



## OPEN ACCESS

## EDITED BY

Pablo Jarrin,  
National Institute of Biodiversity, Ecuador

## REVIEWED BY

Ranjan Kumar Mohapatra,  
Chungnam National University, Republic of  
Korea  
Satya Prakash,  
University of Warwick, United Kingdom

## \*CORRESPONDENCE

Juan Manuel Cevallos-Cevallos,  
✉ jmceva@espol.edu.ec

RECEIVED 31 May 2025

ACCEPTED 05 August 2025

PUBLISHED 20 August 2025

## CITATION

Tigrero-Vaca J, Díaz B, Gu G and  
Cevallos-Cevallos JM (2025) Next-generation  
sequencing applications in food science:  
fundamentals and recent advances.  
*Front. Bioeng. Biotechnol.* 13:1638957.  
doi: 10.3389/fbioe.2025.1638957

## COPYRIGHT

© 2025 Tigrero-Vaca, Díaz, Gu and Cevallos-  
Cevallos. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# Next-generation sequencing applications in food science: fundamentals and recent advances

Joel Tigrero-Vaca<sup>1</sup>, Byron Díaz<sup>1</sup>, Ganyu Gu<sup>2</sup> and  
Juan Manuel Cevallos-Cevallos<sup>1\*</sup>

<sup>1</sup>Escuela Superior Politécnica del Litoral, ESPOL, Centro de Investigaciones Biotecnológicas del Ecuador (CIBE), Guayaquil, Ecuador, <sup>2</sup>Environmental Microbial and Food Safety Laboratory, USDA ARS, Beltsville, MD, United States

Next-generation sequencing (NGS) has revolutionized food science, offering unprecedented insights into microbial communities, food safety, fermentation, and product authenticity. NGS techniques, including metagenetics, metagenomics, and metatranscriptomics, enable culture-independent pathogen detection, antimicrobial resistance surveillance, and detailed microbial profiling, significantly improving food safety monitoring and outbreak prevention. In food fermentation, NGS has enhanced our understanding of microbial interactions, flavor formation, and metabolic pathways, contributing to optimized starter cultures and improved product quality. Furthermore, NGS has become a valuable tool in food authentication and traceability, ensuring product integrity and detecting fraud. Despite its advantages, challenges such as high sequencing costs, data interpretation complexity, and the need for standardized workflows remain. Future research focusing on optimizing real-time sequencing technologies, expanding multi-omics approaches, and addressing regulatory frameworks is suggested to fully harness NGS's potential in ensuring food safety, quality, and innovation.

## KEYWORDS

next-generation sequencing, food microbiome, fermentation, food safety, food authentication

## 1 Introduction

In recent years, food manufacturing, processing, and distribution have evolved to enhance efficiency and meet global consumer demands. However, these advances bring new challenges, including contamination risks from complex supply chains, emerging foodborne pathogens, and increasing consumer expectations for safety, transparency, and sustainability. To address these issues, food science research must integrate modern microbiology tools, which provide insights into microbial ecology, metabolism, and genetics (Kumar et al., 2021).

Microorganisms, such as bacteria and fungi, play an important role in food production, preservation, safety and quality. Advancements in nucleic acid sequencing technologies have significantly improved food microbiology research, allowing for rapid and precise microbial identification, pathogen detection, and food authenticity assessment (Marchelli et al., 2012). Haga clic o pulse aquí para escribir texto. Next-generation sequencing (NGS) has expanded the ability to characterize microbiomes within complex food matrices

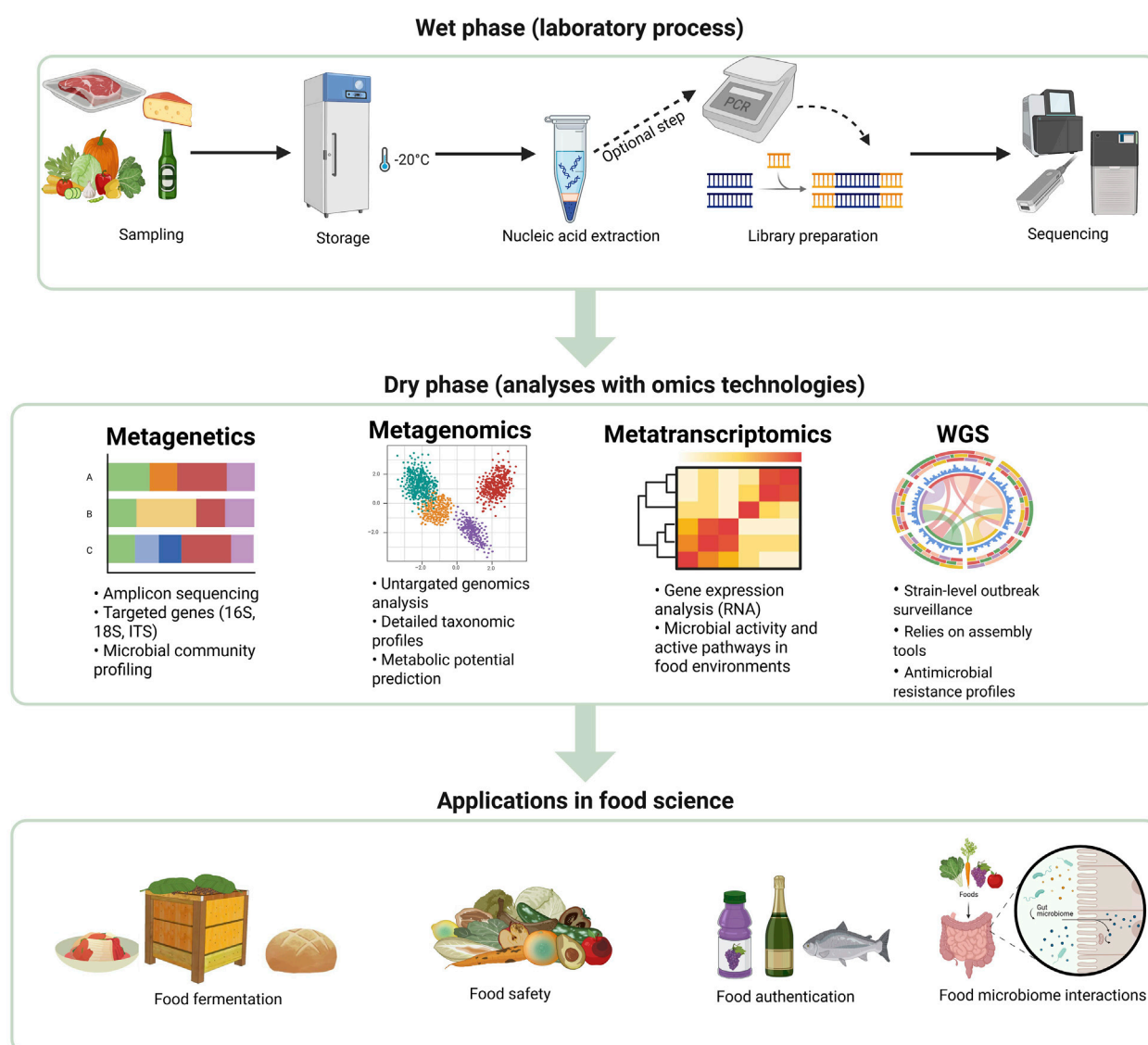
TABLE 1 Main characteristics of NGS platforms and their applications in food science.

NGS technology	Principle	Advantages	Disadvantages	Food science applications	References for applications
Illumina	Sequencing by synthesis	High throughput and accuracy	Short reads, high initial investment (Petersen et al., 2019)	Metagenetics for evaluating food quality and safety	Anagnostopoulos et al. (2022), 2023; Díaz Cárdenas et al. (2024)
				Metagenomics for characterizing microbial dynamics of food fermentation processes	Pothakos et al. (2020), Kim et al. (2021), Lima et al. (2022)
				Metatranscriptomics for flavor formation during food fermentation	Xiao et al. (2021), Pan et al. (2022), Xiong et al. (2024)
				Metagenetics in food fermentation for starter culture development	Pregolini et al. (2021), Tigrero-Vaca et al. (2022), Vaccalluzzo et al. (2022)
				Whole Genome Sequencing (WGS) of foodborne pathogens	Bai et al. (2024), Habib et al. (2024), Yang et al. (2024)
				Metagenomics for studying microbial dynamics during meat processing	Barcenilla et al. (2024), Gaire et al. (2024), Sequino et al. (2024)
Ion Torrent	Sequencing by synthesis, detection of H <sup>+</sup> ions	Small sample size needed. Sequencing takes 2–3 h	Short reads, relatively higher error rate (Shetty et al., 2019)	Metagenetics for seafood quality and safety assessment	Lira et al. (2020)
				Metagenetics for investigating the quality of dairy products	Syromyatnikov et al. (2020), Güley et al. (2023)
				Metagenetics in fermented foods for starter culture assessment	Ferrara et al. (2021)
PacBio	Single-molecule real-time (SMRT) sequencing	Long reads, high accuracy, minimal bias	High initial investment, large sequencer size (Petersen et al., 2019; Ling et al., 2023)	Metagenetics for analyzing the quality of dairy products	Ma et al. (2018), Yang et al. (2020), Liang et al. (2022)
Nanopore	Nanopore electrical signal sequencing	Long reads, portability, easy to use, low capital cost	Relatively high error rates, however Nanopore technology is undergoing constant advancements leading to reduced error rates and enhanced read accuracy (van der Reis et al., 2023)	Metagenomics for identification of foodborne pathogens and antimicrobial resistance genes (AMR)	Solcova et al. (2021), Lee et al. (2024)
				Metagenetics for spoilage microorganism detection in breweries	Kurniawan et al. (2021), Shinohara et al. (2021)
				Metagenetics for evaluating food quality during storage	Gu et al. (2024)
				Metagenomics to analyze microbial communities in fermented foods	Kharnaor and Tamang (2022), Tamang et al. (2022)
				WGS of foodborne pathogens	Martínez-Álvarez et al. (2024), Wang et al. (2024)

through whole-genome sequencing (WGS), metagenetics, metagenomic and metatranscriptomic analysis (Loman and Pallen, 2015; Jagadeesan et al., 2019). Amplicon sequencing is a common approach of metagenetics to examine the diversity of microorganisms by amplification and sequencing of targeted genes or fragments (Carugati et al., 2015). Metagenomics involves untargeted genomic analysis of mixed microbial communities, while metatranscriptomics studies the collective transcriptomes of given habitats (Quince et al., 2017; Ojala et al., 2023). These sequencing approaches allow for deep taxonomic identification, strain-level genome reconstruction, and microbial community characterization (Ferrocino et al., 2023).

Since its introduction in the mid-2000s, NGS has transformed food microbiology research, reducing sequencing costs while improving throughput and accuracy (Goodwin et al., 2016; Heather and Chain, 2016). Continuous advancements in sequencing technologies and bioinformatics have made microbial genomics more accessible, allowing for deeper taxonomic identification and function analysis of food microbiomes (Cao et al., 2017; Cardinali et al., 2017; Van Dijk et al., 2018). This review aims to provide a comprehensive overview of NGS technologies and methodologies, highlighting their applications and potential in food science, and discussing how these innovations are shaping the future of this field.

## Process of NGS analysis in foods



**FIGURE 1**  
Schematic representation of the NGS workflow in food science, illustrating key steps including sample preparation and storage, DNA/RNA extraction, library preparation, sequencing, and the use of bioinformatics tools for various types of analysis including metagenetics, metagenomics, metatranscriptomics and WGS. Created using BioRender: <https://BioRender.com>.

## 2 NGS platforms used in food science

Microbial genome sequencing is now a standard tool in food microbiology research. This is largely owing to advances in NGS technologies, which have made sequencing faster, more accurate, and more affordable (Jagadeesan et al., 2019). Current widely used sequencing technologies can be placed in two major categories: short-read and long-read platforms. Short-read sequencing technologies, such as those developed by Illumina and Ion Torrent, typically rely on sequencing by synthesis (SBS) of complementary DNA strands. In Illumina sequencing, DNA fragments undergo clonal amplification via bridge PCR, followed

by reversible terminator-based sequencing (Ambardar et al., 2016). Illumina systems range from benchtop sequencers like ISeq and MiSeq to production-scale sequencers like HiSeq and NovaSeq (Kulski, 2016). Similarly, Ion Torrent sequencing employs sequencing by synthesis approach, but detects nucleotide incorporation through changes in pH, as hydrogen ions are released during DNA polymerization. This approach eliminates the need for optical detection, using platforms such as the Ion PGM Dx and Ion GeneStudio S5 (Tripathi et al., 2019; Hu et al., 2021).

Long-read sequencing technologies offer longer read lengths (typically >10 kb) and real-time analysis, expanding the scope of

microbial genome studies (Hu et al., 2021). Pacific Biosciences (PacBio) uses single-molecule real-time (SMRT) sequencing, allowing DNA replication without PCR amplification (Ardui et al., 2018). Oxford Nanopore sequencing measures electrical conductivity as nucleic acids pass through a nanopore, offering long-read sequences (Chen et al., 2023; Lamb et al., 2023).

The main characteristics and applications in food science of each type of sequencing platform are presented in Table 1.

### 3 Process of NGS analysis in foods

Irrespective of sequencing technology and platform, each NGS operation includes two main phases: wet and dry. The wet phase refers to the laboratory process, which includes four key steps: (i) collecting and storing samples, (ii) extracting nucleic acids, (iii) library preparation including targeted gene and optional amplification, and (iv) library loading and sequencing run. The dry phase, on the other hand, focuses on the computational analysis of the sequencing data. A schematic overview of this process is illustrated in Figure 1.

#### 3.1 Sample collection and storage

In general, similar considerations and precautions should be applied for food samples subjected to NGS and to conventional microbiological analyses to ensure sample representativeness and integrity. The quantity of samples and repetitions can significantly impact data accuracy and reproducibility. Hence, researchers need to balance the desire for large repetitions and the costs and feasibility of processing these samples. This is especially challenging for food sampling, as raw material microbiota can vary greatly among samples and can change during processing (Sadurski et al., 2024). Both probability (random) and non-probability (non-random)-based approaches have been proven effective in designing sampling schemes tailored to the research question and food type (Taherdoost, 2016).

