressure can also maintain low within-species diversity and act as natural biocontrol, contributing to long-term community stability and preventing the establishment of competitor strains (100, 101). The role of viruses in nutrient redistribution, by releasing organic matter through host cell lysis, is termed "viral shunt" and has been suggested to facilitate succession of microbial populations in food fermentations (102). Fungi, like bacteria, are also susceptible to viral infections (mycoviruses), as demonstrated in *Saccharomyces cerevisiae*, which can be chronically infected with the mycovirus L-A (103). Additionally, certain viruses can occasionally convert specific yeast strains into toxin-producing "killer strains" (104).

Investigating viruses and within-species diversity requires examining bacteriophage-bacterial or mycovirus-fungal coevolution as well as associated clustered regularly interspaced short palindromic repeats (CRISPR) resistance arrays with metagenomics or viromics, along with targeted phage isolation and *in vitro* characterization (105, 106). An example combining all of these approaches is Somerville et al. (101) who explored within-species bacterial diversity and bacteriophage resistance in cheese starter cultures to reveal interspecies functional redundancy alongside significant strain-level variation in virus resistance (101). Furthermore, analyzing CRISPR arrays or isolated viruses from, e.g., production sites may elucidate phylogenies in fermentations down to strain level and deepen our understanding of strain-level dynamics in fermentation processes in the future (102, 107).

CONCLUSION AND OUTLOOK

Characterizing the phenotypes and traits of individual microbial strains has long been recognized as essential, as even closely related microorganisms within the same species can exhibit markedly different functions. Strain variation is particularly evident in food systems, affecting characteristics such as flavor, texture, and safety (9). Moreover, microbial biodiversity is a critical resource for food innovation and human health (108). Unlocking this potential requires a comprehensive understanding of within-species diversity and its implications. Here, we outlined a roadmap integrating emerging technologies with innovative methodologies and present a multifaceted outlook on future technological developments and the opportunities they may offer.

High-resolution community analyses, such as long-read marker-gene sequencing, shotgun metagenomics, or scDNA-seq, are becoming more accessible with advancements in technology and computational power (109). In particular, advances in scDNA-seq hold the potential to revolutionize the study of within-species diversity (65) and could be particularly relevant in assessing genetic heterogeneities within communities, to study, for example, the spread of antibiotic resistances and mechanisms of persistence in spoilage organisms, or the genetic variations in and the dispersion thereof in fermented foods. High-resolution sequencing technologies will be essential to the discovery of novel strains, as well as elucidating evolution and ecology thereof. Therefore, despite remaining challenges, we anticipate these technologies to become more commonly used to investigate fundamental questions in microbiology as well as in practical applications such as tracing contamination routes in food production and authenticating food products (109, 110). Shotgun metagenomics further enables the exploration of the functional potential of individual strains or entire microbial communities and could super-charge strain isolation through culturomics, enabling more precise strain selection for targeted food applications (111) and potentially paving the way for personalized dietary recommendations based on microbial profiles (112, 113).

Advances in bioinformatics and machine learning tools will facilitate predictive modeling to assess food quality and safety based on metagenomic profiles (114, 115), more accurate mapping of transcripts or proteins to strain level-resolved reference databases (75), and prediction of strain-specific functional responses from multi-omics data (116, 117). These tools will aid in distinguishing functional differences between strains, assessing their impact on food quality, identifying key strains and functions responsible for desired flavors and nutritional values, and supporting targeted interventions to enhance quality and safety.

Although metaproteomics is still a nascent field, it is rapidly evolving and promises insights into protein-protein interactions within microbial communities (118). Future advances in protein identification and quantification in complex food matrices will enhance the ability to detect low-abundance proteins and strain-specific enzymes. Ultimately, longitudinal studies could further reveal temporal changes in gene expression and protein abundance, shedding light on community dynamics and strain-specific functional adaptations throughout, for example, the fermentation process.

Similarly, future advancements in metabolomics technologies will enhance the detection of low-abundance metabolites and improve spatial resolution, allowing for the capture of metabolite gradients within microbial communities, which could facilitate the discovery and implementation of distinct biomarkers for food quality and safety (119). Real-time metabolite profiling may also enable the tracking of dynamic changes in microbial ecosystems, offering a future spatio-temporal understanding of communities and their strain-specific contributions. Further multi-omics integrations can link metabolomics to genetic potential, correlate metabolite production with gene expression, and connect metabolomics to metaproteomics to explore the relationship between enzymes and metabolites. Analyzing these multi-omics layers will require advancements in statistical and bioinformatics techniques for interpreting cross-talk between data layers. Supported by artificial intelligence, such methods could enable the prediction of metabolic interactions and community behavior based on genomic or environmental data, thereby identifying key metabolites that drive community dynamics and functionality (120, 121).

Viruses are a frequently overlooked yet crucial component of microbial ecosystems, shaping community composition through strain-specific susceptibilities. Future technological advancements aiming at high-resolution analysis of viral population dynamics within microbiomes will improve the detection, characterization, and isolation of viruses associated with microbial communities. These advancements will also facilitate the screening of interactions with specific strains and the identification of potential biocontrol applications (122). Integration of such strain-specific data will facilitate the modeling of virus-microbe dynamics along with their impact on microbial community stability and (co-)evolution (123). Further integration with metatranscriptomics and metabolomics could help understand microbial gene expression in response to viruses and direct impact on community function. Additionally, CRISPR array analysis has the potential to uncover viral histories and strain-specific resistance mechanisms (102, 107), supporting targeted biocontrol strategies applicable in industrial scales. Collectively, these developments promise to deepen our understanding of the fundamental role of viruses in microbial ecosystems, aiding the design of resilient starter cultures, phage control strategies, and innovative biotechnological applications such as virus-induced diversification of fermented food microbiomes (124).

