

# A Comprehensive View of Food Microbiota: Introducing FoodMicrobionet v5

Eugenio Parente \*  and Annamaria Ricciardi

Scuola di Scienze Agrarie, Forestali, Alimentari ed Ambientali, Università degli Studi della Basilicata, 85100 Potenza, Italy; annamaria.ricciardi@unibas.it

\* Correspondence: eugenio.parente@unibas.it

**Abstract:** Amplicon-targeted metagenomics is now the standard approach for the study of the composition and dynamics of food microbial communities. Hundreds of papers on this subject have been published in scientific journals and the information is dispersed in a variety of sources, while raw sequences and their metadata are available in public repositories for some, but not all, of the published studies. A limited number of web resources and databases allow scientists to access this wealth of information but their level of annotation on studies and samples varies. Here, we report on the release of FoodMicrobionet v5, a comprehensive database of metataxonomic studies on bacterial and fungal communities of foods. The current version of the database includes 251 published studies (11 focusing on fungal microbiota, 230 on bacterial microbiota, and 10 providing data for both bacterial and fungal microbiota) and 14,035 samples with data on bacteria and 1114 samples with data on fungi. The new structure of the database is compatible with interactive apps and scripts developed for previous versions and allows scientists, R&D personnel in industries and regulators to access a wealth of information on food microbial communities.

**Keywords:** amplicon-targeted metagenomics; microbiota; food; food environments; database



**Citation:** Parente, E.; Ricciardi, A. A Comprehensive View of Food Microbiota: Introducing FoodMicrobionet v5. *Foods* **2024**, *13*, 1689. <https://doi.org/10.3390/foods13111689>