For food samples to be analyzed using NGS, changes in cell cultivability is not a major concern, and therefore, harsher conditions can be applied, such as snap freezing, rapid drying, or even certain chemical preservatives can be applied to prevent continued microbial growth or other changes that alter the sample nucleic acid profiles (Fricker et al., 2019).

Proper storage is important to prevent nucleic acid degradation or microbial growth (Kazantseva et al., 2021). To mitigate these risks, food samples are usually cooled to 4°C or frozen at −20°C or −80°C, depending on the available facilities (Lear et al., 2018). In contrast, shelf-stable products like freeze-dried cheonggukjang, a fermented soybean food of Korea, can be stored at room temperature in moisture-free conditions, due to their inherent microbial and chemical stability (Tamang et al., 2022).

Sampling must also account for different processing stages, which influence microbial composition. For instance, Barcenilla et al. (2024) used pooled swabs to improve microbial recovery from raw meat surfaces, while tailoring sampling techniques such as direct sampling or surface swabbing for various end products like fermented sausages or cured meats. Similarly, studies like Gaire et al. (2024) and Sequino et al. (2024) have emphasized the importance of

stage-specific sampling to capture shifts in microbial diversity and antimicrobial resistance.

Ultimately, successful NGS-based food science studies rely on thoughtful sampling design and proper storage. Researchers should tailor these steps to the specific food matrix, processing stage, and study goals to ensure meaningful and reproducible results.

#### 3.2 Extraction of nucleic acids

Nucleic acid extraction from food matrices is a vital process in NGS of food products enabling the detection and study of genetic material (Tan and Yiap, 2009). Nucleic acid extraction methods consist of three steps: lysis, purification, and nucleic acid recovery. Furthermore, extraction can be performed using conventional protocols or commercially available kits, with the choice of method depending on sample complexity and study objectives (Sajali et al., 2018).

Cell lysis is carried out to break open microbial cells and release nucleic acids. Lysis methods can be chemical, enzymatic, mechanical, or a combination, depending on the complexity of the matrix (Lever et al., 2015). For instance, enzymatic lysis combined with mechanical disruption has been effectively applied to romaine lettuce (Gu et al., 2024) and coffee (Pothakos et al., 2020) illustrating the importance of method customization based on the nature of the sample.

After cell lysis, nucleic acid purification is performed to separate DNA or RNA from cellular debris, proteins, and inhibitors, typically using liquid-liquid (LLE) or solid-phase extraction methods (SPE) (Ruggieri et al., 2016). For example, the CTAB and chloroform LLE method was used by Lima et al. (2022) to isolate DNA from fermented cacao and by (Xiao et al., 2022) for RNA isolation from Sichuan paocai, highlighting the continued relevance of traditional reagents in specific contexts. SPE kits utilizing silica-based filters reduce reliance on organic solvents and enhance efficiency in nucleic acid recovery (Emaus et al., 2020). These kits have been applied to extract DNA from raw meat swabs (Barcenilla et al., 2024; Gaire et al., 2024) and fermented soybean products (Kharnaor and Tamang, 2022; Tamang et al., 2022) showing their versatility. However, effectiveness can vary depending on sample composition, emphasizing the need for further optimization, particularly for high-fat or polyphenol-rich matrices.

Following lysis and purification, nucleic acid recovery is carried out by pelleting the nucleic acids through centrifugation, followed by resuspension in a suitable buffer. Nuclease-free water or 1× TE buffer is commonly used, with TE offering added protection (Schenk et al., 2023).

In conclusion, nucleic acid extraction is crucial for NGS-based studies on food science, employing diverse methods to ensure DNA/RNA quality and quantity. Tailoring techniques to specific samples enhances analytical reliability and accuracy, reinforcing confidence in research findings. Table 2 provides an overview of nucleic acid extraction techniques utilized in food science.

#### 3.3 Targeted gene and optional amplification

NGS in food science often incorporates a PCR amplification step to selectively enrich target marker genes from complex mixtures of

TABLE 2 Nucleic acid extraction methods utilized in food science.

Extraction method	Description	Advantages	Challenges	Food matrices	Example kits/References
CTAB and chloroform-based extraction	A variation of the phenol-chloroform method using CTAB for plant material	Effective for isolating DNA from challenging matrices	Use of hazardous chemicals	Fermented foods	Xiao et al. (2021), Lima et al. (2022)
Enzymatic lysis + mechanical disruption	Enzymatic lysis combined with mechanical methods to break down tough food matrices	Efficient for complex samples	Time-consuming; may require optimization	Leafy vegetables, fermented foods	Pothakos et al. (2020), Gu et al. (2024)
Silica-based spin columns-DNA kits	Commercial kits using silica columns to purify DNA from food samples	Safe, automated, and reproducible; fast process	More expensive than other methods	General use for various food matrices	NucleoSpin Food DNA kit (Macherey-Nagel); (Kharnaor and Tamang, 2022; Tamang et al., 2022), DNeasy PowerSoil Pro kit (Qiagen) (Barcenilla et al., 2024; Gaire et al., 2024)
Silica-based spin columns-RNA kits	Commercial kits using silica columns to isolate RNA	RNA protection; high yield; compatible with downstream applications	Requires strict RNase-free conditions	Fermented foods, fruits, dairy products	Maxwell 16 LEV simply RNA tissue kit (Promega) (Jo et al., 2021), RNeasy Plant Mini Kit (Qiagen) (Quijada et al., 2022), InnusPEED Bacteria/Fungi RNA kit (Analytik Jena) (Quijada et al., 2022)

genomic DNA. For instance, 16S ribosomal RNA (rRNA) gene sequencing is a targeted approach where PCR is used to amplify specific hypervariable regions of the 16S rRNA gene from diverse bacterial populations present in food samples (Abellan-Schneyder et al., 2021). This amplification enables in-depth profiling of microbial communities, which is essential for assessing food safety, quality, and shelf-life.

In addition to the 16S rRNA gene, other marker genes such as the internal transcribed spacer (ITS) region for fungal identification and the 18S rRNA gene for eukaryotic microorganisms are also routinely targeted for amplicon sequencing (Banos et al., 2018). By selectively amplifying these marker genes, researchers can generate high-quality, focused libraries for NGS, ultimately providing deeper insights into the microbial ecology of food products and aiding in effective monitoring of potential contaminants (Fagerlund et al., 2021).

Additionally, marker genes play a crucial role in food authenticity determination by enabling the detection of species-specific DNA sequences to verify the origin and composition of food products (Vishnuraj et al., 2023). Commonly used markers include mitochondrial genes such as cytochrome b (*cyt b*) and cytochrome c oxidase I (COI) for identifying animal species, as well as nuclear genes like the *rbcl* and *matK* genes for plant species authentication (Grazina et al., 2020; Afifa et al., 2021; Zhao Z. et al., 2024). The *cyt b* and COI genes encode proteins involved in mitochondrial respiration and are favored for their high interspecies variability (Antil et al., 2023; Tutuş et al., 2025). In plants, *rbcl* encodes the large subunit ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) essential for photosynthesis, while *matK* encodes maturase K, involved in chloroplast RNA splicing and known for its high substitution rate and discriminatory power at the species level (Jia et al., 2025). These genetic markers help prevent food fraud, ensure label accuracy, and protect consumers from misrepresentation or adulteration of food products.

In addition to targeted gene amplification, some NGS applications in food science, such as shotgun metagenomics, incorporate optional PCR amplification during library

preparation to enhance sequencing efficiency. For example, the Nextera XT DNA Library Preparation Kit employs enzymatic fragmentation followed by PCR amplification to generate sequencing-ready libraries, particularly when working with food samples containing low-abundance of target DNA (Kim et al., 2013). This approach facilitates comprehensive microbial profiling and functional gene analysis in food samples. However, as seen in metagenomic studies like the genetic characterization of dengue viruses (Lizarazo et al., 2019), excessive PCR cycles could introduce additional biases and duplicates, potentially affecting sequencing accuracy.

### 3.4 Library preparation

NGS library preparation involves processing nucleic acids (RNA or DNA) to align with the chosen sequencing platform (Head et al., 2014). This includes fragmentation, multiplexing, normalization and adapter ligation (Abdi et al., 2024). While amplicon sequencing relies on short DNA fragments of target regions selectively amplified via PCR, shotgun sequencing requires fragmentation of total DNA using sonication, enzymatic digestion, or mechanical shearing (Srinivas et al., 2022). On the other hand, for RNA sequencing in metatranscriptomic analyses, mRNA molecules are converted to cDNA fragment by reverse transcription, using either poly(A) tailing (targeting eukaryotes) or random priming (targeting prokaryotes) (Saliba et al., 2014). Short-read platforms like Illumina favor fragments under 450 bp, whereas long-read platforms such as PacBio and Oxford Nanopore Technologies (ONT) require high molecular weight (HMW) DNA, sometimes sheared to ~20 kb for improved sequencing efficiency (Madoui et al., 2015; Ribarska et al., 2022).

Multiplexing, or indexing, allows multiple libraries to be pooled and sequenced together using unique index sequences. To ensure equal representation, libraries must first be normalized to the same DNA/RNA concentration before being pooled in equal volumes (Muller et al., 2019). However, challenges like index hopping (where



reads are misassigned) can introduce errors, particularly in Illumina platforms with patterned flow cells. Unique dual indexing and nested metabarcoding help mitigate these issues (MacConaill et al., 2018; Guenay-Greunke et al., 2021).

Adapter ligation attaches platform-specific sequences to DNA fragments, enabling binding to flow cells. Illumina uses adapters for anchoring, whereas PacBio employs hairpin adapters for circular DNA sequencing. ONT ligates adapters that guide DNA strands through nanopores (Oikonomopoulos et al., 2020; MacKenzie and Argyropoulos, 2023). Amplicon sequencing on Illumina integrates adapters directly during PCR, bypassing the ligation step (Glenn et al., 2019).

Minimizing contamination is critical, as microbial DNA from reagents can bias NGS analyses. Therefore, proper laboratory practices, negative controls, and mock microbial communities help reduce this risk (Salter et al., 2014). Advances in automation and quality control continue to enhance the reliability and efficiency of NGS library preparation in food science applications.

## 3.5 Sequencing

NGS methods for food microbiome analysis can be broadly divided into targeted (amplicon-based) and random (shotgun) sequencing approaches, each serving different analytical goals (Sekse et al., 2017).

Amplicon sequencing, also known as targeted sequencing, focuses on specific genetic markers, such as the 16S rRNA gene for bacteria, ITS regions for fungi, and COI genes for metazoans (Francioli et al., 2021). This approach, using platforms like Illumina MiSeq and Ion Torrent PGM is widely adopted due to cost-effectiveness and established analytical pipelines (Lira et al., 2020; Syromyatnikov et al., 2020; Anagnostopoulos et al., 2022; 2023; Güley et al., 2023). However, PCR-induced biases in primer selection can significantly affect microbiota characterization (Sergeant et al., 2012). In fungi, ITS length variation can distort community structure by favoring shorter fragments (De Filippis et al., 2017). Amplicon sequencing also lacks strain-level resolution, limiting food safety assessments (La Reau et al., 2023). To mitigate biases, researchers can consider alternative fungal-specific targets, such as 26S and 18S rRNA genes, for improved accuracy (Banos et al., 2018; Mota-Gutierrez et al., 2019).

Shotgun metagenomic sequencing, in contrast, sequences all DNA present in a sample, providing a comprehensive view of the genetic content, including functional genes, metabolic pathways, and antimicrobial resistance determinants (Tyagi et al., 2019). This approach enables deep characterization of food microbiomes, making it ideal for studying fermentation and resistome profiling (Pothakos et al., 2020; Kim et al., 2021; Kharnaor and Tamang, 2022; Lima et al., 2022; Tamang et al., 2022; Lee et al., 2024). However, it cannot easily distinguish live from dead cells. When viability is critical, combining metagenomics with culture-dependent techniques ensures accurate risk assessment (Lindner et al., 2024). Shotgun metagenomics requires robust computational resources for assembly and annotation (Tremblay et al., 2022). Platforms like Illumina NovaSeq and PacBio Sequel II offer high-throughput sequencing, while long-read technologies such as ONT

and PacBio HiFi enhance genome assembly and resolution in complex food matrices (Zhang et al., 2022).

In addition, whole genome sequencing (WGS) and RNA sequencing (RNA-Seq) are also important for microbial analysis and food safety. Platforms used for WGS of foodborne pathogens include Oxford Nanopore MinION, for *Escherichia coli* characterization from meat products (Wang et al., 2024; Martínez-Álvarez et al., 2025). Illumina systems like NextSeq, HiSeq and NovaSeq, are widely employed for RNA-Seq for elucidating important microbial functions during food fermentation processes, as well as in transcriptomic studies of foodborne pathogens in fresh produce and related matrices (Xiao et al., 2021; Redding et al., 2024; Xiong et al., 2024; 2024; Ding et al., 2025). Comprehensive metatranscriptomic analyses of mRNA dynamics can also provide insights on microbial interactions in complex microbial communities or biological processes.

To enhance sequencing efficiency and accuracy, several strategies can be employed. Optimizing sample preparation helps minimize biases, especially in complex food matrices. Choosing the right sequencing platform ensures a balance between cost, read length, and throughput. Combining amplicon and shotgun approaches can provide complementary insights, and advancements in bioinformatics tools as discussed in the following section, such as machine learning-based classification and improved genome binning, can facilitate improved data interpretation for better tracking of foodborne pathogens and microbial ecology shifts.

## 3.6 Bioinformatic analysis

The vast amount of data generated by NGS requires robust bioinformatics pipelines for accurate analysis and interpretation. The bioinformatics workflow varies depending on the sequencing approach as well as the specific NGS platform used.

### 3.6.1 Metagenetic analysis

Metagenetic analysis, including amplicon high-throughput sequencing, typically follows a standardized workflow, beginning with quality control (QC) and filtering to remove low-quality reads and sequencing artifacts. This step is commonly performed using tools like FastQC (Andrews, 2010) and DADA2 (Callahan et al., 2015) or USEARCH (Zhou et al., 2024), which help denoise sequences and cluster sequence readings into operational taxonomic units (OTUs) or amplicon sequence variants (ASVs). Taxonomic classification is then conducted using reference databases such as SILVA, Greengenes, or UNITE (for fungal ITS sequences) (Santamaria et al., 2012).