While the described omics technologies support revealing the origins and implications of within-species diversity, the isolation and *in vitro* testing of strains remains a central technique in microbiology. Physical strain isolates are key to obtaining high-quality assembled genomes and enabling experimental manipulation (125), which can reveal true causal connections and allow the study of interactions within communities (2). Ultimately, to gain a deeper understanding of how microbial communities evolve and to generate practical insights for food producers, experimental manipulations in both synthetic environments and real-world food production settings are essential. These experimental approaches enable the exploration of strain behaviors and interactions under more complex conditions, complementing the insights gained from omics tools.

Notably, these various experimental approaches need to be tailored to the various microbial groups present in fermented foods, such as bacteria, yeast, and molds. Given that the definition of "strain" often varies across research fields, future research must account for these differences and tackle specific challenges, such as capturing fungal proteins and addressing discrepancies in database coverage between kingdoms. Integrating these approaches will ultimately be crucial for comprehensively understanding within-species diversity and its role in inter-kingdom interactions within food systems.

In conclusion, this review highlights the importance of investigating within-species diversity in microbial communities of food systems, as it is necessary for driving deeper insight into the functional roles and interactions of microorganisms within these ecosystems. The integrated application of novel or existing techniques, including paired integration of cultivation-based techniques and omics technologies, can radically

advance our fundamental understanding of microbiomes, while driving significant innovations in food production to enhance quality, safety, and production efficiency.

AUTHOR AFFILIATION

[1]Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

AUTHOR ORCIDs

Lena Flörl http://orcid.org/0000-0003-4458-4071
Annina Meyer http://orcid.org/0009-0006-3110-6598
Nicholas A. Bokulich http://orcid.org/0000-0002-1784-8935

AUTHOR CONTRIBUTIONS

Lena Flörl, Funding acquisition, Writing – original draft, Writing – review and editing | Annina Meyer, Writing – original draft, Writing – review and editing | Nicholas A. Bokulich, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing

DIRECT CONTRIBUTION

This article is a direct contribution from Nicholas A. Bokulich, a member of the *Applied and Environmental Microbiology* Editorial Board, who arranged for and secured reviews by Aaron Walsh, Trinity College Dublin; Benjamin E. Wolfe, Tufts University; and Francesca De Filippis, Universita degli Studi di Napoli Federico II.

REFERENCES

1. Arrigan D, Kothe CI, Oliverio A, Evans JD, Wolfe BE. 2024. Novel fermentations integrate traditional practice and rational design of fermented-food microbiomes. Curr Biol 34:R1094–R1108. https://doi.org/10.1016/j.cub.2024.09.047
2. Niccum BA, Kastman EK, Kfoury N, Robbat A, Wolfe BE. 2020. Strain-level diversity impacts cheese rind microbiome assembly and function. mSystems 5:mSystems–20. https://doi.org/10.1128/mSystems.00149-20
3. Boulton C. 2017. Brewing yeast physiology. In Brewing microbiology: current research, omics and microbial ecology. Caister Academic Press.
4. Rappaport HB, Senewiratne NPJ, Lucas SK, Wolfe BE, Oliverio AM. 2024. Genomics and synthetic community experiments uncover the key metabolic roles of acetic acid bacteria in sourdough starter microbiomes. mSystems 9:e0053724. https://doi.org/10.1128/msystems.00537-24

5. Beck KL, Haiminen N, Chambliss D, Edlund S, Kunitomi M, Huang BC, Kong N, Ganesan B, Baker R, Markwell P, Kawas B, Davis M, Prill RJ, Krishnareddy H, Seabolt E, Marlowe CH, Pierre S, Quintanar A, Parida L, Dubois G, Kaufman J, Weimer BC. 2021. Monitoring the microbiome for food safety and quality using deep shotgun sequencing. NPJ Sci Food 5:3. https://doi.org/10.1038/s41538-020-00083-y

6. Casey MG, Isolini D, Amrein R, Wechsler D, Berthoud H. 2008. Naturally occurring genetic markers in lactobacilli and their use to verify the authenticity of Swiss emmental PDO cheese. Dairy Sci Technol 88:457–466. https://doi.org/10.1051/dst:2008014

7. Fusco V, Fanelli F, Chieffi D. 2023. Recent and advanced DNA-based technologies for the authentication of probiotic, protected designation of origin (PDO) and protected geographical indication (PGI) fermented foods and beverages. Foods 12:3782. https://doi.org/10.3390/foods122 03782

8. Koutsoumanis K, Allende A, Bolton D, Bover-Cid S, Chemaly M, De Cesare A, Herman L, Hilbert F, Lindqvist R, Nauta M, et al. 2024. Persistence of microbiological hazards in food and feed production and processing environments. EFSA J 22:e8521. https://doi.org/10.2903/j.efs a.2024.8521

9. Wolfe BE, Dutton RJ. 2015. Fermented foods as experimentally tractable microbial ecosystems. Cell 161:49–55. https://doi.org/10.1016/j.cell.201 5.02.034

10. Todd ECD, Greig JD, Bartleson CA, Michaels BS. 2008. Outbreaks where food workers have been implicated in the spread of foodborne disease. Part 5. Sources of contamination and pathogen excretion from infected persons. J Food Prot 71:2582–2595. https://doi.org/10.4315/0362-028x-71.12.2582

11. Reese AT, Madden AA, Joossens M, Lacaze G, Dunn RR. 2020. Influences of ingredients and bakers on the bacteria and fungi in sourdough starters and bread. mSphere 5:e00950-19. https://doi.org/10.1128/mSp here.00950-19

12. Louw NL, Lele K, Ye R, Edwards CB, Wolfe BE. 2023. Microbiome assembly in fermented foods. Annu Rev Microbiol 77:381–402. https://d oi.org/10.1146/annurev-micro-032521-041956