Academic Editor: Wei Zhang

Received: 26 April 2024

Revised: 21 May 2024

Accepted: 24 May 2024

Published: 28 May 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

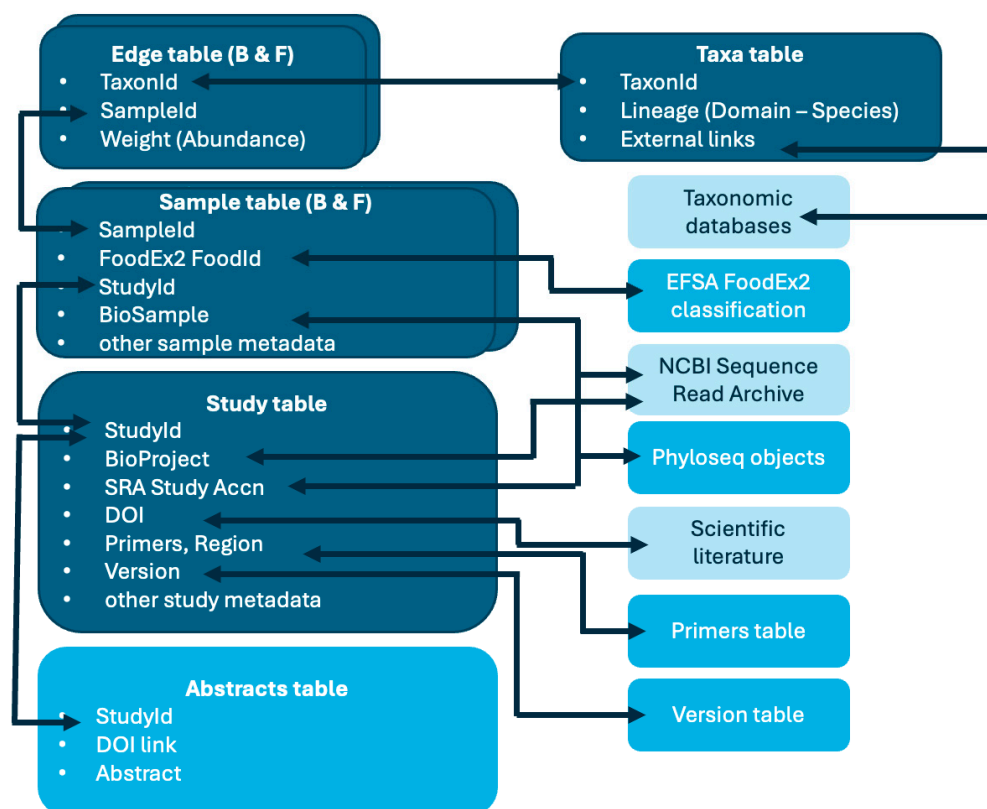
## 1. Introduction

Since the first study published in 2009 [1], the use of Next Generation Sequencing to analyze libraries of amplicons of taxonomically relevant regions of the genomes of *Bacteria* and *Archaea* (typically the 16S RNA gene [2–4]) and *Fungi* (typically one or more regions including the Internal Transcribed Spacer [5–7]) has become the standard tool for the cultivation independent analysis of microbial food communities [8–11]. The approaches used for sequencing, for the bioinformatic analysis of sequences and for the statistical analysis of the results have been reviewed multiple times [2,12–15] and the increasing use of third generation platforms is helping in overcoming some of the limitations of short-read sequencing [16–21]. Raw sequences are often but not always made available in public repositories, like, among others, the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA, <https://www.ncbi.nlm.nih.gov/sra>, last accessed on 18 April 2024), the European Nucleotide Archive (ENA, <https://www.ebi.ac.uk/ena/browser/home>, last accessed on 18 April 2024). In addition, several web-based database and tools [22] provide access to sequences and to processing pipelines. Notable examples are QIITA (<https://qiita.ucsd.edu/>, last accessed on 18 April 2024) [23], MGnify (<https://www.ebi.ac.uk/metagenomics>, last accessed on 18 April 2024) [24], IMNGS (<https://www.imngs.org>, last accessed on 18 April 2024) [25], DBbact (<http://dbbact.org/main>, last accessed on 18 April 2024) [26], and the Microbe Atlas (<https://microbeatlas.org/index.html>, last accessed on 18 April 2024) [27]. However, only recently have Minimum Information about MARKer gene Sequence (MIMARKS [28]) standards become compulsory for sample metadata. Among other things, these standards include annotation with a standardized terminology. For food and food environments samples, NCBI SRA incorporates the FoodOn ontology

for foods and food processing [29]. Similar controlled ontologies have been developed by the European Food Safety Agency [30]. Unfortunately, older accessions have a variable quality of sample metadata and, even for newer deposits in sequence archives, the quality of metadata is proportional to the good will of researchers. To overcome some of the limitations, we actively develop FoodMicrobionet (FMBN), which up to version 4 only included data on bacterial communities [31–33]. Even if the level of sophistication of the database and related software is inferior to other tools, which perform automated searches and analysis on existing repositories or on user-provided projects (like QIITA, MGnify, IMNGS and Microbe Atlas), we feel that the level of annotation of the FoodMicrobionet allows scientists to rapidly access complete information on studies, samples and taxa, integrate them with external resources and rapidly retrieve data for further analysis in a variety of projects. We have demonstrated use cases in a number of recent papers [32,34–37]. Here, we report on the release of version 5 of FoodMicrobionet, which, for the first time, includes data on fungal microbial communities of food and food environments.

## 2. Materials and Methods

The structure of FoodMicrobionet v 5.0 has changed little compared to its previous iterations [32,33] and it is schematically shown in Figure 1.



**Figure 1.** FoodMicrobionet table structure. The main tables are in dark blue; the service tables are in cyan; external resources accessible via links in FoodMicrobionet tables are in light sky blue. Arrows connect primary and foreign key fields in each table.

The main changes include:

1. Addition of data on fungi, with new samples and edges tables;
2. Addition of two further accessory tables, abstracts and version history.

The relational database now includes 4 main tables:

1. Studies: study metadata, including NCBI SRA (<https://www.ncbi.nlm.nih.gov/sra>, last accessed on 18 April 2024) accession numbers for bioproject and study, references,

- nucleic acid targets, etc.; each study has a unique id and corresponds to a single NCBI project accession;
2. Samples: sample description, FoodEx2 classification [30], NCBI SRA accession numbers for biosample and run (which are unique for each sample), etc.;
  3. Taxa: label and taxonomic lineage for taxa and links to external resources;
  4. Edges: relative abundance for each taxon found in a given sample
- and a number of service tables:
1. Primers: with sequences of primers used for amplification and target regions;
  2. Abstracts: a text table with abstracts for each of the publications from which sequence data were obtained;
  3. Version history: change log and references for each FoodMicrobionet version;
  4. FoodEx2 classification.

Detailed information on tables and fields is available in Supplementary Material (Supplementary File S1).