Common bioinformatics platforms, including QIIME2 (Bolyen et al., 2019) and Mothur (Schloss et al., 2009), streamline these processes and are widely applied in food microbiome studies. For instance, a study analyzing the bacterial diversity in fermented beverages used QIIME2 to assess microbial shifts during different stages of fermentation (Zhao X. et al., 2024). Similarly, UNITE was employed in a study to identify fungal communities in stored rice grains and their potential mycotoxin production (Qi et al., 2022).

The choice of reference databases and bioinformatics tools significantly impacts taxonomic classification accuracy,

influencing the results of amplicon-based studies (De Filippis et al., 2017). While bacterial databases such as SILVA and Greengenes are well-curated, fungal databases remain comparatively less refined (Tedersoo et al., 2011). Enhancing genomic databases with foodborne microbial genomes and developing food-specific gene catalogs would improve taxonomic resolution and strengthen food microbiome research.

### 3.6.2 Metagenomic analysis

Metagenomic analysis involves a more complex computational pipeline due to the high volume of sequencing data and the need for assembly. Quality control is performed using Trimmomatic (Bolger et al., 2014) or FASTP (Chen, 2023) to remove adapter sequences and low-quality reads. Reads are then assembled using tools like MEGAHIT (Li et al., 2015) or SPAdes (Bankevich et al., 2012), which reconstruct longer contigs from fragmented sequences.

Taxonomic and functional annotation is performed using databases such as Kraken2, MetaPhlAn, and KEGG, allowing researchers to identify microbial taxons, functional genes, and antimicrobial resistance markers in food samples (Nam et al., 2023). A study investigating the microbiome of food and environmental microbiomes in various meat processing facilities used Kraken2 for taxonomic classification, revealing that microbial communities change throughout processing, from raw materials to final products, with food contact surfaces significantly influencing the final microbiome (Barcenilla et al., 2024). Similarly, a metagenomic analysis of fermented vegetables used KEGG to map functional genes related to amino acid metabolism and probiotic activity (Yasir et al., 2022).

Long-read sequencing platforms, such as PacBio and Oxford Nanopore Technologies (ONT), improve metagenomic studies by generating high-continuity assemblies, enabling strain-level identification and plasmid detection in foodborne pathogens (Kwon et al., 2020; Espinosa et al., 2024). Tools like Flye (Kolmogorov et al., 2019) and Canu (Koren et al., 2017) are commonly used for long-read assembly, while Medaka and Racon aid in sequence polishing to improve accuracy (Lee et al., 2021). However, metagenomic data is inherently compositional, meaning relative abundances represent proportions rather than absolute feature loads (Morton et al., 2024). This perspective is crucial when interpreting pathogen detection, as any read set represents only a subset of the total community DNA. Proper contextualization ensures accurate conclusions about microbial presence and potential risks (Lindner et al., 2024).

### 3.6.3 Metagenomic binning and downstream analysis of MAGs

Metagenome-assembled genomes (MAGs) provide insights into microbial community structure and function in food and in food production environments (Zhou et al., 2022). Metagenomic binning tools such as MetaBAT2 (Kang et al., 2019), CONCOCT (Alneberg et al., 2014), and MaxBin (Wu et al., 2014) group contigs into individual genome bins. A study examining microbial communities in artisanal cheese production successfully recovered high-quality MAGs using MetaBAT2, identifying bacterial strains responsible for flavor development (Walsh et al., 2020).

Quality assessment of MAGs can be performed using CheckM (Parks et al., 2015) and GUNC (Orakov et al., 2021) to ensure completeness and reduce contamination. Further downstream analysis includes genome annotation using tools like Prokka

(Seemann, 2014) and EggNOG-mapper (Huerta-Cepas et al., 2017), which provide insights into microbial metabolism in food matrices.

### 3.6.4 Functional profiling

Functional profiling is a key step in understanding microbial metabolism in food matrices. Tools like HUMAnN3, MEGAN, FAPROTAX, and PICRUSt2 predict metabolic functions based on sequencing data (Beier et al., 2017; Nam et al., 2023). For example, a study on fermented dairy products used HUMAnN3 to profile the microbial pathways involved in amino acid metabolism, linking them to health benefits (Yang et al., 2025).

### 3.6.5 Metatranscriptomic analysis

Metatranscriptomics provides insights into microbial activity in food environments by analyzing RNA sequences. Bioinformatics tools streamline this process through key steps: quality control, assembly, annotation, and functional analysis. Preprocessing tools like FastQC, Trimmomatic, and SortMeRNA (Kopylova et al., 2012) ensure high-quality RNA-seq data. Read alignment and assembly rely on HISAT2 (Wen, 2017), STAR (Dobin et al., 2013), MEGAHIT, and Trinity (Haas et al., 2013), while functional annotation is performed using DIAMOND and Kaiju (Buchfink et al., 2015; Menzel et al., 2016). Differential expression and pathway analysis involve DESeq2, edgeR, and KEGG Mapper (Robinson et al., 2010; Anders and Huber, 2010; Kanehisa and Sato, 2020).

### 3.6.6 WGS and antimicrobial resistance (AMR) analysis

WGS analysis, essential for tracing foodborne outbreaks, relies on assembly tools such as Unicycler (Wick et al., 2017) and annotation tools like Bakta (Schwengers et al., 2021) and ROARY (Page et al., 2015) for comparative genomics. For example, a WGS-based study on antibiotic-resistant *E. coli* from market chickens in Lima, Peru, used ROARY for comparative genomics to assess resistance gene distribution and potential transmission sources (Murray et al., 2021).

AMR surveillance in foodborne pathogens is critical for food safety. Databases like CARD (Comprehensive Antibiotic Resistance Database) (Alcock et al., 2023) and ResFinder (Florensa et al., 2022) help identify AMR genes, while tools such as AMRFinderPlus (Feldgarden et al., 2021) and ARIBA (Hunt et al., 2017) predict resistance profiles based on WGS and metagenomic data. A recent study utilizing nanopore sequencing-based metagenomics leveraged CARD to comprehensively identify AMR genes in food products, highlighting the prevalence of resistance determinants across diverse microbial communities (Lee et al., 2024). Similarly, ARIBA was used to analyze *E. coli* from store-bought produce, identifying diverse AMR genes and shedding light on potential resistance transmission through fresh vegetables (Reid et al., 2020).

There is an urgent need for standardized, open-access AMR databases to improve method comparison and surveillance. Advancing these technologies in microbiology labs will enhance detection, support personalized medicine, and streamline AMR monitoring.

### 3.6.7 Advancing bioinformatics in food science

To enhance bioinformatics analysis, integrating multi-omics approaches (e.g., metagenomics with metabolomics or

**TABLE 3** Bioinformatics tools and workflows for NGS analysis in food science.

Analysis type	Tools/Databases	Workflow steps	Key references
Metagenetic analysis	<ul style="list-style-type: none"> <li>FastQC</li> <li>DADA2</li> <li>USEARCH</li> <li>QIIME2, Mothur</li> <li>SILVA, Greengenes</li> <li>UNITE</li> </ul>	<ul style="list-style-type: none"> <li>Quality control and filtering</li> <li>Sequence denoising and clustering into OTUs or ASVs</li> <li>Comprehensive platforms that streamline analysis workflows and databases for taxonomic classification</li> </ul>	FastQC (Andrews, 2010), DADA2 (Callahan et al., 2015), USEARCH (Zhou et al., 2024), QIIME2 (Bolyen et al., 2019), Mothur (Schloss et al., 2009), SILVA, Greengenes, UNITE (Santamaria et al., 2012)
Metagenomic analysis	<ul style="list-style-type: none"> <li>Trimmomatic, FASTP</li> <li>MEGAHIT, SPAdes</li> <li>Kraken2, (Segata et al., 2012)</li> <li>MetaPhlAn</li> <li>KEGG</li> </ul>	<ul style="list-style-type: none"> <li>Quality control</li> <li>Sequence assembly</li> <li>Taxonomic classification</li> <li>Functional annotation</li> </ul>	Trimmomatic (Bolger et al., 2014), FASTP (Chen, 2023), MEGAHIT (Li et al., 2015), SPAdes (Bankevich et al., 2012), Kraken 2 (Wood et al., 2019), MetaPhlAn (Segata et al., 2012), KEGG (Kanehisa et al., 2017)
Long-read sequencing and assembly	<ul style="list-style-type: none"> <li>Flye, Canu</li> <li>Medaka, Racon</li> </ul>	<ul style="list-style-type: none"> <li>Long-read assembly</li> <li>Sequence polishing</li> </ul>	Flye (Kolmogorov et al., 2019), Canu (Koren et al., 2017), Medaka, Racon (Lee et al., 2021)
Metagenomic binning and MAGs analysis	<ul style="list-style-type: none"> <li>MetaBAT2, CONCOCT, MaxBin</li> <li>CheckM, GUNC</li> <li>Prokka, EggNOG-mapper</li> </ul>	<ul style="list-style-type: none"> <li>Binning of contigs into genome bins</li> <li>Quality assessment of MAGs</li> <li>Genome annotation</li> </ul>	MetaBAT2 (Kang et al., 2019), CONCOCT (Alneberg et al., 2014), Maxbin (Wu et al., 2014), CheckM (Parks et al., 2015), GUNC (Orakov et al., 2021), Prokka (Seemann, 2014), EggNOG-mapper (Huerta-Cepas et al., 2017)
Functional profiling	<ul style="list-style-type: none"> <li>HUMAnN3, MEGAN, FAPROTAX, PICRUSt2</li> </ul>	<ul style="list-style-type: none"> <li>Prediction of metabolic functions</li> </ul>	HUMAnN3 (Beghini et al., 2021), MEGAN (Beier et al., 2017), FAPROTAX (Louca et al., 2016), PICRUSt2 (Douglas et al., 2020)
Metatranscriptomic analysis	<ul style="list-style-type: none"> <li>FASTQC, Trimmomatic, SortMeRNA</li> <li>HISAT2, STAR, MEGAHIT, Trinity</li> <li>Diamond, Kaiju</li> <li>DESeq2, edgeR, and KEGG Mapper</li> </ul>	<ul style="list-style-type: none"> <li>Preprocessing</li> <li>Alignment and assembly</li> <li>Functional annotation</li> <li>Differential expression and pathway analysis</li> </ul>	SortMeRNA (Kopylova et al., 2012), HISAT2 (Wen, 2017), STAR (Dobin et al., 2013), Trinity (Haas et al., 2013), Diamond (Buchfink et al., 2015), Kaiju (Menzel et al., 2016), DESeq2 (Anders and Huber, 2010), edgeR (Robinson et al., 2010), KEGG Mapper (Kanehisa and Sato, 2020)
WGS and AMR analysis	<ul style="list-style-type: none"> <li>Unicycler</li> <li>Bakta, ROARY</li> <li>CARD, ResFinder, AMRFinderPlus, ARIBA</li> </ul>	<ul style="list-style-type: none"> <li>Genome assembly and annotation</li> <li>Comparative genomics</li> <li>AMR gene identification</li> </ul>	Unicycler (Wick et al., 2017), Bakta (Schwengers et al., 2021), ROARY (Page et al., 2015), CARD (Alcock et al., 2023), Resfinder (Florensa et al., 2022), AMRFinderPlus (Feldgarden et al., 2021), ARIBA (Hunt et al., 2017)

transcriptomics) can provide a deeper understanding of microbial activity and food quality (Balkir et al., 2021). Cloud-based bioinformatics platforms, such as MG-RAST (Keegan et al., 2016) and Galaxy (Goecks et al., 2010), enable scalable and user-friendly data analysis. The development of machine learning algorithms for microbial classification, AMR prediction, and genome annotation is further enhancing NGS applications in food safety and quality control. Table 3 provides an overview of the key bioinformatics tools and databases employed at various stages of NGS analysis in food science.

## 4 Application of NGS in food science

### 4.1 NGS for food safety

The application of NGS in food safety assessment has significantly improved pathogen detection, microbial community analysis, and AMR surveillance (Pennone et al., 2022).

Metagenetics, primarily based on 16S rRNA gene sequencing, is used to study microbial communities in food and their processing environments. For example, Anagnostopoulos et al. (2022) and

Anagnostopoulos et al. (2023) used 16S metabarcoding to analyze microbial succession in ice-stored seabream, revealing shifts in bacteria like *Pseudomonas* and *Shewanella* associated with spoilage. Similarly, Syromyatnikov et al. (2020) applied 16S sequencing to butter microbiota, detecting opportunistic pathogens missed by traditional methods. Díaz Cárdenas et al. (2024) studied street-vended foods in Ecuador, identifying dominant genera like *Acinetobacter*, *Lactococcus*, and *Vibrio*. They found twenty-nine spoilage bacteria and twenty-four opportunistic pathogens, underscoring the food safety risks in these environments and highlighting the importance of metagenetics for food safety monitoring.

WGS offers unmatched precision for examining individual microbial isolates, allowing for meticulous strain tracking during outbreaks. Bai et al. (2024) applied WGS to fresh-cut fruits and vegetables, detecting *Salmonella*, *Listeria monocytogenes*, and *E. coli* along with virulence and resistance genes. Similarly, Habib et al. (2024) found multidrug-resistant *Salmonella enterica* in chilled broiler chicken. These bacterial pathogens are among the most relevant to public health in the various food production sectors in the EU/EEA (Koutsoumanis et al., 2024). Moreover, the emergence of long-read sequencing technologies like Oxford



TABLE 4 Overview of foodborne pathogens, food sources, and their detection via NGS.