13. Fuchsman CA, Collins RE, Rocap G, Brazelton WJ. 2017. Effect of the environment on horizontal gene transfer between bacteria and archaea. PeerJ 5:e3865. https://doi.org/10.7717/peerj.3865

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

15. Rodriguez-Valera F, Martín-Cuadrado A-B, Rodriguez-Brito B, Pasic L, Thingstad TF, Rohwer F, Mira A. 2009. Explaining microbial population genomics through phage predation. Nat Prec. https://doi.org/10.1038/ npre.2009.3489.1

16. Pham VD, Simpson DJ, Gänzle MG. 2025. Strain-level identification of bacteria persisting in food and in food processing facilities: when do two isolates represent the same strain and which tools identify a strain? Curr Opin Food Sci 61:101245. https://doi.org/10.1016/j.cofs.2024.1012 45

17. Frantzen CA, Kleppen HP, Holo H. 2018. Lactococcus lactis diversity in undefined mixed dairy starter cultures as revealed by comparative genome analyses and targeted amplicon sequencing of epsD. Appl Environ Microbiol 84:e02199-17. https://doi.org/10.1128/AEM.02199-17

18. Legras JL, Merdinoglu D, Cornuet JM, Karst F. 2007. Bread, beer and wine: Saccharomyces cerevisiae diversity reflects human history. Mol Ecol 16:2091–2102. https://doi.org/10.1111/j.1365-294X.2007.03266.x

19. Carreto L, Eiriz MF, Domingues I, Schuller D, Moura GR, Santos MAS. 2011. Expression variability of co-regulated genes differentiates Saccharomyces cerevisiae strains. BMC Genomics 12:201. https://doi.org /10.1186/1471-2164-12-201

20. Hernández A, Pérez-Nevado F, Ruiz-Moyano S, Serradilla MJ, Villalobos MC, Martín A, Córdoba MG. 2018. Spoilage yeasts: what are the sources of contamination of foods and beverages? Int J Food Microbiol 286:98–110. https://doi.org/10.1016/j.ijfoodmicro.2018.07.031

21. Riquelme AJ, Calvo MA, Guzmán AM, Depix MS, García P, Pérez C, Arrese M, Labarca JA. 2003. Saccharomyces cerevisiae fungemia after Saccharomyces boulardii treatment in immunocompromised patients. J Clin Gastroenterol 36:41–43. https://doi.org/10.1097/00004836-200301 000-00013

22. Didari T, Solki S, Mozaffari S, Nikfar S, Abdollahi M. 2014. A systematic review of the safety of probiotics. Expert Opin Drug Saf 13:227–239. htt ps://doi.org/10.1517/14740338.2014.872627

23. Dumas E, Feurtey A, Rodríguez de la Vega RC, Le Prieur S, Snirc A, Coton M, Thierry A, Coton E, Le Piver M, Roueyre D, Ropars J, Branca A, Giraud T. 2020. Independent domestication events in the blue-cheese fungus Penicillium roqueforti. Mol Ecol 29:2639–2660. https://doi.org/10.1111/ mec.15359

24. Coton E, Coton M, Hymery N, Mounier J, Jany JL. 2020. Penicillium roqueforti: an overview of its genetics, physiology, metabolism and biotechnological applications. Fungal Biol Rev 34:59–73. https://doi.org /10.1016/j.fbr.2020.03.001

25. Douglass AP, Offei B, Braun-Galleani S, Coughlan AY, Martos AAR, Ortiz-Merino RA, Byrne KP, Wolfe KH. 2018. Population genomics shows no distinction between pathogenic Candida krusei and environmental Pichia kudriavzevii: one species, four names. PLOS Pathog 14:e1007138. https://doi.org/10.1371/journal.ppat.1007138

26. Chu Y, Li M, Jin J, Dong X, Xu K, Jin L, Qiao Y, Ji H. 2023. Advances in the application of the non-conventional yeast Pichia kudriavzevii in food and biotechnology industries. J Fungi (Basel) 9:170. https://doi.org/10.3 390/jof9020170

27. Breuer U, Harms H. 2006. Debaryomyces hansenii--an extremophilic yeast with biotechnological potential. Yeast 23:415–437. https://doi.org /10.1002/yea.1374

28. Albertin W, Setati ME, Miot-Sertier C, Mostert TT, Colonna-Ceccaldi B, Coulon J, Girard P, Moine V, Pillet M, Salin F, Bely M, Divol B, Masneuf-Pomarede I. 2016. Hanseniaspora uvarum from winemaking environ-ments show spatial and temporal genetic clustering. Front Microbiol 6. https://doi.org/10.3389/fmicb.2015.01569

29. Iijima K, Suzuki K, Ozaki K, Yamashita H. 2006. horC confers beer-spoilage ability on hop-sensitive Lactobacillus brevis ABBC45cc. J Appl Microbiol 100:1282–1288. https://doi.org/10.1111/j.1365-2672.2006.02 869.x

30. Yilmaz B, Bangar SP, Echegaray N, Suri S, Tomasevic I, Manuel Lorenzo J, Melekoglu E, Rocha JM, Ozogul F. 2022. The impacts of Lactiplantibacil-lus plantarum on the functional properties of fermented foods: a review of current knowledge. Microorganisms 10:826. https://doi.org/10.3390/ microorganisms10040826

31. Bartowsky EJ, Borneman AR. 2011. Genomic variations of Oenococcus oeni strains and the potential to impact on malolactic fermentation and aroma compounds in wine. Appl Microbiol Biotechnol 92:441–447. http s://doi.org/10.1007/s00253-011-3546-2