To accommodate data on fungal communities of foods and food environments, both the samples and the edges tables were duplicated (i.e., samples\_B and samples\_F; edges\_B and edges\_F). This approach was needed for two main reasons:

1. In some cases, data available in NCBI were deposited without fully adhering to SRA specifications (<https://www.ncbi.nlm.nih.gov/sra/docs/submitmeta/>, accessed on 18 April 2024): i.e., data for the same project were deposited within different bioproject accessions; data for the same biological sample were deposited with different biosample accessions;
2. Metataxonomic taxa abundance data are, in their essence, compositional: no absolute quantification of target for bacteria and fungi is typically available and runs targeting, for example, the V3-V4 region of the 16S rRNA gene for bacteria and Internal Transcribed Spacer 1 region (ITS1) for fungi are separate. In FoodMicrobionet, for both types of runs, the total number of sequences (which differ in the two runs) and the proportion for each taxon as relative abundance (%) is stored in sample and edges tables, respectively.

The taxa table now includes both bacteria and fungi.

As for previous versions, FoodMicrobionet grows using the addition of raw sequences downloaded from SRA. We regularly scan the literature for new publications and those whose sequence data have been deposited in SRA or ENA are submitted for further screening to evaluate if metadata available in the sequence archives or in the scientific papers enable the obtainment of the information needed to complete the fields in the study and sample table. The raw sequences are then processed with an analysis pipeline adapted from the Bioconductor workflow [38], based on the DADA2 package [39], with several modifications to accommodate for variable length sequences in ITS targeted studies and for sequence data in fastq format with compressed quality scores obtained with the most recent Illumina platforms (including those of the NovaSeq series; <https://tinyurl.com/4c3nekt4>, last accessed on 23 May 2024). The current version of the pipeline includes options for primer removal with cutadapt [40]. In this version of the database, only studies using the Internal Transcribed Spacer region as a target are used. Taxonomic assignment for fungi is performed using the UNITE [41,42] general release for fungi using the assignTaxonomy() function of the DADA2 package for R [39].

After processing, metadata were manually annotated through merging metadata from NCBI SRA with data obtained from the scientific publications.

A further feature of this version is the addition of a matchId field in sample tables for both bacteria and fungi. This procedure allows rapid identification of matching samples in both tables even when deposit in SRA by the original authors violated the principle that a single biological sample must have a unique accession.

As in previous versions, all scripts for sequence processing and assembly of the database are publicly available in GitHub (<https://github.com/ep142/FoodMicrobionet>, last accessed on 23 May 2024).

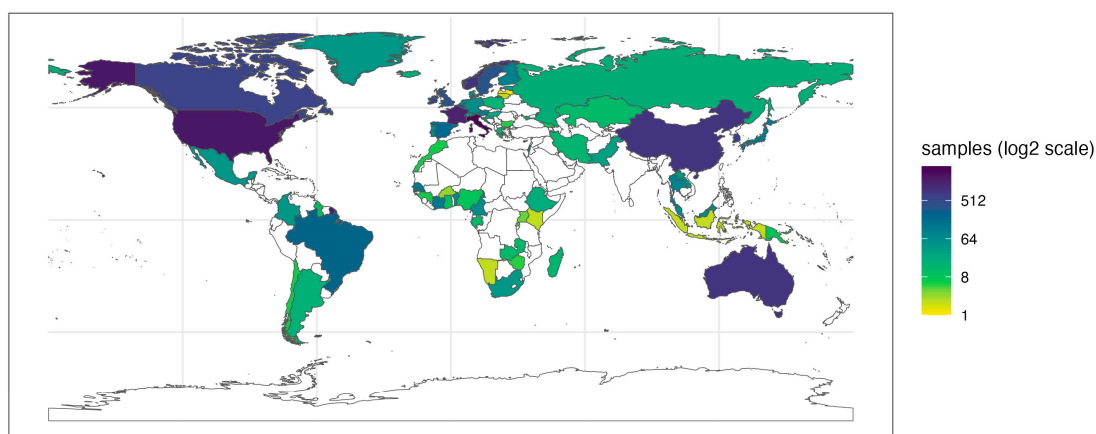
The database can be easily accessed using an interactive R Shiny app which is publicly available (Parente, Eugenio (2024), “ShinyFMBN, a Shiny app to access FoodMicrobionet”, Mendeley Data, V9, doi: 10.17632/8fwwjpm79y.9), but both the bacterial and fungal versions of the database are compatible, with minor changes, with scripts developed and used in previous proof of concept papers [34,36,43].