Foodborne pathogen	Food source	NGS focus	Sequencing platform	Application	References
<i>Bacillus cereus</i> , <i>L. monocytogenes</i> , <i>Salmonella</i> spp., <i>Shigella</i> spp., <i>Vibrio cholerae</i> , <i>Vibrio parahaemolyticus</i> and <i>Vibrio vulnificus</i>	Fresh seafood	Metagenetics	Illumina	Explored the microbial diversity in seafood factories, revealing significant insights into potential pathogenic threats	Kuan et al. (2025)
<i>Campylobacter jejuni</i>	Contaminated poultry products	Transcriptomics	Illumina	Investigated the mechanism of antimicrobial formulations against <i>Campylobacter</i>	Tang et al. (2021)
<i>C. jejuni</i>	Retail chicken	Comparative genomics	Illumina, Nanopore	Compared genomic data between <i>C. jejuni</i> isolates to understand virulence factors and resistance mechanisms	Neal-McKinney et al. (2021)
<i>Cronobacter sakazakii</i>	Garlic extract	Transcriptomics	Illumina	Analyzed responses to antimicrobial compounds from garlic, establishing gene expression profiles	Feng et al. (2014)
<i>E. coli</i>	Canola sprouts	Transcriptomics	Illumina	Identified genes essential for survival in plant tissues, emphasizing iron acquisition mechanisms	Na et al. (2018)
<i>E. coli</i> HEHA16, <i>L. monocytogenes</i> , <i>Salmonella enterica</i> Typhi, <i>Cronobacter sakazakii</i>	Meat and meat analogues	Whole Genome Sequencing	Illumina	Investigated the dynamics of foodborne pathogens in meats using traditional microbiology and NGS for better pathogen detection	Bonaldo et al. (2023)
<i>L. monocytogenes</i>	Food grade stainless steel	Transcriptomics	Illumina	Evaluated transcriptional changes upon desiccation stress, revealing adaptation mechanisms	Kragh and Truelstrup Hansen (2020)
<i>L. monocytogenes</i>	Food processing environments	Metagenetics	Illumina	Identified diverse <i>Listeria</i> spp. in food processing facilities, linking environmental microbiota to pathogen presence	Tan et al. (2019)
Shiga toxin-producing <i>E. coli</i>	Raw palm sap Samples	Metabarcoding	Nanopore	Used nanopore sequencing to identify pathogenic species in palm sap, emphasizing NGS approaches for foodborne pathogen monitoring	Ábrahám et al. (2024)
<i>Vibrio</i> spp. and opportunistic pathogens	Street-vended foods	Metabarcoding	Illumina	Characterized the microbial composition of street vended foods	Díaz Cárdenas et al. (2024)

Nanopore suggests a pathway toward real-time, portable, on-site food safety testing (Lee et al., 2024). To provide a clearer overview of key foodborne pathogens, their typical food sources, and their detection using NGS, Table 4 summarizes selected examples aligned with recent research.

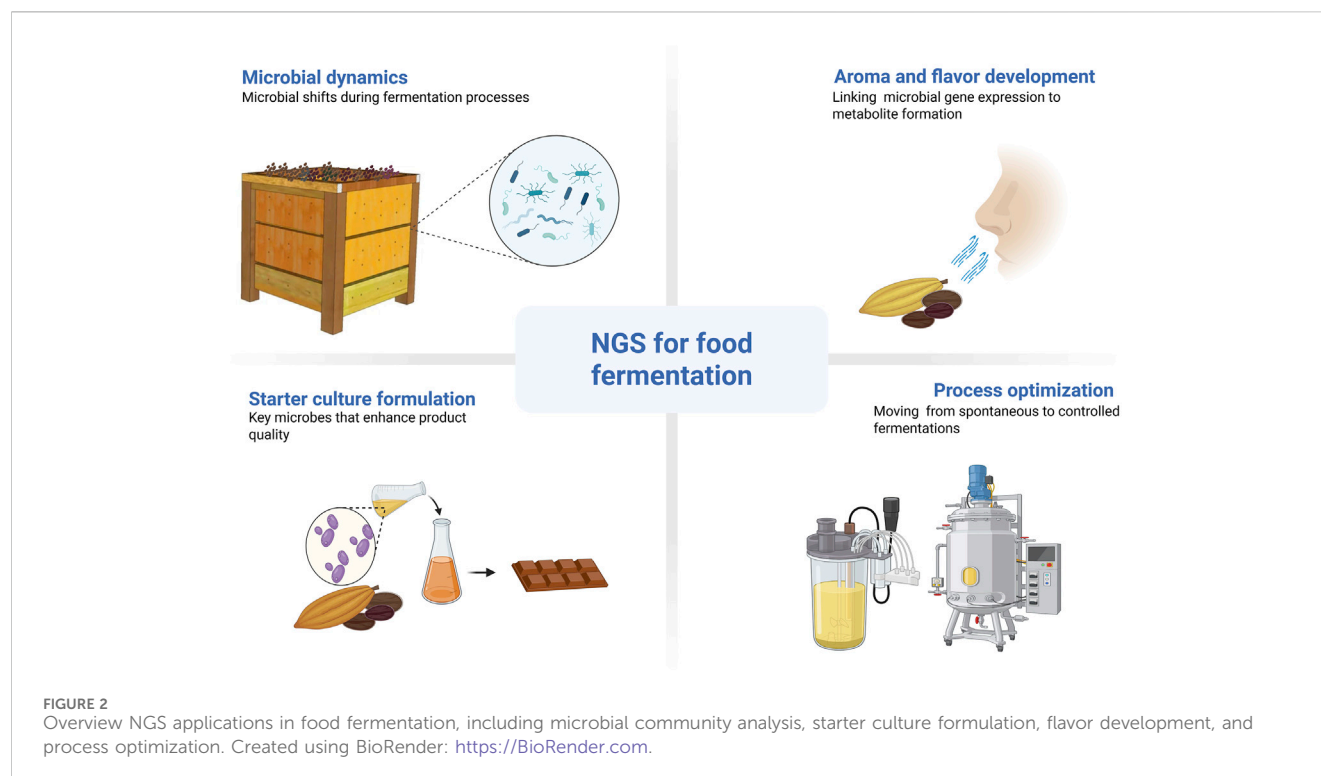
Beyond outbreak detection, NGS also strengthens food safety through its role in preventive quality control and source attribution. By characterizing microbial contaminants at high resolution, NGS can help detect the origin of contamination, differentiating whether it arose from raw ingredients, surfaces, or environmental niches. This technology is especially useful when multiple sources of contamination are involved, as genomic comparison of isolates from diverse sampling points can uncover distinct strain lineages or reveal shared transmission routes (Rantsiou et al., 2018; Koutsoumanis et al., 2024).

Metatranscriptomics, while less frequently deployed, introduces a unique analytical layer by targeting gene expression (RNA) rather than solely genetic potential (DNA). This approach is particularly useful for assessing spoilage mechanisms and microbial metabolism in perishable foods. For example, de Lira et al. (2020) integrated metatranscriptomics to examine histamine production in fish, which

highlights its utility in connecting specific microbial functions to observable food quality deterioration. Its application in verifying the activity of foodborne viruses like norovirus Yang et al. (2024), also demonstrates its value extends beyond bacteria. This technique, therefore, provides critical insights into which microbes are metabolically active and the functions they are executing within the food matrix at a specific moment.

The widespread adoption of NGS across varied food industries from dairy (Ma et al., 2018; Yang et al., 2020; Liang et al., 2022) and meat/poultry products (Barcenilla et al., 2024; Martínez-Álvarez et al., 2025; Wang et al., 2024) to fresh produce (Solcova et al., 2021; Gu et al., 2024) attests to its flexibility, but also brings common challenges into focus. A consistent theme emerging from these diverse studies is the significant influence of processing environments on microbial ecology and the spread of AMR genes (Gaire et al., 2024; Koutsoumanis et al., 2024), alongside the risks posed by pathogens and resistance elements in raw ingredients. This implies that microbial risks are ubiquitous yet shaped by the specifics of each food type and its journey through the production chain.

For effective management, NGS monitoring strategies should be customized for different sectors, concentrating on critical control



points identified via thorough environmental and product analyses. Promoting the sharing of data and insights from these varied applications, following the example of successful collaborative projects (Allard et al., 2016) will be instrumental in refining best practices and informing risk management universally.

Comparing NGS with conventional methods starkly illustrates its advantages in speed, resolution, and culture-independent analysis, facilitating more potent outbreak responses and surveillance (Forbes et al., 2017; Hilt and Ferrieri, 2022; Panwar et al., 2023). The practical impact is evident in real-world scenarios, such as the FDA resolving complex outbreaks using WGS data in regulatory decisions (Ottesen and Ramachandran, 2019; Mak et al., 2025). Nevertheless, significant hurdles impede its routine implementation: costs, the demand for specialized bioinformatics skills and absence of standardization present considerable barriers. Future progress hinges on collaborative initiatives aimed at standardizing protocols and bioinformatics workflows, substantial investment in training and infrastructure, and the creation of unambiguous guidelines for validation and data interpretation.

## 4.2 NGS for food fermentation

Food fermentation enhances food products through complex microbial activities, traditionally studied with methods that were limited in scope. NGS offers a transformative approach, providing extensive insights into microbial community structure, function, and dynamics (Pothakos et al., 2020; Ferrara et al., 2021; Kim et al., 2021; Lima et al., 2022; Tamang et al., 2022; Xiong et al., 2024). The application of NGS reveals the previously hidden microbial diversity, including unculturable organisms, vital for fully understanding fermentation processes. Figure 2 summarizes the

main contributions of NGS to food fermentation, including insights into microbial dynamics, starter culture development, process optimization, and the linkage between gene expression and flavor compound formation.

NGS methods like amplicon sequencing and shotgun metagenomics provide detailed taxonomic profiles (Ferrara et al., 2021; Tigrero-Vaca et al., 2022; Vaccalluzzo et al., 2022), and metatranscriptomics identifies active genes and microbes during fermentation stages (Pothakos et al., 2020; Kim et al., 2021; Xiong et al., 2024). Studies exemplify this power: identifying biogenic amine producers in soy sauce highlights NGS's role in assessing safety aspects alongside flavor development (Kim et al., 2021); revealing microbial shifts influencing volatile compounds in olives demonstrates its utility in targeted quality improvement (Vaccalluzzo et al., 2022); and uncovering complex networks in focaccia contrasts sharply with the limited view from plating methods (Ferrara et al., 2021).

The consistent finding across fermentation studies is that NGS offers a far more comprehensive view than traditional methods, which is crucial for informed development of starter cultures. A potential gap suggested by the focus on dominant bacteria (like lactic acid bacteria) is the need for deeper investigation into the roles of sub-dominant bacteria, yeasts, and potentially other microbes; employing deeper sequencing or targeted enrichment could address this.

Functionally, metagenomics predicts the metabolic potential (Kharnaor and Tamang, 2022; Lima et al., 2022; Tamang et al., 2022), while metatranscriptomics confirms active pathways (Kim et al., 2021; Xiong et al., 2024). Linking specific microbial gene expression to metabolite formation provides direct targets for intervention (Kim et al., 2021; Kharnaor and Tamang, 2022; Xiao et al., 2022; Xiong et al., 2024). Uncovering enzymatic

functions in cacao or amino acid profiles in cheonggukjang directly informs strategies for enhancing specific sensory attributes (Lima et al., 2022; Tamang et al., 2022; Tigrero-Vaca et al., 2022). A significant implication is the ability to move from spontaneous to controlled fermentations by understanding and manipulating these functional links.

Ultimately, the application of NGS profoundly shapes the future of food fermentation. Identifying key microbes and metabolic pathways facilitates the rational design and selection of starter cultures tailored for specific outcomes, such as desired flavor profiles in cacao or controlled activity in coffee (Pregolini et al., 2021; Lima et al., 2022; Tigrero-Vaca et al., 2022; Constante Catuto et al., 2024). Understanding microbial succession and function, as shown in paocai and suancai, is critical for process optimization (Xiao et al., 2022; Xiong et al., 2024). A clear direction arising from these studies is the synergy gained from multi-omics approaches; combining genomics, transcriptomics, and metabolomics offers the most holistic view (Kharnaor and Tamang, 2022; Xiong et al., 2024) and represents a key strategy to address remaining knowledge gaps about complex microbial interactions and their precise impact on final product characteristics.

### 4.3 Other NGS applications in food science

In addition to food safety and fermentation, NGS is driving innovative approaches in food authentication, traceability, and product integrity. Techniques like metagenomics and metabarcoding are being employed to ensure the quality and authenticity of food products.

Metabarcoding, is a molecular technique that enables the identification of multiple species within a mixed sample, such as bulk or environmental DNA, through high-throughput sequencing of a targeted genetic marker (Liu et al., 2020). In food science, metabarcoding plays a crucial role in species authentication, detection of adulteration, and monitoring of microbial communities (Mottola et al., 2024). For example, Detcharoen et al. (2024) used Nanopore sequencing to authenticate fish species in surimi-based products, offering a rapid and accurate method for detecting species substitution and ensuring food traceability. Additionally, Giusti et al. (2024) employed metabarcoding to uncover species mislabeling in insect-based products marketed in the EU, underscoring its value in ensuring transparency and regulatory compliance in novel food markets.

Nanopore sequencing has also proven valuable in real-time, on-site detection of contaminants in brewing processes. Kurniawan et al. (2021) developed a rapid nanopore-based platform to identify beer-spoiling bacteria directly in breweries, improving quality control and minimizing spoilage risks. In another brewing application, (Shinohara et al., 2021), utilized nanopore sequencing to identify yeast species quickly and accurately in breweries, enhancing yeast management and beer production.

Beyond these applications, NGS has also emerged as a powerful tool in studying food microbiome interactions. For example, Lerma-Aguilera et al. (2024) examined how cooking methods influence microbial diversity, while Küçükğöz et al. (2025) examined the impact of fermented beetroot ketchup on gut microbiota. Similarly, Gong et al. (2024) investigated the effects of processed

oats and pinto beans on gut microbiota, demonstrating that dietary fibers from these foods promote beneficial bacteria. Additionally, Milani et al. (2025) demonstrated how bacteria present in cheese can modulate the gut microbiome, emphasizing the role of food as a vehicle for beneficial microbes. These findings underscore the growing importance of NGS in assessing the functional effects of food on human health.

### 4.4 Integration of artificial intelligence (AI) and machine learning (ML) with NGS in food science

The convergence of NGS with AI and ML is rapidly reshaping food science, offering predictive power and real-time decision-making in areas such as food safety, traceability, and microbial risk assessment (Liao et al., 2024). These approaches go beyond traditional detection and profiling, enabling the extraction of meaningful patterns from complex, high-dimensional sequencing data.