32. Ruiz P, Izquierdo PM, Seseña S, Palop ML. 2010. Selection of autochtho-nous Oenococcus oeni strains according to their oenological properties and vinification results. Int J Food Microbiol 137:230–235. https://doi.or g/10.1016/j.ijfoodmicro.2009.11.027

33. El Khoury M, Campbell-Sills H, Salin F, Guichoux E, Claisse O, Lucas PM. 2017. Biogeography of Oenococcus oeni reveals distinctive but nonspecific populations in wine-producing regions. Appl Environ Microbiol 83:e02322-16. https://doi.org/10.1128/AEM.02322-16

34. Rönkä E, Malinen E, Saarela M, Rinta-Koski M, Aarnikunnas J, Palva A. 2003. Probiotic and milk technological properties of Lactobacillus brevis. Int J Food Microbiol 83:63–74. https://doi.org/10.1016/s0168-160 5(02)00315-x

35. Fang F, Xu J, Li Q, Xia X, Du G. 2018. Characterization of a Lactobacillus brevis strain with potential oral probiotic properties. BMC Microbiol 18:221. https://doi.org/10.1186/s12866-018-1369-3

36. Giraffa G, Gatti M, Rossetti L, Senini L, Neviani E. 2000. Molecular diversity within Lactobacillus helveticus as revealed by genotypic characterization. Appl Environ Microbiol 66:1259–1265. https://doi.org/ 10.1128/AEM.66.4.1259-1265.2000

37. Kido Y, Maeno S, Tanno H, Kichise Y, Shiwa Y, Endo A. 2021. Niche-specific adaptation of Lactobacillus helveticus strains isolated from malt whisky and dairy fermentations. Microb Genom 7:000560. https://doi.or g/10.1099/mgen.0.000560

38. Cachat E, Priest FG. 2005. Lactobacillus suntoryeus sp. nov., isolated from malt whisky distilleries. Int J Syst Evol Microbiol 55:31–34. https://doi.or g/10.1099/ijs.0.63266-0

39. Cavanagh D, Fitzgerald GF, McAuliffe O. 2015. From field to fermenta-tion: the origins of Lactococcus lactis and its domestication to the dairy environment. Food Microbiol 47:45–61. https://doi.org/10.1016/j.fm.20 14.11.001

40. Roux E, Nicolas A, Valence F, Siekaniec G, Chuat V, Nicolas J, Le Loir Y, Guédon E. 2022. The genomic basis of the Streptococcus thermophilus health-promoting properties. BMC Genomics 23:210. https://doi.org/10 .1186/s12864-022-08459-y

41. Michaylova M, Minkova S, Kimura K, Sasaki T, Isawa K. 2007. Isolation and characterization of *Lactobacillus delbrueckii* ssp. bulgaricus and *Streptococcus thermophilus* from plants in Bulgaria. FEMS Microbiol Lett 269:160–169. https://doi.org/10.1111/j.1574-6968.2007.00631.x

42. Bokulich NA, Bamforth CW. 2013. The microbiology of malting and brewing. Microbiol Mol Biol Rev 77:157–172. https://doi.org/10.1128/MMBR.00060-12

43. Blaya J, Barzideh Z, LaPointe G. 2018. Symposium review: interaction of starter cultures and nonstarter lactic acid bacteria in the cheese environment. J Dairy Sci 101:3611–3629. https://doi.org/10.3168/jds.2017-13345

44. Gill A. 2017. The importance of bacterial culture to food microbiology in the age of genomics. Front Microbiol 8:777. https://doi.org/10.3389/fmicb.2017.00777

45. Zotta T, Ricciardi A, Condelli N, Parente E. 2022. Metataxonomic and metagenomic approaches for the study of undefined strain starters for cheese manufacture. Crit Rev Food Sci Nutr 62:3898–3912. https://doi.org/10.1080/10408398.2020.1870927

46. Heyse J, Buysschaert B, Props R, Rubbens P, Skirtach AG, Waegeman W, Boon N. 2019. Coculturing bacteria leads to reduced phenotypic heterogeneities. Appl Environ Microbiol 85:e02814-18. https://doi.org/10.1128/AEM.02814-18

47. da Silva Duarte V, Carlot M, Pakroo S, Tarrah A, Lombardi A, Santiago H, Corich V, Giacomini A. 2020. Comparative evaluation of cheese whey microbial composition from four Italian cheese factories by viable counts and 16S rRNA gene amplicon sequencing. Int Dairy J 104:104656. https://doi.org/10.1016/j.idairyj.2020.104656

48. Mendoza LM, Neef A, Vignolo G, Belloch C. 2017. Yeast diversity during the fermentation of Andean chicha: a comparison of high-throughput sequencing and culture-dependent approaches. Food Microbiol 67:1–10. https://doi.org/10.1016/j.fm.2017.05.007

49. Bokulich NA, Lewis ZT, Boundy-Mills K, Mills DA. 2016. A new perspective on microbial landscapes within food production. Curr Opin Biotechnol 37:182–189. https://doi.org/10.1016/j.copbio.2015.12.008

50. Parente E, Ricciardi A, Zotta T. 2020. The microbiota of dairy milk: a review. Int Dairy J 107:104714. https://doi.org/10.1016/j.idairyj.2020.104714

51. Johnson JS, Spakowicz DJ, Hong B-Y, Petersen LM, Demkowicz P, Chen L, Leopold SR, Hanson BM, Agresta HO, Gerstein M, Sodergren E, Weinstock GM. 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun 10:5029. https://doi.org/10.1038/s41467-019-13036-1

52. Illumina Knowledge. 2024. Maximum read length for illumina sequencing platforms. Available from: https://knowledge.illumina.com/instrumentation/general/instrumentation-general-reference_material-list/000002826

53. PacBio GM. 2024. Sequencing 101: comparing long-read sequencing technologies. Available from: https://www.pacb.com/blog/sequencing-101-comparing-long-read-sequencing-technologies/

54. Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. 2021. Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol 39:1348–1365. https://doi.org/10.1038/s41587-021-01108-x

55. Callahan BJ, Grinevich D, Thakur S, Balamotis MA, Yehezkel TB. 2021. Ultra-accurate microbial amplicon sequencing with synthetic long reads. Microbiome 9:130. https://doi.org/10.1186/s40168-021-01072-3

56. Abellan-Schneyder I, Matchado MS, Reitmeier S, Sommer A, Sewald Z, Baumbach J, List M, Neuhaus K. 2021. Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. mSphere 6:e01202-20. https://doi.org/10.1128/mSphere.01202-20

57. Tedersoo L, Anslan S, Bahram M, Kõljalg U, Abarenkov K. 2020. Identifying the "unidentified" fungi: a global-scale long-read third-generation sequencing approach. Fungal Divers 103:273–293. https://doi.org/10.1007/s13225-020-00456-4

58. Azinheiro S, Roumani F, Carvalho J, Prado M, Garrido-Maestu A. 2021. Suitability of the MinION long read sequencer for semi-targeted detection of foodborne pathogens. Anal Chim Acta 1184:339051. https://doi.org/10.1016/j.aca.2021.339051

59. Walsh AM, Leech J, Huttenhower C, Delhomme-Nguyen H, Crispie F, Chervaux C, Cotter PD. 2023. Integrated molecular approaches for fermented food microbiome research. FEMS Microbiol Rev 47:fuad001. https://doi.org/10.1093/femsre/fuad001

60. Zhou Y, Liu M, Yang J. 2022. Recovering metagenome-assembled genomes from shotgun metagenomic sequencing data: Methods, applications, challenges, and opportunities. Microbiol Res 260:127023. https://doi.org/10.1016/j.micres.2022.127023

61. Zheng W, Zhao S, Yin Y, Zhang H, Needham DM, Evans ED, Dai CL, Lu PJ, Alm EJ, Weitz DA. 2022. High-throughput, single-microbe genomics with strain resolution, applied to a human gut microbiome. Science 376:eabm1483. https://doi.org/10.1126/science.abm1483

62. Truong DT, Tett A, Pasolli E, Huttenhower C, Segata N. 2017. Microbial strain-level population structure and genetic diversity from metagenomes. Genome Res 27:626–638. https://doi.org/10.1101/gr.216242.116

63. Chen L, Zhao N, Cao J, Liu X, Xu J, Ma Y, Yu Y, Zhang X, Zhang W, Guan X, Yu X, Liu Z, Fan Y, Wang Y, Liang F, Wang D, Zhao L, Song M, Wang J. 2022. Short- and long-read metagenomics expand individualized structural variations in gut microbiomes. Nat Commun 13:3175. https://doi.org/10.1038/s41467-022-30857-9

64. Evrony GD, Hinch AG, Luo C. 2021. Applications of single-Cell DNA sequencing. Annu Rev Genomics Hum Genet 22:171–197. https://doi.org/10.1146/annurev-genom-111320-090436

65. Lloréns-Rico V, Simcock JA, Huys GRB, Raes J. 2022. Single-cell approaches in human microbiome research. Cell 185:2725–2738. https://doi.org/10.1016/j.cell.2022.06.040

66. Doyle CJ, O'Toole PW, Cotter PD. 2017. Metagenome-based surveillance and diagnostic approaches to studying the microbial ecology of food production and processing environments. Environ Microbiol 19:4382–4391. https://doi.org/10.1111/1462-2920.13859

67. Maguire M, Kase JA, Roberson D, Muruvanda T, Brown EW, Allard M, Musser SM, González-Escalona N. 2021. Precision long-read metagenomics sequencing for food safety by detection and assembly of Shiga toxin-producing *Escherichia coli* in irrigation water. PLoS One 16:e0245172. https://doi.org/10.1371/journal.pone.0245172

68. Barnum TP, Crits-Christoph A, Molla M, Carini P, Lee HH, Ostrov N. 2024. Predicting microbial growth conditions from amino acid composition. Microbiology. https://doi.org/10.1101/2024.03.22.586313

69. Lagier J-C, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, Levasseur A, Rolain J-M, Fournier P-E, Raoult D. 2018. Culturing the human microbiota and culturomics. Nat Rev Microbiol 16:540–550. https://doi.org/10.1038/s41579-018-0041-0

70. Landis EA, Fogarty E, Edwards JC, Popa O, Eren AM, Wolfe BE. 2022. Microbial diversity and interaction specificity in kombucha tea fermentations. mSystems 7:e0015722. https://doi.org/10.1128/msystems.00157-22

71. Leech J, Cabrera-Rubio R, Walsh AM, Macori G, Walsh CJ, Barton W, Finnegan L, Crispie F, O'Sullivan O, Claesson MJ, Cotter PD. 2020. Fermented-food metagenomics reveals substrate-associated differences in taxonomy and health-associated and antibiotic resistance determinants. mSystems 5:e00522-20. https://doi.org/10.1128/mSystems.00522-20

72. De Filippis F, La Storia A, Villani F, Ercolini D. 2019. Strain-level diversity analysis of pseudomonas fragi after *in situ* pangenome reconstruction shows distinctive spoilage-associated metabolic traits clearly selected by different storage conditions. Appl Environ Microbiol 85:e02212-18. https://doi.org/10.1128/AEM.02212-18

73. Carlino N, Blanco-Míguez A, Punčochář M, Mengoni C, Pinto F, Tatti A, Manghi P, Armanini F, Avagliano M, Barcenilla C, et al. 2024. Unexplored microbial diversity from 2,500 food metagenomes and links with the human microbiome. Cell 187:5775–5795. https://doi.org/10.1016/j.cell.2024.07.039