### 3. Results and Discussion

#### 3.1. FoodMicrobionet, a Global Database

Version 5 of FoodMicrobionet includes data from 251 published studies, 11 of which are on fungal microbiota only, while 230 are on bacterial microbiota. There are 10 studies for which both the data for bacterial and fungal microbiota are available. However, due to inconsistencies in the deposit of sequences in SRA (in several cases the same sample was deposited with two separate biosample accessions and/or data for bacteria and fungi were deposited with different bioproject or study accessions), the same samples were present in two studies, one for bacteria and one for fungi. We did our best to manually match samples in these situations. The addition of datasets on fungi is in progress and, in time, we will add fungal data for all the studies which are already in FoodMicrobionet with bacterial community data and add more studies specialized on fungal microbiota. The combinations of sequencing platforms and targets (Supplementary Tables, Table S1) reflect current and past practices in microbiota analysis. For bacteria, studies performed using the V3-V4 or V4 regions of the 16S RNA gene and with the Illumina platform are the majority. We are currently only adding studies using the ITS target for fungi. Even with its known limitations [20,44], this target is still the most frequently used by the scientific community. On the other hand, FoodMicrobionet is fully transparent in terms of information on targets (which can be selected or excluded in both the studies and sample table) and primers (available in the studies table). This feature is not immediately available in other databases or online tools, like QIITA [23], MGnify [24] or IMNGS [25].

The latest versions of FoodMicrobionet include information on the geographic location of samples in SRA metadata or in the published papers whenever available. The geographical distribution of the samples is shown in Figure 2.



**Figure 2.** Geographical distribution of samples in FoodMicrobionet 5.0. The continuous color scale for countries represents the number of food and food environmental samples present in FoodMicrobionet for each country.

FoodMicrobionet is a truly global database with samples from 77 countries and 6 continents. The level of detail in the geographical annotation varies, but several samples include geographical coordinate data. This, in principle, would allow interested parties to perform

biogeographical studies, comparing the microbiota of a specific product over several locations [45,46]. The level of detail of sample descriptions for both food samples and food environments or food contact surfaces allows the mapping of samples within factories to perform studies of microbial landscapes within a factory or similar factories [47,48].

### 3.2. Diversity of Samples

One of the main strengths of FoodMicrobionet is the level of annotation of sample metadata, which includes a description obtained through incorporating a combination of metadata fields from sample metadata deposited in NCBI SRA, the FoodEx2 [30] code, and the L1, L4 and L6 descriptions from the same classification. Aggregation at other levels of FoodEx2 can be easily performed using the service table FoodEx2. In addition, further fields indicate the type of sample (food or food environment), the stage of the process (raw, intermediate, finished), the use of a lethal treatment, and the occurrence of spoilage and fermentation.

The diversity of samples in FoodMicrobionet is extensive. FoodMicrobionet version 5 includes 14,035 samples with data on bacteria and 1114 samples with data on fungi. For 366 samples, we were able to match samples for bacteria and fungi. There are 2634 environmental samples and 12,125 food samples. Samples in FMBN belong to 19 major food groups (L1 level of FoodEx2 exposure classification, Table 1). Samples in FMBN are further classified using levels L4 and L6 of the FoodEx2 exposure classification, and additional fields (which allow to identify raw products, intermediates or finished products, the level of thermal treatment and the occurrence of spoilage and/or fermentation) allow a finer classification. Samples in FMBN belong to 134 L4 food groups and 239 L6 food groups.

**Table 1.** Distribution (count, proportion and cumulative proportion) of FoodMicrobionet 5 samples in major categories (L1 level) of the FoodEx2 classification.

L1	n	Prop	Cum.Prop.
Milk and dairy products	5807	0.3931	0.3931
Meat and meat products	3605	0.2440	0.6371
Vegetables and vegetable products	1508	0.1021	0.7391
Fruit and fruit products	1178	0.0797	0.8189
Fish, seafood, amphibians, reptiles and invertebrates	814	0.0551	0.8740
Alcoholic beverages	513	0.0347	0.9087
Major isolated ingredients, additives, flavours, baking and processing aids	339	0.0229	0.9316
Seasoning, sauces and condiments	202	0.0137	0.9453
Grains and grain-based products	155	0.0105	0.9558
Composite dishes	147	0.0099	0.9658
Fruit and vegetable juices and nectars (including concentrates)	82	0.0056	0.9713
Legumes, nuts, oilseeds and spices	81	0.0055	0.9768
Generic food environments	64	0.0043	0.9811
Food products for young population	61	0.0041	0.9852
Eggs and egg products	56	0.0038	0.9890
Animal and vegetable fats and oils and primary derivatives thereof	53	0.0036	0.9926
Sugar and similar, confectionery and water-based sweet desserts	42	0.0028	0.9955
Starchy roots or tubers and products thereof, sugar plants	17	0.0012	0.9983
Coffee, cocoa, tea and infusions	15	0.0010	0.9993

There are 199 foodIds (food types), and, combining further information on samples (nature, heat treatment, spoilage/fermentation), there are 388 combinations. The diversity is particularly extensive for dairy products, including cheeses.