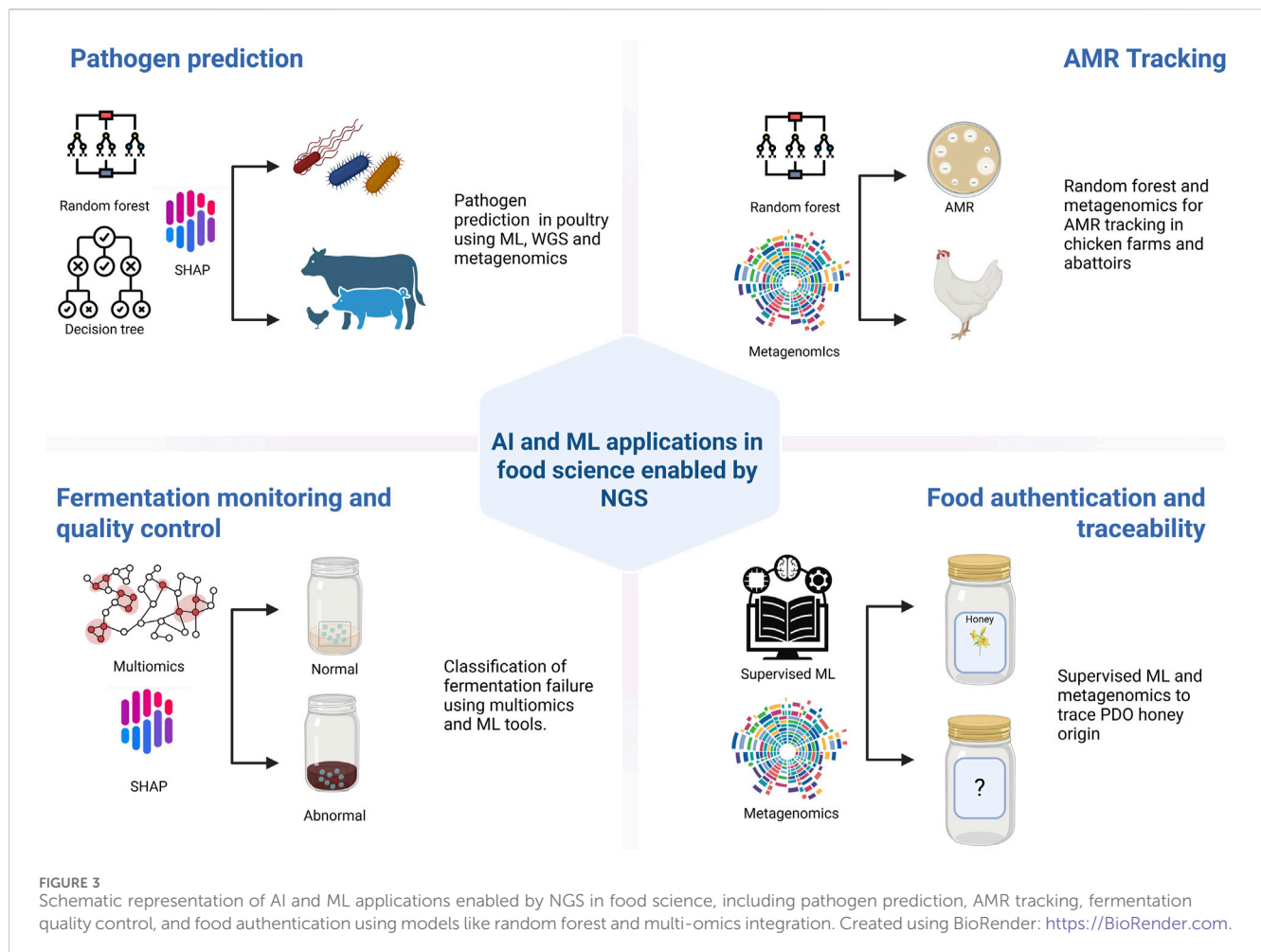
Supervised learning models, such as random forests and decision trees, have been used to predict foodborne pathogen contamination. For example, Bolinger et al. (2021) applied random forest classifiers to 16S rRNA microbiome profiles from poultry rinsates to estimate *Salmonella* contamination risk at slaughter. Their model identified specific microbial signatures as reliable proxies for pathogen presence, demonstrating the predictive value of sequencing data integrated with ML algorithms.

Explainable AI approaches are also emerging. Ince et al. (2025) combined WGS with a decision-tree model, using SHAP (SHapley Additive exPlanations) values to quantify the influence of microbiota features and temporal variables on the growth of *Clostridium perfringens* in pork. This interpretable framework offered microbiologically relevant insights, advancing food spoilage prediction and process control.

Baker et al. (2023) further demonstrated the use of ML in food safety surveillance by combining shotgun metagenomics with unsupervised clustering and supervised models to examine microbial communities and AMR genes across poultry farms and abattoirs in China. The study revealed consistent resistome profiles across sites, suggesting common selective pressures and potential routes of cross-contamination. This illustrates how AI-enhanced NGS analyses can support comprehensive surveillance of AMR and microbial hazards across the food production chain.

Beyond safety, ML tools have also been applied to food authentication. Sabater et al. (2024) combined metagenomic sequencing with supervised machine learning to trace the geographic origin of Spanish Protected Designation of Origin (PDO) honey, identifying microbial community signatures as indicators of authenticity. This approach underscores the potential of ML-integrated NGS data to support fraud detection and ensure product traceability. Similar strategies could be extended to other high-value or regulated food products, making AI-enhanced NGS an asset in modern food authentication frameworks.

In the domain of food fermentation, ML and multi-omics approaches are proving especially powerful (Li et al., 2025). Integrated metagenomics sequencing and metabolomics with machine learning algorithm (logistic regression, K-nearest



neighbors (KNN), and random forest) to classify abnormal stacking fermentations in sauce-flavor Baijiu. Their models successfully distinguished between types of fermentation failure, and SHAP analysis was used to identify key microbial and metabolite biomarkers driving the predictions. This study exemplifies how AI enhanced multi-omics can uncover mechanistic insights into microbial dynamics and functional disruptions, supporting both quality control and the optimization of traditional fermentation processes.

Figure 3 presents a schematic representation of AI and ML applications in food science enabled by NGS, including pathogen prediction, AMR tracking, fermentation monitoring, and food authentication. The diagram highlights the integration of models such as random forest, SHAP, and supervised learning with omics data to support predictive and diagnostic capabilities across food systems.

Despite these advances, challenges remain. Many models are dataset-dependent, limiting generalizability across food systems. The need for standardized pipelines, validated microbial markers, and robust training datasets is critical. Furthermore, regulatory acceptance and interpretability of AI outputs are ongoing concerns.

Looking forward, integrating multi-omics data (e.g., genomics, transcriptomics, metabolomics) with hybrid AI models may provide a more holistic understanding of microbial behavior, food spoilage mechanisms, and safety risks. As sequencing technologies and

computational methods evolve, AI and ML will become increasingly central to precision food safety, real-time diagnostics, and intelligent quality assurance.

## 5 Conclusion and future research

NGS has transformed food science by enabling culture-independent pathogen detection, AMR surveillance, and microbial community profiling, significantly improving food safety and outbreak prevention. Compared to traditional methods, NGS offers greater sensitivity and a broader scope for detecting pathogens, spoilage organisms, and resistance genes across various food matrices.

In food fermentation, NGS has deepened our understanding of microbial interactions, metabolic pathways, and flavor formation, optimizing starter cultures and enhancing product quality. Similarly, food authentication and traceability have benefited from metabarcoding and metagenomics, improving fraud detection and product integrity.

Despite these advantages, challenges such as high sequencing costs, complex data interpretation, and the need for standardized bioinformatics workflows hinder widespread adoption. Addressing these requires harmonized analytical pipelines, improved computational tools, detailed and clear regulatory guidelines.



Integrating multi-omics approaches and leveraging AI will further enhance microbial analysis, food authenticity verification, and risk assessment.

Future research should focus on advancing real-time, on-site sequencing technologies for rapid pathogen detection in food processing environments. Portable platforms like nanopore sequencing could revolutionize contamination monitoring and outbreak response. Additionally, microbiome-based food safety strategies and predictive modeling using microbial signatures may help prevent contamination and assess food quality. Expanding NGS applications to emerging food technologies, including alternative proteins and novel fermented foods, will be critical for ensuring their safety and quality. Establishing globally accepted regulatory frameworks will further facilitate industry adoption and standardization.

As NGS technology evolves, interdisciplinary collaboration will be key to maximizing its impact on food safety, quality assurance, and innovation, contributing to a safer and more sustainable food system.

## Author contributions

JT-V: Investigation, Visualization, Writing – original draft. BD: Investigation, Writing – original draft. GG: Conceptualization, Writing – review and editing. JC-C: Conceptualization, Writing – review and editing.

## Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

## References

- Abdi, G., Tarighat, M. A., Jain, M., Tendulkar, R., Tendulkar, M., and Barwant, M. (2024). "Revolutionizing genomics: exploring the potential of next-generation sequencing," in *Advances in bioinformatics* (Singapore: Springer Nature Singapore), 1–33. doi:10.1007/978-981-99-8401-5\_1
- Abellan-Schneyder, I., Machado, M. S., Reitmeier, S., Sommer, A., Sewald, Z., Baumbach, J., et al. (2021). Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. *mSphere* 6, e01202-20. doi:10.1128/msphere.01202-20
- Ábrahám, Á., Islam, Md. N., Gazdag, Z., Khan, S. A., Chowdhury, S., Kemenesi, G., et al. (2024). Bacterial metabarcoding of raw palm sap samples from Bangladesh with nanopore sequencing. *Foods* 13, 1285. doi:10.3390/foods13091285
- Afifa, K., Hossain, A., Hossain, Md. S., Kamruzzaman Munshi, M., and Huque, R. (2021). Detection of species adulteration in meat products and Mozzarella-type cheeses using duplex PCR of mitochondrial cyt b gene: a food safety concern in Bangladesh. *Food Chem. Mol. Sci.* 2, 100017. doi:10.1016/j.fochms.2021.100017
- Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., et al. (2023). CARD 2023: expanded curation, support for machine learning, and resistance prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* 51, D690–D699. doi:10.1093/nar/gkac920
- Allard, M. W., Strain, E., Melka, D., Bunning, K., Musser, S. M., Brown, E. W., et al. (2016). Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *J. Clin. Microbiol.* 54, 1975–1983. doi:10.1128/jcm.00081-16
- Alneberg, J., Bjarnason, B. S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nat. Methods* 11, 1144–1146. doi:10.1038/nmeth.3103
- Ambardar, S., Gupta, R., Trakroo, D., Lal, R., and Vakhlu, J. (2016). High throughput sequencing: an overview of sequencing chemistry. *Indian J. Microbiol.* 56, 394–404. doi:10.1007/s12088-016-0606-4
- Anagnostopoulos, D. A., Parlapani, F. F., Mallouchos, A., Angelidou, A., Syropoulou, F., Minos, G., et al. (2022). Volatile organic compounds and 16S metabarcoding in ice-stored red seabream *Pagrus major*. *Foods* 11, 666. doi:10.3390/foods11050666
- Anagnostopoulos, D. A., Syropoulou, F., Parlapani, F. F., Tsiartsafis, A., Exadactylos, A., Nychas, G.-J. E., et al. (2023). Microbiota profile of filleted gilthead seabream (*Sparus aurata*) during storage at various conditions by 16S rRNA metabarcoding analysis. *Food Res. Int.* 164, 112312. doi:10.1016/j.foodres.2022.112312
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Nat. Rev.* 11, R106. doi:10.1186/gb-2010-11-10-r106
- Andrews, S. (2010). FastQC A quality control tool for high throughput sequence data. *Babraham Bioinforma.* Available online at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (Accessed July 23, 2025).
- Antil, S., Abraham, J. S., Sripoorna, S., Maurya, S., Dagar, J., Makhija, S., et al. (2023). DNA barcoding, an effective tool for species identification: a review. *Mol. Biol. Rep.* 50, 761–775. doi:10.1007/s11033-022-08015-7
- Ardui, S., Ameer, A., Vermeesch, J. R., and Hestand, M. S. (2018). Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Res.* 46, 2159–2168. doi:10.1093/nar/gky066
- Bai, Y., Li, J., Huang, M., Yan, S., Li, F., Xu, J., et al. (2024). Prevalence and characterization of foodborne pathogens isolated from fresh-cut fruits and vegetables in Beijing, China. *Int. J. Food Microbiol.* 421, 110804. doi:10.1016/j.ijfoodmicro.2024.110804
- Baker, M., Zhang, X., Maciel-Guerra, A., Dong, Y., Wang, W., Hu, Y., et al. (2023). Machine learning and metagenomics reveal shared antimicrobial resistance profiles across multiple chicken farms and abattoirs in China. *Nat. Food* 4, 707–720. doi:10.1038/s43016-023-00814-w