74. Bertuzzi AS, Walsh AM, Sheehan JJ, Cotter PD, Crispie F, McSweeney PLH, Kilcawley KN, Rea MC. 2018. Omics-based insights into flavor development and microbial succession within surface-ripened cheese. mSystems 3:e00211-17. https://doi.org/10.1128/mSystems.00211-17

75. Escobar-Zepeda A, Sanchez-Flores A, Quirasco Baruch M. 2016. Metagenomic analysis of a Mexican ripened cheese reveals a unique complex microbiota. Food Microbiol 57:116–127. https://doi.org/10.1016/j.fm.2016.02.004

76. Gu C, Kim GB, Kim WJ, Kim HU, Lee SY. 2019. Current status and applications of genome-scale metabolic models. Genome Biol 20:121. https://doi.org/10.1186/s13059-019-1730-3

77. Shakya M, Lo CC, Chain PSG. 2019. Advances and challenges in metatranscriptomic analysis. Front Genet 10:904. https://doi.org/10.3389/fgene.2019.00904

78. Jung JY, Lee SH, Jin HM, Hahn Y, Madsen EL, Jeon CO. 2013. Metatranscriptomic analysis of lactic acid bacterial gene expression during

kimchi fermentation. Int J Food Microbiol 163:171–179. https://doi.org/10.1016/j.ijfoodmicro.2013.02.022

79. Kastman EK, Kamelamela N, Norville JW, Cosetta CM, Dutton RJ, Wolfe BE. 2016. Biotic interactions shape the ecological distributions of *Staphylococcus* species. MBio 7:e01157-16. https://doi.org/10.1128/mBio.01157-16

80. Bodinaku I, Shaffer J, Connors AB, Steenwyk JL, Biango-Daniels MN, Kastman EK, Rokas A, Robbat A, Wolfe BE. 2019. Rapid phenotypic and metabolomic domestication of wild *Penicillium* molds on cheese. MBio 10:e02445-19. https://doi.org/10.1128/mBio.02445-19

81. VicenteJ, BenitoS, MarquinaD, SantosA. 2025. Subpopulation-specific gene expression in *Lachancea thermotolerans* uncovers distinct metabolic adaptations to wine fermentation. Current research in food science 10:100954. https://doi.org/10.1016/j.crfs.2024.100954

82. Vicente J, Ruiz J, Tomasi S, de Celis M, Ruiz-de-Villa C, Gombau J, Rozès N, Zamora F, Santos A, Marquina D, Belda I. 2023. Impact of rare yeasts in *Saccharomyces cerevisiae* wine fermentation performance: population prevalence and growth phenotype of *Cyberlindnera fabianii*, *Kazachstania unispora*, and *Naganishia globosa*. Food Microbiol 110:104189. https://doi.org/10.1016/j.fm.2022.104189

83. Vannier N, Mesny F, Getzke F, Chesneau G, Dethier L, Ordon J, Thiergart T, Hacquard S. 2023. Genome-resolved metatranscriptomics reveals conserved root colonization determinants in a synthetic microbiota. Nat Commun 14:8274. https://doi.org/10.1038/s41467-023-43688-z

84. Pollard MO, Gurdasani D, Mentzer AJ, Porter T, Sandhu MS. 2018. Long reads: their purpose and place. Hum Mol Genet 27:R234–R241. https://doi.org/10.1093/hmg/ddy177

85. Cho H, Qu Y, Liu C, Tang B, Lyu R, Lin BM, Roach J, Azcarate-Peril MA, Aguiar Ribeiro A, Love MI, Divaris K, Wu D. 2023. Comprehensive evaluation of methods for differential expression analysis of metatranscriptomics data. Brief Bioinform 24:bbad279. https://doi.org/10.1093/bib/bbad279

86. Rocca JD, Hall EK, Lennon JT, Evans SE, Waldrop MP, Cotner JB, Nemergut DR, Graham EB, Wallenstein MD. 2015. Relationships between protein-encoding gene abundance and corresponding process are commonly assumed and rarely observed. ISME J 9:1693–1699. https://doi.org/10.1038/ismej.2014.252

87. Shi H, An F, Lin H, Li M, Wu J, Wu R. 2022. Advances in fermented foods revealed by multi-omics: a new direction toward precisely clarifying the roles of microorganisms. Front Microbiol 13:1044820. https://doi.org/10.3389/fmicb.2022.1044820

88. Zhang K, Delgado-Baquerizo M, Zhu YG, Chu H. 2020. Space is more important than season when shaping soil microbial communities at a large spatial scale. mSystems 5:msystems. https://doi.org/10.1128/msystems.00783-19

89. Li L, Wang T, Ning Z, Zhang X, Butcher J, Serrana JM, Simopoulos CMA, Mayne J, Stintzi A, Mack DR, Liu Y-Y, Figeys D. 2023. Revealing proteome-level functional redundancy in the human gut microbiome using ultra-deep metaproteomics. Nat Commun 14:3428. https://doi.org/10.1038/s41467-023-39149-2

90. Singer F, Kuhring M, Renard BY, Muth T. 2025. Moving toward metaproteogenomics: a computational perspective on analyzing microbial samples via proteogenomics, p 297–318. In Allmer J, Kumar A (ed), Proteogenomics: methods and protocols. Springer US, New York, NY. https://doi.org/10.1007/978-1-0716-4152-1_17.