## Acknowledgments

We gratefully acknowledge Dr. Nou Xiangwu for his valuable suggestions and insightful advice that contributed to the improvement of this work.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Balkir, P., Kemahlioglu, K., and Yucel, U. (2021). Foodomics: a new approach in food quality and safety. *Trends Food Sci. and Technol.* 108, 49–57. doi:10.1016/j.tifs.2020.11.028
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly Algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi:10.1089/cmb.2012.0021
- Banos, S., Lentendu, G., Kopf, A., Wubet, T., Glöckner, F. O., and Reich, M. (2018). A comprehensive fungi-specific 18S rRNA gene sequence primer toolkit suited for diverse research issues and sequencing platforms. *BMC Microbiol.* 18, 190. doi:10.1186/s12866-018-1331-4
- Barcenilla, C., Cobo-Díaz, J. F., Puente, A., Valentino, V., De Filippis, F., Ercolini, D., et al. (2024). In-depth characterization of food and environmental microbiomes across different meat processing plants. *Microbiome* 12, 199. doi:10.1186/s40168-024-01856-3
- Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., et al. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* 10, e65088. doi:10.7554/eLife.65088
- Beier, S., Tappu, R., and Huson, D. H. (2017). “Functional Analysis in Metagenomics using MEGAN 6,” in *Functional metagenomics: tools and applications* (Cham: Springer International Publishing), 65–74. doi:10.1007/978-3-319-61510-3\_4
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170
- Bolinger, H., Tran, D., Harary, K., Paoli, G. C., Guron, G. K. P., Namazi, H., et al. (2021). Utilizing the microbiota and machine learning algorithms to assess risk of salmonella contamination in poultry rinsate. *J. Food Prot.* 84, 1648–1657. doi:10.4315/jfp-20-367
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Author Correction: reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 1091. doi:10.1038/s41587-019-0252-6
- Bonaldo, F., Avot, B. J. P., De Cesare, A., Aarestrup, F. M., and Otani, S. (2023). Foodborne pathogen dynamics in meat and meat analogues analysed using traditional microbiology and metagenomic sequencing. *Antibiotics* 13, 16. doi:10.3390/antibiotics13010016
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi:10.1038/nmeth.3176
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., and Holmes, S. P. (2015). DADA2: high resolution sample inference from amplicon data. doi:10.1101/024034
- Cao, Y., Fanning, S., Proos, S., Jordan, K., and Srikumar, S. (2017). A review on the applications of next generation sequencing technologies as applied to food-related microbiome studies. *Front. Microbiol.* 8, 1829. doi:10.3389/fmicb.2017.01829
- Cardinali, G., Corte, L., and Robert, V. (2017). Next generation sequencing: problems and opportunities for next generation studies of microbial communities in food and food industry. *Curr. Opin. Food Sci.* 17, 62–67. doi:10.1016/j.cofs.2017.09.009
- Carugati, L., Corinaldesi, C., Dell’Anno, A., and Danovaro, R. (2015). Metagenetic tools for the census of marine meiofaunal biodiversity: an overview. *Mar. Genomics* 24, 11–20. doi:10.1016/j.margen.2015.04.010
- Chen, S. (2023). Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* 2, e107. doi:10.1002/imt2.107
- Chen, P., Sun, Z., Wang, J., Liu, X., Bai, Y., Chen, J., et al. (2023). Portable nanopore-sequencing technology: trends in development and applications. *Front. Microbiol.* 14, 1043967. doi:10.3389/fmicb.2023.1043967
- Constante Catuto, M. P., Tigrero-Vaca, J., Villavicencio-Vasquez, M., Montoya, D. C., Cevallos, J. M., and Coronel-León, J. (2024). Evaluation of stress tolerance and design of alternative culture media for the production of fermentation starter cultures in cacao. *Heliyon* 10, e29900. doi:10.1016/j.heliyon.2024.e29900
- De Filippis, F., Laiola, M., Blaiotta, G., and Ercolini, D. (2017). Different amplicon targets for sequencing-based studies of fungal diversity. *Appl. Environ. Microbiol.* 83, e00905-17. doi:10.1128/AEM.00905-17
- Detcharoen, M., Khruaekaw, P., Sukkapat, P., Benjakul, S., and Saetang, J. (2024). Metabarcoding for authentication of fish species in surimi-based products by Nanopore sequencing. *Food Biosci.* 61, 104628. doi:10.1016/j.fbio.2024.104628
- Díaz Cárdenas, B., Salazar Llorente, E., Gu, G., Nou, X., Ortiz, J., Maldonado, P., et al. (2024). Microbial composition and diversity of high-demand street-vended foods in Ecuador. *J. Food Prot.* 87, 100247. doi:10.1016/j.jfp.2024.100247
- Ding, Q., Gu, G., Luo, Y., Nou, X., and Micallef, S. A. (2025). Transcriptomic response of *Escherichia coli* O157:H7 on Romaine lettuce from harvest to storage during the pre-processing interval. *Postharvest Biol. Technol.* 227, 113594. doi:10.1016/j.postharvbio.2025.113594
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635
- Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2020). PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 38, 685–688. doi:10.1038/s41587-020-0548-6
- Emaus, M. N., Varona, M., Eitzmann, D. R., Hsieh, S.-A., Zeger, V. R., and Anderson, J. L. (2020). Nucleic acid extraction: fundamentals of sample preparation methodologies, current advancements, and future endeavors. *TrAC Trends Anal. Chem.* 130, 115985. doi:10.1016/j.trac.2020.115985
- Espinosa, E., Bautista, R., Larrosa, R., and Plata, O. (2024). Advancements in long-read genome sequencing technologies and algorithms. *Genomics* 116, 110842. doi:10.1016/j.ygeno.2024.110842
- Fagerlund, A., Langsrud, S., and Møretro, T. (2021). Microbial diversity and ecology of biofilms in food industry environments associated with *Listeria monocytogenes* persistence. *Curr. Opin. Food Sci.* 37, 171–178. doi:10.1016/j.cofs.2020.10.015
- Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., et al. (2021). AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.* 11, 12728. doi:10.1038/s41598-021-91456-0
- Feng, S., Eucker, T. P., Holly, M. K., Konkel, M. E., Lu, X., and Wang, S. (2014). Investigating the responses of *Cronobacter sakazakii* to garlic-derived organosulfur compounds: a systematic study of pathogenic-bacterium injury by use of high-throughput whole-transcriptome sequencing and Confocal Micro-Raman spectroscopy. *Appl. Environ. Microbiol.* 80, 959–971. doi:10.1128/aem.03460-13
- Ferrara, M., Sisto, A., Mulè, G., Lavermicocca, P., and De Bellis, P. (2021). Metagenetic analysis for microbial characterization of Focaccia doughs obtained by using two different starters: traditional Baker’s yeast and a selected *Leuconostoc citreum* strain. *Foods* 10, 1189. doi:10.3390/foods10061189
- Ferrocino, I., Rantsiou, K., McClure, R., Kostic, T., De Souza, R. S. C., Lange, L., et al. (2023). The need for an integrated multi-OMICs approach in microbiome science in the food system. *Comp. Rev. Food Sci. Food Safe* 22, 1082–1103. doi:10.1111/1541-4337.13103
- Florensa, A. F., Kaas, R. S., Clausen, P. T. L. C., Aytan-Aktug, D., and Aarestrup, F. M. (2022). ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microb. Genomics* 8, 000748. doi:10.1099/mgen.0.000748
- Forbes, J. D., Knox, N. C., Ronholm, J., Pagotto, F., and Reimer, A. (2017). Metagenomics: the next culture-independent game changer. *Front. Microbiol.* 8, 1069. doi:10.3389/fmicb.2017.01069
- Francioli, D., Lentendu, G., Lewin, S., and Kolb, S. (2021). DNA metabarcoding for the characterization of terrestrial Microbiota—Pitfalls and solutions. *Microorganisms* 9, 361. doi:10.3390/microorganisms9020361
- Fricker, A. M., Podlesny, D., and Fricke, W. F. (2019). What is new and relevant for sequencing-based microbiome research? A mini-review. *J. Adv. Res.* 19, 105–112. doi:10.1016/j.jare.2019.03.006
- Gaire, T. N., Odland, C., Zhang, B., Slizovskiy, I., Jorgenson, B., Wehri, T., et al. (2024). Slaughtering processes impact microbial communities and antimicrobial resistance genes of pig carcasses. *Sci. Total Environ.* 946, 174394. doi:10.1016/j.scitotenv.2024.174394
- Giusti, A., Spatola, G., Mancini, S., Nuvoloni, R., and Armani, A. (2024). Novel foods, old issues: metabarcoding revealed mislabeling in insect-based products sold by e-commerce on the EU market. *Food Res. Int.* 184, 114268. doi:10.1016/j.foodres.2024.114268
- Glenn, T. C., Pierson, T. W., Bayona-Vásquez, N. J., Kieran, T. J., Hoffberg, S. L., Thomas IV, J. C., et al. (2019). Adapterama II: universal amplicon sequencing on Illumina platforms (TaggMatrix). *PeerJ* 7, e7786. doi:10.7717/peerj.7786
- Goecks, J., Nekrutenko, A., Taylor, J., and The Galaxy Team (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* 11, R86. doi:10.1186/gb-2010-11-8-r86
- Gong, Y., Zhao, L., Karwe, M., and Tan, J. (2024). Effects of processed oats and pinto beans on the composition of human gut microbiome: an *in vitro* study. *Bioact. Carbohydrates Diet. Fibre* 32, 100434. doi:10.1016/j.bcdf.2024.100434
- Goodwin, S., McPherson, J. D., and McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* 17, 333–351. doi:10.1038/nrg.2016.49
- Grazina, L., Amaral, J. S., and Mafra, I. (2020). Botanical origin authentication of dietary supplements by DNA-based approaches. *Comp. Rev. Food Sci. Food Safe* 19, 1080–1109. doi:10.1111/1541-4337.12551
- Gu, G., Ding, Q., Redding, M., Yang, Y., O’Brien, R., Gu, T., et al. (2024). Differential microbiota shift on whole romaine lettuce subjected to source or forward processing and on fresh-cut products during cold storage. *Int. J. Food Microbiol.* 416, 110665. doi:10.1016/j.jfoodmicro.2024.110665
- Guenay-Greunke, Y., Bohan, D. A., Traugott, M., and Wallinger, C. (2021). Handling of targeted amplicon sequencing data focusing on index hopping and demultiplexing using a nested metabarcoding approach in ecology. *Sci. Rep.* 11, 19510. doi:10.1038/s41598-021-98018-4

- Güley, Z., Fallico, V., Cabrera-Rubio, R., O'Sullivan, D., Marotta, M., Pennone, V., et al. (2023). Diversity of the microbiota of traditional Izmir tulum and Izmir brined tulum cheeses and selection of potential probiotics. *Foods* 12, 3482. doi:10.3390/foods12183482
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi:10.1038/nprot.2013.084
- Habib, I., Mohamed, M.-Y. I., Lakshmi, G. B., Ghazawi, A., Khan, M., Abdalla, A., et al. (2024). High prevalence and genomic features of multidrug-resistant *Salmonella enterica* isolated from chilled broiler chicken on retail sale in the United Arab Emirates. *Int. J. Food Microbiol.* 423, 110828. doi:10.1016/j.ijfoodmicro.2024.110828
- Head, S. R., Komori, H. K., LaMere, S. A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D. R., et al. (2014). Library construction for next-generation sequencing: overviews and challenges. *BioTechniques* 56, 61–77. doi:10.2144/000114133
- Heather, J. M., and Chain, B. (2016). The sequence of sequencers: the history of sequencing DNA. *Genomics* 107, 1–8. doi:10.1016/j.ygeno.2015.11.003
- Hilt, E. E., and Ferrieri, P. (2022). Next generation and other sequencing technologies in diagnostic microbiology and infectious diseases. *Genes* 13, 1566. doi:10.3390/genes13091566
- Hu, T., Chitnis, N., Monos, D., and Dinh, A. (2021). Next-generation sequencing technologies: an overview. *Hum. Immunol.* 82, 801–811. doi:10.1016/j.humimm.2021.02.012
- Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., Von Mering, C., et al. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol. Biol. Evol.* 34, 2115–2122. doi:10.1093/molbev/msx148
- Hunt, M., Mather, A. E., Sánchez-Busó, L., Page, A. J., Parkhill, J., Keane, J. A., et al. (2017). ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb. Genomics* 3, e000131. doi:10.1099/mgen.0.000131
- Ince, V., Bader-El-Den, M., Alderton, J., Arabikhan, F., Sari, O. F., and Sansom, A. (2025). Machine learning-based prediction of *Clostridium* growth in pork meat using explainable artificial intelligence. *J. Food Sci. Technol.* doi:10.1007/s13197-024-06187-7
- Jagadeesan, B., Gerner-Smidt, P., Allard, M. W., Leuillet, S., Winkler, A., Xiao, Y., et al. (2019). The use of next generation sequencing for improving food safety: translation into practice. *Food Microbiol.* 79, 96–115. doi:10.1016/j.fm.2018.11.005
- Jia, J., Li, C., Tang, M., Liao, H., Yang, Z., Wang, Y., et al. (2025). Mitochondrial genome characterization, evolution and intron dynamics of the entomopathogenic genus *Cordyceps*. *Front. Microbiol.* 16, 1605218. doi:10.3389/fmicb.2025.1605218
- Kanehisa, M., and Sato, Y. (2020). KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci.* 29, 28–35. doi:10.1002/pro.3711
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361. doi:10.1093/nar/gkw1092
- Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7, e7359. doi:10.7717/peerj.7359
- Kazantseva, J., Malv, E., Kaleda, A., Kallastu, A., and Meikas, A. (2021). Optimisation of sample storage and DNA extraction for human gut microbiota studies. *BMC Microbiol.* 21, 158. doi:10.1186/s12866-021-02233-y
- Keegan, K. P., Glass, E. M., and Meyer, F. (2016). “MG-RAST, a metagenomics service for analysis of microbial community structure and function,” in *Methods in molecular biology* (New York, NY: Springer), 207–233. doi:10.1007/978-1-4939-3369-3\_13
- Kharnaior, P., and Tamang, J. P. (2022). Metagenomic-Metabolomic mining of Kinema, a naturally fermented soybean food of the Eastern himalayas. *Front. Microbiol.* 13, 868383. doi:10.3389/fmicb.2022.868383
- Kim, H., Jebail, M. J., Sinha, A., Bent, Z. W., Solberg, O. D., Williams, K. P., et al. (2013). A microfluidic DNA Library preparation platform for next-generation sequencing. *PLoS ONE* 8, e68988. doi:10.1371/journal.pone.0068988
- Kim, K. H., Chun, B. H., Kim, J., and 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. doi:10.1016/j.foodcont.2020.107681
- Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* 37, 540–546. doi:10.1038/s41587-019-0072-8
- Kopylova, E., Noé, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28, 3211–3217. doi:10.1093/bioinformatics/bts611
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi:10.1101/gr.215087.116
- Koutsoumanis, K., Bolton, D., Bover-Cid, S., Chemaly, M., De Cesare, A., Herman, L., et al. (2024). Persistence of microbiological hazards in food and feed production and processing environments. *EFSA J.* 22, e8521. doi:10.2903/j.efsa.2024.8521
- Kragh, M. L., and Truelstrup Hansen, L. (2020). Initial transcriptomic response and adaption of *Listeria monocytogenes* to desiccation on food grade stainless steel. *Front. Microbiol.* 10, 3132. doi:10.3389/fmicb.2019.03132
- Kuan, S., Chin, N. L., Tee, T. P., and Hasnan, N. Z. N. (2025). Microbiome diversity in seafood factories via next-generation sequencing for Food Safety Management System (FSMS) certifications in Malaysia. *Foods* 14, 1517. doi:10.3390/foods14091517
- Küçükgöz, K., Venema, K., Chamorro, F., Cassani, L., Donn, P., Prieto, M. A., et al. (2025). Unlocking the potential of fermented beetroot ketchup: enhancing polyphenol recovery and gut microbiota interactions. *Food Chem.* 463, 141141. doi:10.1016/j.foodchem.2024.141141
- Kulski, J. K. (2016). “Next-Generation sequencing — an overview of the history, tools, and ‘Omic’ applications,” in *Next generation sequencing - advances, applications and challenges* (London, United Kingdom: InTech). doi:10.5772/61964
- Kumar, A., Pratush, A., and Bera, S. (2021). “Significance of nanoscience in food microbiology: current trend and future prospects,” in *Environmental and microbial biotechnology* (Singapore: Springer Nature Singapore), 249–267. doi:10.1007/978-981-15-9916-3\_10
- Kurniawan, Y. N., Shinohara, Y., Takesue, N., Sakai, H., Magarifuchi, T., and Suzuki, K. (2021). Development of a rapid and accurate Nanopore-based sequencing platform for on-field identification of beer-spoilage bacteria in the breweries. *J. Am. Soc. Brew. Chem.* 79, 240–248. doi:10.1080/03610470.2021.1904491
- Kwon, H. J., Chen, Z., Evans, P., Meng, J., and Chen, Y. (2020). Characterization of Mobile genetic elements using long-read sequencing for tracking *Listeria monocytogenes* from food processing environments. *Pathogens* 9, 822. doi:10.3390/pathogens9100822
- La Reau, A. J., Strom, N. B., Filvaroff, E., Mavrommatis, K., Ward, T. L., and Knights, D. (2023). Shallow shotgun sequencing reduces technical variation in microbiome analysis. *Sci. Rep.* 13, 7668. doi:10.1038/s41598-023-33489-1
- Lamb, H. J., Nguyen, L. T., Briody, T. E., Ambrose, R. K., Hayes, B. J., Mahony, T. J., et al. (2023). Skim-Nanopore sequencing for routine genomic evaluation and bacterial pathogen detection in cattle. *Anim. Prod. Sci.* 63, 1074–1085. doi:10.1071/an22451
- Lear, G., Dickie, I., Banks, J., Boyer, S., Buckley, H., Buckley, T., et al. (2018). Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. *NZ.* doi:10.20417/nzjcol.42.9
- Lee, J. Y., Kong, M., Oh, J., Lim, J., Chung, S. H., Kim, J.-M., et al. (2021). Comparative evaluation of Nanopore polishing tools for microbial genome assembly and polishing strategies for downstream analysis. *Sci. Rep.* 11, 20740. doi:10.1038/s41598-021-00178-w
- Lee, A. W.-T., Ng, I. C.-F., Wong, E. Y.-K., Wong, I. T.-F., Sze, R. P.-P., Chan, K.-Y., et al. (2024). Comprehensive identification of pathogenic microbes and antimicrobial resistance genes in food products using nanopore sequencing-based metagenomics. *Food Microbiol.* 121, 104493. doi:10.1016/j.fm.2024.104493
- Lerma-Aguilera, A. M., Pérez-Burillo, S., Navajas-Porras, B., León, E. D., Ruiz-Pérez, S., Pastoriza, S., et al. (2024). Effects of different foods and cooking methods on the gut microbiota: an *in vitro* approach. *Front. Microbiol.* 14, 1334623. doi:10.3389/fmicb.2023.1334623
- Lever, M. A., Torti, A., Eickenbusch, P., Michaud, A. B., Å anti-Temkiv, T., and Jä rgensen, B. B. (2015). A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types. *Front. Microbiol.* 6, 476. doi:10.3389/fmicb.2015.00476
- Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi:10.1093/bioinformatics/btv033
- Li, S., Han, Y., Yan, M., Qiu, S., and Lu, J. (2025). Machine learning and multi-omics integration to reveal biomarkers and microbial community assembly differences in abnormal stacking fermentation of sauce-flavor baijiu. *Foods* 14, 245. doi:10.3390/foods14020245
- Liang, L., Wang, P., Zhao, X., He, L., Qu, T., and Chen, Y. (2022). Single-molecule real-time sequencing reveals differences in bacterial diversity in raw milk in different regions and seasons in China. *J. Dairy Sci.* 105, 5669–5684. doi:10.3168/jds.2021-21445
- Liao, C., Wang, L., and Quon, G. (2024). Microbiome-based classification models for fresh produce safety and quality evaluation. *Microbiol. Spectr.* 12, e03448-23. doi:10.1128/spectrum.03448-23
- Lima, C. O. D. C., De Castro, G. M., Solar, R., Vaz, A. B. M., Lobo, F., Pereira, G., et al. (2022). Unraveling potential enzymes and their functional role in fine cocoa beans fermentation using temporal shotgun metagenomics. *Front. Microbiol.* 13, 994524. doi:10.3389/fmicb.2022.994524
- Lindner, B. G., Gerhardt, K., Feistel, D. J., Rodriguez-R, L. M., Hatt, J. K., and Konstantinidis, K. T. (2024). A user's guide to the bioinformatic analysis of shotgun metagenomic sequence data for bacterial pathogen detection. *Int. J. Food Microbiol.* 410, 110488. doi:10.1016/j.ijfoodmicro.2023.110488
- Lira, A. D., Kothe, C. I., Rué, O., Midoux, C., Mann, M. B., Mallmann, L. P., et al. (2020). Assessing the quality of fresh Whitemouth croaker (*Micropogonias furnieri*) meat based on micro-organism and histamine analysis using NGS, qPCR and HPLC-DAD. *J. Appl. Microbiol.* 128, 1448–1459. doi:10.1111/jam.14566



- Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., and Burrige, C. P. (2020). A practical guide to DNA metabarcoding for entomological ecologists. *Ecol. Entomol.* 45, 373–385. doi:10.1111/een.12831
- Lizarazo, E., Couto, N., Vincenti-Gonzalez, M., Raangs, E. C., Velasco, Z., Bethencourt, S., et al. (2019). Applied shotgun metagenomics approach for the genetic characterization of dengue viruses. *J. Biotechnol.* 306, 100009. doi:10.1016/j.biotech.2019.100009
- Loman, N. J., and Pallen, M. J. (2015). Twenty years of bacterial genome sequencing. *Nat. Rev. Microbiol.* 13, 787–794. doi:10.1038/nrmicro3565
- Louca, S., Parfrey, L. W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Sci.* (1979). 353, 1272–1277. doi:10.1126/science.aaf4507
- Ma, H., Li, J., Xi, X., Xu, H., Wuri, L., Bian, Y., et al. (2018). Evaluation of bacterial contamination in Goat milk powder using PacBio single molecule real-time sequencing and droplet digital PCR. *J. Food Prot.* 81, 1791–1799. doi:10.4315/0362-028x.jfp-17-535
- MacConaill, L. E., Burns, R. T., Nag, A., Coleman, H. A., Slevin, M. K., Giorda, K., et al. (2018). Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. *BMC Genomics* 19, 30. doi:10.1186/s12864-017-4428-5
- MacKenzie, M., and Argyropoulos, C. (2023). An introduction to nanopore sequencing: past, present, and future considerations. *Micromachines* 14, 459. doi:10.3390/mi14020459
- Madoui, M.-A., Engelen, S., Cruaud, C., Belser, C., Bertrand, L., Alberti, A., et al. (2015). Genome assembly using Nanopore-guided long and error-free DNA reads. *BMC Genomics* 16, 327. doi:10.1186/s12864-015-1519-z
- Mak, T.-M., Stegger, M., Mølbak, K., and Maurer-Stroh, S. (2025). “Pathogen genomics for surveillance and outbreak investigations,” in *Infectious disease emergencies: preparedness and response* (Singapore: NUS Press), 270–288. doi:10.56159/emergencies-21
- Marchelli, R., Tedeschi, T., and Tonelli, A. (2012). “DNA analyses in food safety and quality: current status and expectations,” in *Detection of non-amplified genomic DNA* (Dordrecht: Springer). doi:10.1007/978-94-007-1226-3
- Martínez-Álvarez, S., Châtre, P., François, P., Zarazaga, M., Madec, J.-Y., Haenni, M., et al. (2025). Comparative phylogenomics of extended-spectrum beta-lactamase-producing *Escherichia coli* revealed a wide diversity of clones and plasmids in Spanish chicken meat. *Int. J. Food Microbiol.* 426, 110900. doi:10.1016/j.ijfoodmicro.2024.110900
- Menzel, P., Ng, K. L., and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7, 11257. doi:10.1038/ncomms11257
- Milani, C., Longhi, G., Alessandri, G., Fontana, F., Viglioli, M., Tarracchini, C., et al. (2025). Functional modulation of the human gut microbiome by bacteria vectored by cheese. *Appl. Environ. Microbiol.* 91, e00180-25. doi:10.1128/aem.00180-25
- Mota-Gutiérrez, J., Ferrocino, I., Rantsiou, K., and Cocolin, L. (2019). Metataxonomic comparison between internal transcribed spacer and 26S ribosomal large subunit (LSU) rDNA gene. *Int. J. Food Microbiol.* 290, 132–140. doi:10.1016/j.ijfoodmicro.2018.10.010
- Mottola, A., Intermite, C., Piredda, R., Lorusso, L., Ranieri, L., Carpino, S., et al. (2024). DNA metabarcoding approach as a potential tool for supporting official food control programs: a case study. *Foods* 13, 2941. doi:10.3390/foods13182941
- Muller, B. H., Mollon, P., Santiago-Allexant, E., Javerliat, F., and Kaneko, G. (2019). In-depth comparison of library pooling strategies for multiplexing bacterial species in NGS. *Diagnostic Microbiol. Infect. Dis.* 95, 28–33. doi:10.1016/j.diagmicrobio.2019.04.014
- Murray, M., Salvatierra, G., Dávila-Barclay, A., Ayzanoa, B., Castillo-Vilcahuaman, C., Huang, M., et al. (2021). Market chickens as a source of antibiotic-resistant *Escherichia coli* in a peri-urban community in Lima, Peru. *Front. Microbiol.* 12, 635871. doi:10.3389/fmicb.2021.635871
- Na, H., Kim, Y., Kim, D., Yoon, H., and Ryu, S. (2018). Transcriptomic analysis of Shiga toxin-producing *Escherichia coli* FORC\_035 reveals the essential role of iron acquisition for survival in Canola sprouts and water dropwort. *Front. Microbiol.* 9, 2397. doi:10.3389/fmicb.2018.02397
- Nam, N., Do, H., Loan Trinh, K., and Lee, N. (2023). Metagenomics: an effective approach for exploring microbial diversity and functions. *Foods* 12, 2140. doi:10.3390/foods12112140
- Neal-McKinney, J. M., Liu, K. C., Lock, C. M., Wu, W.-H., and Hu, J. (2021). Comparison of MiSeq, MinION, and hybrid genome sequencing for analysis of *Campylobacter jejuni*. *Sci. Rep.* 11, 5676. doi:10.1038/s41598-021-84956-6
- Oikonomopoulos, S., Bayega, A., Fahiminiya, S., Djambazian, H., Berube, P., and Ragoussis, J. (2020). Methodologies for transcript profiling using long-read technologies. *Front. Genet.* 11, 606. doi:10.3389/fgene.2020.00606
- Ojala, T., Kankuri, E., and Kankainen, M. (2023). Understanding human health through metatranscriptomics. *Trends Mol. Med.* 29, 376–389. doi:10.1016/j.molmed.2023.02.002
- Orakov, A., Fullam, A., Coelho, L. P., Khedkar, S., Szklarczyk, D., Mende, D. R., et al. (2021). GUNC: detection of chimerism and contamination in prokaryotic genomes. *Genome Biol.* 22, 178. doi:10.1186/s13059-021-02393-0
- Ottesen, A. R., and Ramachandran, P. (2019). “Food microbiomes: a new paradigm for food and food ecology,” in *Food microbiology* (Washington, DC, USA: ASM Press), 963–970. doi:10.1128/9781555819972.ch37
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi:10.1093/bioinformatics/btv421
- Pan, Y., Wang, Y., Hao, W., Duan, C., Wang, S., Wei, J., et al. (2022). Metatranscriptomics unravel composition, drivers, and functions of the active microorganisms in light-flavor liquor fermentation. *Microbiol. Spectr.* 10, e02151-21. doi:10.1128/spectrum.02151-21
- Panwar, S., Duggirala, K. S., Yadav, P., Debnath, N., Yadav, A. K., and Kumar, A. (2023). Advanced diagnostic methods for identification of bacterial foodborne pathogens: contemporary and upcoming challenges. *Crit. Rev. Biotechnol.* 43, 982–1000. doi:10.1080/07388551.2022.2095253
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi:10.1101/gr.186072.114
- Pennone, V., Cobo-Díaz, J. F., Prieto, M., and Alvarez-Ordóñez, A. (2022). Application of genomics and metagenomics to improve food safety based on an enhanced characterisation of antimicrobial resistance. *Curr. Opin. Food Sci.* 43, 183–188. doi:10.1016/j.cofs.2021.12.002
- Pothakos, V., De Vuyst, L., Zhang, S. J., De Bruyn, F., Verce, M., Torres, J., et al. (2020). Temporal shotgun metagenomics of an Ecuadorian coffee fermentation process highlights the predominance of lactic acid bacteria. *Curr. Res. Biotechnol.* 2, 1–15. doi:10.1016/j.crbiot.2020.02.001
- Pregolini, V. B., De Melo Pereira, G. V., Da Silva Vale, A., De Carvalho Neto, D. P., and Soccol, C. R. (2021). Influence of environmental microbiota on the activity and metabolism of starter cultures used in coffee beans fermentation. *Fermentation* 7, 278. doi:10.3390/fermentation7040278
- Qi, Z., Zhou, X., Tian, L., Zhang, H., Cai, L., and Tang, F. (2022). Temporal and spatial variation of microbial communities in stored rice grains from two major depots in China. *Food Res. Int.* 152, 110876. doi:10.1016/j.foodres.2021.110876
- Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., and Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* 35, 833–844. doi:10.1038/nbt.3935
- Rantsiou, K., Kathariou, S., Winkler, A., Skandamis, P., Saint-Cyr, M. J., Rouzeau-Szynalski, K., et al. (2018). Next generation microbiological risk assessment: opportunities of whole genome sequencing (WGS) for foodborne pathogen surveillance, source tracking and risk assessment. *Int. J. Food Microbiol.* 287, 3–9. doi:10.1016/j.ijfoodmicro.2017.11.007
- Redding, M., Zheng, J., Mowery, J., Gu, G., Bolten, S., Luo, Y., et al. (2024). Microscopic and transcriptomic characterization of *Listeria monocytogenes* aggregation and biofilm formation in cantaloupe juice. *Food control* 158, 110243. doi:10.1016/j.foodcont.2023.110243
- Reid, C. J., Blau, K., Jechalke, S., Smalla, K., and Djordjevic, S. P. (2020). Whole genome sequencing of *Escherichia coli* from store-bought produce. *Front. Microbiol.* 10, 3050. doi:10.3389/fmicb.2019.03050
- Ribarska, T., Bjørnstad, P. M., Sundaram, A. Y. M., and Gilfillan, G. D. (2022). Optimization of enzymatic fragmentation is crucial to maximize genome coverage: a comparison of library preparation methods for Illumina sequencing. *BMC Genomics* 23, 92. doi:10.1186/s12864-022-08316-y
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi:10.1093/bioinformatics/btp616
- Ruggieri, J., Kemp, R., Forman, S., and Van Eden, M. E. (2016). “Techniques for nucleic acid purification from plant, animal, and microbial samples,” in *Springer protocols handbooks* (New York, NY: Springer New York), 41–52. doi:10.1007/978-1-4939-3185-9\_4
- Sabater, C., Calvete, I., Vázquez, X., Ruiz, L., and Margolles, A. (2024). Tracing the origin and authenticity of Spanish PDO honey using metagenomics and machine learning. *Int. J. Food Microbiol.* 421, 110789. doi:10.1016/j.ijfoodmicro.2024.110789
- Sadurski, J., Polak-Berecka, M., Staniszeński, A., and Wąsko, A. (2024). Step-by-Step metagenomics for food microbiome analysis: a detailed review. *Foods* 13, 2216. doi:10.3390/foods13142216
- Sajali, N., Wong, S. C., Hanapi, U. K., Abu Bakar @ Jamaluddin, S., Tasrip, N. A., and Mohd Desa, M. N. (2018). The challenges of DNA extraction in different assorted food matrices: a review. *J. Food Sci.* 83, 2409–2414. doi:10.1111/1750-3841.14338
- Saliba, A.-E., Westermann, A. J., Gorski, S. A., and Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 42, 8845–8860. doi:10.1093/nar/gku555
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., et al. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 12, 87. doi:10.1186/s12915-014-0087-z
- Santamaria, M., Fosso, B., Consiglio, A., De Caro, G., Grillo, G., Licciulli, F., et al. (2012). Reference databases for taxonomic assignment in metagenomics. *Briefings Bioinforma.* 13, 682–695. doi:10.1093/bib/bbs036