91. Zhang L, Bao Y, Chen H, Huang J, Xu Y. 2020. Functional microbiota for polypeptide degradation during hypertonic moromi-fermentation of pixian broad bean paste. Foods 9:930. https://doi.org/10.3390/foods9070930

92. Bauermeister A, Mannochio-Russo H, Costa-Lotufo LV, Jarmusch AK, Dorrestein PC. 2022. Mass spectrometry-based metabolomics in microbiome investigations. Nat Rev Microbiol 20:143–160. https://doi.org/10.1038/s41579-021-00621-9

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

94. Liu SP, Yu JX, Wei XL, Ji ZW, Zhou ZL, Meng XY, Mao J. 2016. Sequencing-based screening of functional microorganism to decrease the formation of biogenic amines in Chinese rice wine. Food Control 64:98–104. https://doi.org/10.1016/j.foodcont.2015.12.013

95. Yap M, O'Sullivan O, O'Toole PW, Sheehan JJ, Fenelon MA, Cotter PD. 2024. Seasonal and geographical impact on the Irish raw milk microbiota correlates with chemical composition and climatic variables. mSystems 9:e01290–23. https://doi.org/10.1128/msystems.01290-23

96. Maske BL, de Melo Pereira GV, da Silva Vale A, Marques Souza DS, De Dea Lindner J, Soccol CR. 2021. Viruses in fermented foods: are they good or bad? two sides of the same coin. Food Microbiol 98:103794. https://doi.org/10.1016/j.fm.2021.103794

97. Olszak T, Latka A, Roszniowski B, Valvano MA, Drulis-Kawa Z. 2017. Phage life cycles behind bacterial biodiversity. Curr Med Chem 24:3987–4001. https://doi.org/10.2174/0929867324666170413100136

98. White K, Yu J-H, Eraclio G, Dal Bello F, Nauta A, Mahony J, van Sinderen D. 2022. Bacteriophage-host interactions as a platform to establish the role of phages in modulating the microbial composition of fermented foods. Microbiome Res Rep 1:3. https://doi.org/10.20517/mrr.2021.04

99. Folimonova SY. 2012. Superinfection exclusion is an active virus-controlled function that requires a specific viral protein. J Virol 86:5554–5561. https://doi.org/10.1128/JVl.00310-12

100. Cristobal-Cueto P, García-Quintanilla A, Esteban J, García-Quintanilla M. 2021. Phages in food industry biocontrol and bioremediation. Antibiotics (Basel) 10:786. https://doi.org/10.3390/antibiotics10070786

101. Somerville V, Berthoud H, Schmidt RS, Bachmann H-P, Meng YH, Fuchsmann P, von Ah U, Engel P. 2022. Functional strain redundancy and persistent phage infection in Swiss hard cheese starter cultures. ISME J 16:388–399. https://doi.org/10.1038/s41396-021-01071-0

102. Zhang H, Zhang H, Du H, Yu X, Xu Y. 2025. The insights into the phage communities of fermented foods in the age of viral metagenomics. Crit Rev Food Sci Nutr 65:1656–1668. https://doi.org/10.1080/10408398.2023.2299323

103. Chau S, Gao J, Diao AJ, Cao SB, Azhieh A, Davidson AR, Meneghini MD. 2023. Diverse yeast antiviral systems prevent lethal pathogenesis caused by the L-A mycovirus. Proc Natl Acad Sci U S A 120:e2208695120. https://doi.org/10.1073/pnas.2208695120

104. Rodríguez-Cousiño N, Maqueda M, Ambrona J, Zamora E, Esteban R, Ramírez M. 2011. A new wine *Saccharomyces cerevisiae* killer toxin (Klus), encoded by a double-stranded RNA virus, with broad antifungal activity is evolutionarily related to a chromosomal host gene. Appl Environ Microbiol 77:1822–1832. https://doi.org/10.1128/AEM.02501-10

105. Mahony J, Moscarelli A, Kelleher P, Lugli GA, Ventura M, Settanni L, van Sinderen D. 2017. Phage biodiversity in artisanal cheese wheys reflects the complexity of the fermentation process. Viruses 9:45. https://doi.org/10.3390/v9030045

106. White K, Eraclio G, Lugli GA, Ventura M, Mahony J, Bello FD. 2024. A metagenomics approach to enumerate bacteriophages in A food niche, p 185–199. In Tumban E (ed), Bacteriophages: methods and protocols. Springer US, New York, NY.

107. Collins AJ, Whitaker RJ. 2023. CRISPR comparison toolkit: rapid identification, visualization, and analysis of CRISPR array diversity. CRISPR J 6:386–400. https://doi.org/10.1089/crispr.2022.0080

108. Hernández-Velázquez R, Flörl L, Lavrinienko A, Sebechlebská Z, Merk L, Greppi A, Bokulich NA. 2024. The future is fermented: microbial biodiversity of fermented foods is a critical resource for food innovation and human health. Trends Food Sci Technol 150:104569. https://doi.org/10.1016/j.tifs.2024.104569

109. Billington C, Kingsbury JM, Rivas L. 2022. Metagenomics approaches for improving food safety: a review. J Food Prot 85:448–464. https://doi.org/10.4315/JFP-21-301

110. Pham VD, Xu ZS, Simpson DJ, Zhang JS, Gänzle MG. 2024. Does strain-level persistence of lactobacilli in long-term back-slopped sourdoughs inform on domestication of food-fermenting lactic acid bacteria? Appl Environ Microbiol 90:e0189224. https://doi.org/10.1128/aem.01892-24

111. Liu S, Moon CD, Zheng N, Huws S, Zhao S, Wang J. 2022. Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. Microbiome 10:76. https://doi.org/10.1186/s40168-022-01272-5

112. Kolodziej MC, Singla J, Sánchez-Martín J, Zbinden H, Šimková H, Karafiátová M, Doležel J, Gronnier J, Poretti M, Glauser G, Zhu W, Köster P, Zipfel C, Wicker T, Krattinger SG, Keller B. 2021. A membrane-bound ankyrin repeat protein confers race-specific leaf rust disease resistance in wheat. Nat Commun 12:956. https://doi.org/10.1038/s41467-020-20777-x

113. Leshem A, Segal E, Elinav E. 2020. The gut microbiome and individual-specific responses to diet. mSystems 5:00665–20. https://doi.org/10.1128/mSystems.00665-20

114. Mukherjee A, Breselge S, Dimidi E, Marco ML, Cotter PD. 2024. Fermented foods and gastrointestinal health: underlying mechanisms. Nat Rev Gastroenterol Hepatol 21:248–266. https://doi.org/10.1038/s41575-023-00869-x

115. Suárez N, Weckx S, Minahk C, Hebert EM, Saavedra L. 2020. Metagenomics-based approach for studying and selecting bioprotective strains from the bacterial community of artisanal cheeses. Int J Food Microbiol 335:108894. https://doi.org/10.1016/j.ijfoodmicro.2020.108894

116. Lambert BS, Groussman RD, Schatz MJ, Coesel SN, Durham BP, Alverson AJ, White AE, Armbrust EV. 2022. The dynamic trophic architecture of open-ocean protist communities revealed through machine-guided metatranscriptomics. Proc Natl Acad Sci U S A 119:e2100916119. https://doi.org/10.1073/pnas.2100916119

117. Mallick H, Ma S, Franzosa EA, Vatanen T, Morgan XC, Huttenhower C. 2017. Experimental design and quantitative analysis of microbial community multiomics. Genome Biol 18:228. https://doi.org/10.1186/s13059-017-1359-z

118. Yang L, Fan W, Xu Y. 2020. Metaproteomics insights into traditional fermented foods and beverages. Compr Rev Food Sci Food Saf 19:2506–2529. https://doi.org/10.1111/1541-4337.12601

119. Mialon N, Roig B, Capodanno E, Cadiere A. 2023. Untargeted metabolomic approaches in food authenticity: a review that showcases biomarkers. Food Chem 398:133856. https://doi.org/10.1016/j.foodchem.2022.133856

120. Coler EA, Chen W, Melnik AV, Morton JT, Aksenov AA. 2024. Metabolomics in the era of artificial intelligence. Microbiota and Host 2. https://doi.org/10.1530/MAH-23-0017

121. Ferrocino I, Rantsiou K, McClure R, Kostic T, de Souza RSC, Lange L, FitzGerald J, Kriaa A, Cotter P, Maguin E, Schelkle B, Schloter M, Berg G, Sessitsch A, Cocolin L, MicrobiomeSupport Consortium. 2023. The need for an integrated multi-OMICs approach in microbiome science in the food system. Compr Rev Food Sci Food Saf 22:1082–1103. https://doi.org/10.1111/1541-4337.13103

122. Wagemans J, Holtappels D, Vainio E, Rabiey M, Marzachì C, Herrero S, Ravanbakhsh M, Tebbe CC, Ogliastro M, Ayllón MA, Turina M. 2022. Going viral: virus-based biological control agents for plant protection. Annu Rev Phytopathol 60:21–42. https://doi.org/10.1146/annurev-phyto-021621-114208

123. Koskella B, Hernandez CA, Wheatley RM. 2022. Understanding the impacts of bacteriophage viruses: from laboratory evolution to natural ecosystems. Annu Rev Virol 9:57–78. https://doi.org/10.1146/annurev-virology-091919-075914

124. Mahony J, van Sinderen D. 2022. Virome studies of food production systems: time for "farm to fork" analyses. Curr Opin Biotechnol 73:22–27. https://doi.org/10.1016/j.copbio.2021.06.014

125. Anderson BD, Bisanz JE. 2023. Challenges and opportunities of strain diversity in gut microbiome research. Front Microbiol 14:1117122. https://doi.org/10.3389/fmicb.2023.1117122

AUTHOR BIOS

Lena Flörl holds a Bachelor's and Master's degree in Food Science and Biotechnology from the University of Natural Resources and Life Sciences, Vienna (Austria). After a research stay at Lawrence Berkeley National Laboratory in California, she began her PhD at ETH Zurich (Switzerland), where she explores microbial ecology in vineyards and traditional fermented foods. Passionate about interdisciplinary research, she studies what shapes microbial communities and how they contribute to the unique characteristics of fermented foods – especially when fungi are involved. Her work integrates bioinformatics, microbiology, food science, and ecology to uncover the hidden dynamics of these systems. When she's not studying microbes, she's probably talking about them or tasting the results of their hard work.

Annina Meyer holds both a Bachelor's and a Master's degree in Microbiology and Immunology from ETH Zurich, Switzerland. After formative research internships at the Institut Pasteur in Paris, France and Novartis in Basel, Switzerland, she returned to ETH Zurich to focus on microbiome research, mapping and characterizing the microbial diversity of spontaneous food and beverage fermentations from across Europe. Her work actively involves citizen scientists through home experimentations, sample collection, and knowledge exchange. By integrating newly generated multi-omics data with public datasets, she aims to globally contextualize fermentation microbiomes and better understand the ecological and evolutionary drivers shaping biodiversity in fermentations. Her primary interests lie in applying computational approaches to study microbial diversity, stability, and biogeography in complex ecosystems, with a focus on how microbial communities persist and adapt across environments.

Nicholas A. Bokulich is an Assistant Professor at ETH Zurich, Switzerland, and the head of the Laboratory of Food Systems Biotechnology. He earned his PhD at the University of California, Davis, examining the microbial ecology and biogeography of winegrowing and food fermentations. After completing postdoctoral research stays at New York University Medical Center and Northern Arizona University, he started his group at ETH Zurich in 2020. He and his team develop and use computational and experimental "omics" approaches for studying microbial ecosystems at the interface of foods and human health, with the aim of establishing a predictive understanding of how complex communities develop and interact with each other and their host environment.