- Schenk, J. J., Becklund, L. E., Carey, S. J., and Fabre, P. P. (2023). What is the “modified” CTAB protocol? Characterizing modifications to the CTAB DNA extraction protocol. *Appl. Plant Sci.* 11, e11517. doi:10.1002/aps3.11517
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: Open-Source, Platform-Independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi:10.1128/aem.01541-09
- Schwengers, O., Jelonek, L., Dieckmann, M. A., Beyvers, S., Blom, J., and Goesmann, A. (2021). Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb. Genomics* 7, 000685. doi:10.1099/mgen.0.000685
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi:10.1093/bioinformatics/btu153
- Segata, N., Waldron, L., Ballarín, A., Narasimhan, V., Jousson, O., and Huttenhower, C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* 9, 811–814. doi:10.1038/nmeth.2066
- Sekse, C., Holst-Jensen, A., Dobrindt, U., Johannessen, G. S., Li, W., Spilberg, B., et al. (2017). High throughput sequencing for detection of foodborne pathogens. *Front. Microbiol.* 8, 2029. doi:10.3389/fmicb.2017.02029
- Sequino, G., Cobo-Díaz, J. F., Valentino, V., Tassou, C., Volpe, S., Torrieri, E., et al. (2024). Microbiome mapping in beef processing reveals safety-relevant variations in microbial diversity and genomic features. *Food Res. Int.* 186, 114318. doi:10.1016/j.foodres.2024.114318
- Sergeant, M. J., Constantinidou, C., Cogan, T., Penn, C. W., and Pallen, M. J. (2012). High-throughput sequencing of 16S rRNA gene amplicons: effects of extraction procedure, primer length and annealing temperature. *PLoS ONE* 7, e38094. doi:10.1371/journal.pone.0038094
- Shinohara, Y., Kurniawan, Y. N., Sakai, H., Magarifuchi, T., and Suzuki, K. (2021). Nanopore based sequencing enables easy and accurate identification of yeasts in breweries. *J. Inst. Brew.* 127, 160–166. doi:10.1002/jib.639
- Srinivas, M., O’Sullivan, O., Cotter, P. D., Sinderen, D. V., and Kenny, J. G. (2022). The application of metagenomics to study microbial communities and develop desirable traits in fermented foods. *Foods* 11, 3297. doi:10.3390/foods11203297
- Syromyatnikov, M. Y., Kokina, A. V., Solodskikh, S. A., Panevina, A. V., Popov, E. S., and Popov, V. N. (2020). High-Throughput 16S rRNA gene sequencing of butter Microbiota reveals a variety of opportunistic pathogens. *Foods* 9, 608. doi:10.3390/foods9050608
- Taherdoost, H. (2016). Sampling methods in research methodology; how to choose a sampling technique for research. *SSRN J.* doi:10.2139/ssrn.3205035
- Tamang, J. P., Das, S., Kharnaier, P., Pariyar, P., Thapa, N., Jo, S.-W., et al. (2022). Shotgun metagenomics of Cheonggukjang, a fermented soybean food of Korea: community structure, predictive functionalities and amino acids profile. *Food Res. Int.* 151, 110904. doi:10.1016/j.foodres.2021.110904
- Tan, S. C., and Yiap, B. C. (2009). DNA, RNA, and protein extraction: the past and the present. *BioMed Res. Int.* 2009, 574398. doi:10.1155/2009/574398
- Tan, X., Chung, T., Chen, Y., Macarasin, D., LaBorde, L., and Kovac, J. (2019). The occurrence of *Listeria monocytogenes* is associated with built environment microbiota in three tree fruit processing facilities. *Microbiome* 7, 115. doi:10.1186/s40168-019-0726-2
- Tang, Y., Li, F., Gu, D., Wang, W., Huang, J., and Jiao, X. (2021). Antimicrobial effect and the mechanism of diallyl trisulfide against *Campylobacter jejuni*. *Antibiotics* 10, 246. doi:10.3390/antibiotics10030246
- Tedersoo, L., Abarenkov, K., Nilsson, R. H., Schüssler, A., Grelet, G. A., Kohout, P., et al. (2011). Tidying up international Nucleotide sequence databases: ecological, geographical and sequence quality annotation of ITS sequences of mycorrhizal fungi. *PLoS ONE* 6, e24940. doi:10.1371/journal.pone.0024940
- Tigrero-Vaca, J., Maridueña-Zavala, M. G., Liao, H.-L., Prado-Lince, M., Zambrano-Vera, C. S., Monserrate-Maggi, B., et al. (2022). Microbial diversity and contribution to the Formation of volatile compounds during fine-flavor Cacao Bean Fermentation. *Foods* 11, 915. doi:10.3390/foods11070915
- Tremblay, J., Schreiber, L., and Greer, C. W. (2022). High-resolution shotgun metagenomics: the more data, the better? *Briefings Bioinforma.* 23, bbac443. doi:10.1093/bib/bbac443
- Tripathi, V., Kumar, P., Tripathi, P., Kishore, A., and Kamle, M. (2019). *Microbial genomics in sustainable agroecosystems* (Singapore: Springer Singapore), 2. doi:10.1007/978-981-32-9860-6
- Tutus, R., Kahraman, A., and Memon, A. (2025). Comparative phylogenetic analysis of key chloroplast and nuclear DNA regions of some Brassica species. *Genet. Resour. Crop Evol.* doi:10.1007/s10722-025-02458-0
- Tyagi, A., Singh, B., Billekallu Thammegowda, N. K., and Singh, N. K. (2019). Shotgun metagenomics offers novel insights into taxonomic compositions, metabolic pathways and antibiotic resistance genes in fish gut microbiome. *Arch. Microbiol.* 201, 295–303. doi:10.1007/s00203-018-1615-y
- Vaccalluzzo, A., Celano, G., Pino, A., Calabrese, F. M., Foti, P., Caggia, C., et al. (2022). Metagenetic and volatilomic approaches to elucidate the effect of *Lactiplantibacillus plantarum* starter cultures on Sicilian table olives. *Front. Microbiol.* 12, 771636. doi:10.3389/fmicb.2021.771636
- Van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., and Thermes, C. (2018). The third revolution in sequencing technology. *Trends Genet.* 34, 666–681. doi:10.1016/j.tig.2018.05.008
- Vishnuraj, M. R., Aravind Kumar, N., Vaithyanathan, S., and Barbuddhe, S. B. (2023). Authentication issues in foods of animal origin and advanced molecular techniques for identification and vulnerability assessment. *Trends Food Sci. and Technol.* 138, 164–177. doi:10.1016/j.tifs.2023.05.019
- Walsh, A. M., Macori, G., Kilcawley, K. N., and Cotter, P. D. (2020). Meta-analysis of cheese microbiomes highlights contributions to multiple aspects of quality. *Nat. Food* 1, 500–510. doi:10.1038/s43016-020-0129-3
- Wang, Q., Han, Y.-Y., Zhang, T.-J., Chen, X., Lin, H., Wang, H.-N., et al. (2024). Whole-genome sequencing of *Escherichia coli* from retail meat in China reveals the dissemination of clinically important antimicrobial resistance genes. *Int. J. Food Microbiol.* 415, 110634. doi:10.1016/j.ijfoodmicro.2024.110634
- Wen, G. (2017). “A simple process of RNA-Sequence analyses by Hisat2, Htseq and DESeq2,” in *Proceedings of the 2017 international conference on biomedical engineering and bioinformatics* (New York, NY: ACM), 11–15. doi:10.1145/3143344.3143354
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13, e1005595. doi:10.1371/journal.pcbi.1005595
- Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biol.* 20, 257. doi:10.1186/s13059-019-1891-0
- Wu, Y.-W., Tang, Y.-H., Tringe, S. G., Simmons, B. A., and Singer, S. W. (2014). MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* 2, 26. doi:10.1186/2049-2618-2-26
- Xiao, M., Huang, T., Xu, Y., Peng, Z., Liu, Z., Guan, Q., et al. (2021). Metatranscriptomics reveals the gene functions and metabolic properties of the major microbial community during Chinese Sichuan Paocai fermentation. *Food Microbiol.* 98, 103573. doi:10.1016/j.fm.2020.103573
- Xiao, M., Peng, Z., Hardie, W. J., Huang, T., Liu, Z., Zhang, Y., et al. (2022). Exploring the typical flavours formation by combined with metatranscriptomics and metabolomics during Chinese Sichuan paocai fermentation. *LWT* 153, 112474. doi:10.1016/j.lwt.2021.112474
- Xiong, S., Xu, X., Zhang, L., Du, T., Huang, T., Huang, J., et al. (2024). Integrated metatranscriptomics and metabolomics reveal microbial succession and flavor formation mechanisms during the spontaneous fermentation of Laotan Suancai. *Food Res. Int.* 177, 113865. doi:10.1016/j.foodres.2023.113865
- Yang, C., Zhao, F., Hou, Q., Wang, J., Li, M., and Sun, Z. (2020). PacBio sequencing reveals bacterial community diversity in cheeses collected from different regions. *J. Dairy Sci.* 103, 1238–1249. doi:10.3168/jds.2019-17496
- Yang, Z., Kulka, M., Yang, Q., Papafragkou, E., Yu, C., Wales, S. Q., et al. (2024). Whole-Genome sequencing-based confirmatory methods on RT-qPCR results for the detection of foodborne viruses in frozen berries. *Food Environ. Virol.* 16, 225–240. doi:10.1007/s12560-024-09591-6
- Yang, C., Borjihan, Q., Liu, H., Kwok, L.-Y., Wang, D., and Chen, Y. (2025). Terroir-driven flavor, microbiome, and metabolome signatures of jiaoke: a multi-omics exploration of traditional fermented dairy from Inner Mongolia. *Food Biosci.* 66, 105922. doi:10.1016/j.fbio.2025.105922
- Yasir, M., Al-Zahrani, I. A., Bibi, F., Abd El Ghany, M., and Azhar, E. I. (2022). New insights of bacterial communities in fermented vegetables from shotgun metagenomics and identification of antibiotic resistance genes and probiotic bacteria. *Food Res. Int.* 157, 111190. doi:10.1016/j.foodres.2022.111190
- Zhang, T., Zhou, J., Gao, W., Jia, Y., Wei, Y., and Wang, G. (2022). Complex genome assembly based on long-read sequencing. *Briefings Bioinforma.* 23, bbac305. doi:10.1093/bib/bbac305
- Zhao, X., Song, L., Han, D., Han, P., and Bai, F. (2024a). Microbiological and physicochemical dynamics in traditional and industrial fermentation processes of koumiss. *Fermentation* 10, 66. doi:10.3390/fermentation10010066
- Zhao, Z., Zhang, H., Zhao, S., Qu, X., Yu, J., Tang, Y., et al. (2024b). Species identification using PCR targeting specific region in cytochrome oxidase I gene for high-value fish product market in China. *J. Food Compos. Anal.* 136, 106762. doi:10.1016/j.jfca.2024.106762
- Zhou, Y., Liu, M., and Yang, J. (2022). Recovering metagenome-assembled genomes from shotgun metagenomic sequencing data: methods, applications, challenges, and opportunities. *Microbiol. Res.* 260, 127023. doi:10.1016/j.micres.2022.127023
- Zhou, Y., Liu, Y., and Li, X. (2024). USEARCH 12: open-source software for sequencing analysis in bioinformatics and microbiome. *iMeta* 3, e236. doi:10.1002/imt2.236