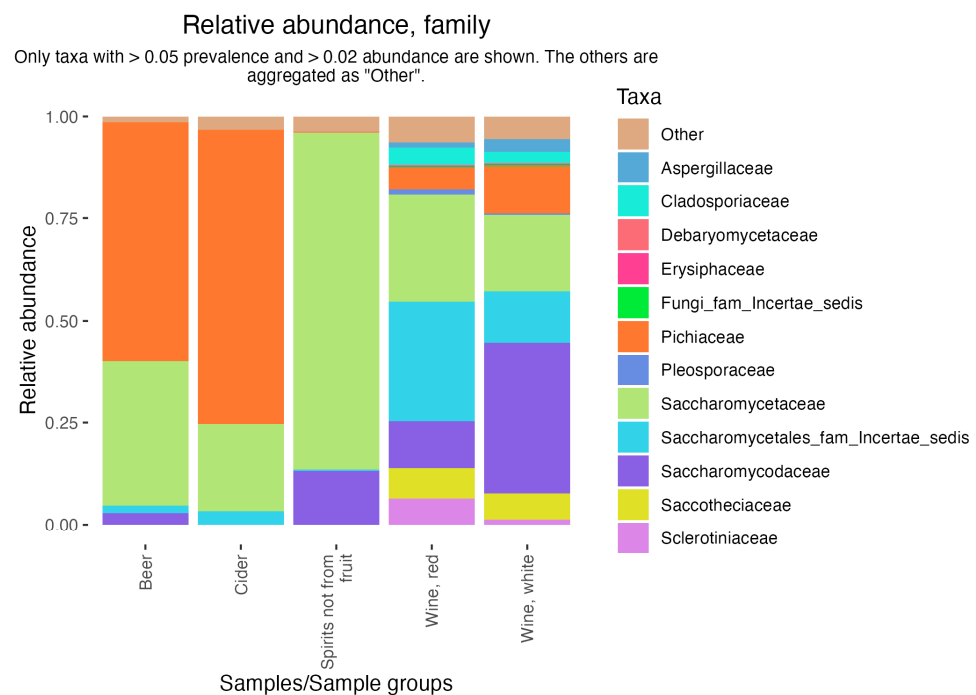
### 3.3. Accessibility of FoodMicrobionet Data

Researchers with the ability to use the R programming language and environment can make the most of FoodMicrobionet and use it extensively for metastudies to test specific hypotheses or even to obtain quantitative comparisons of their own data with

those available in FoodMicrobionet. While we have demonstrated this in previous proof of concept papers [32,34,36,37,49], we will add two common use cases below.

### 3.3.1. FMBN Made Simple: Accessing Data Using the ShinyFMBN App

The ShinyFMBN app (<https://data.mendeley.com/datasets/8fwwjpm79y/9>, last accessed on 23 May 2024) is designed to provide offline access to the database to users who have limited experience with R. The use requires minimal informatic abilities: it is sufficient to install R and launch a script which installs all needed packages and launches the app. By changing a few lines of code in the app, the user can choose to access samples with data on fungi or bacteria. The app can be opened in a browser and allows the exploration of data, the selection of subsets of samples, the export of the subset in several formats (including a phyloseq object [50], a versatile format which is used by many other R packages for the analysis of microbiome data, including the interactive ShinyPhyloseq app [51]). However, ShinyFMBN has also limited capabilities to produce tabular and graphical analysis (prevalence and abundance plots and tables, bar plots, box plots). Figure 3 shows an example: data from six studies on alcoholic beverages [52–57] were used to show the distribution of fungal families in beer, wine, cider, pulque, and palm wine. Differences in the relative abundance of yeasts are clearly evident, with beer dominated by *Dekkera* and *Saccharomyces*, cider by *Pichia*, *Saccharomyces*, *Dekkera*, and *Brettanomyces*, palm wine by *Saccharomyces* and *Hanseniaspora*, and wines with a wider diversity of genera.

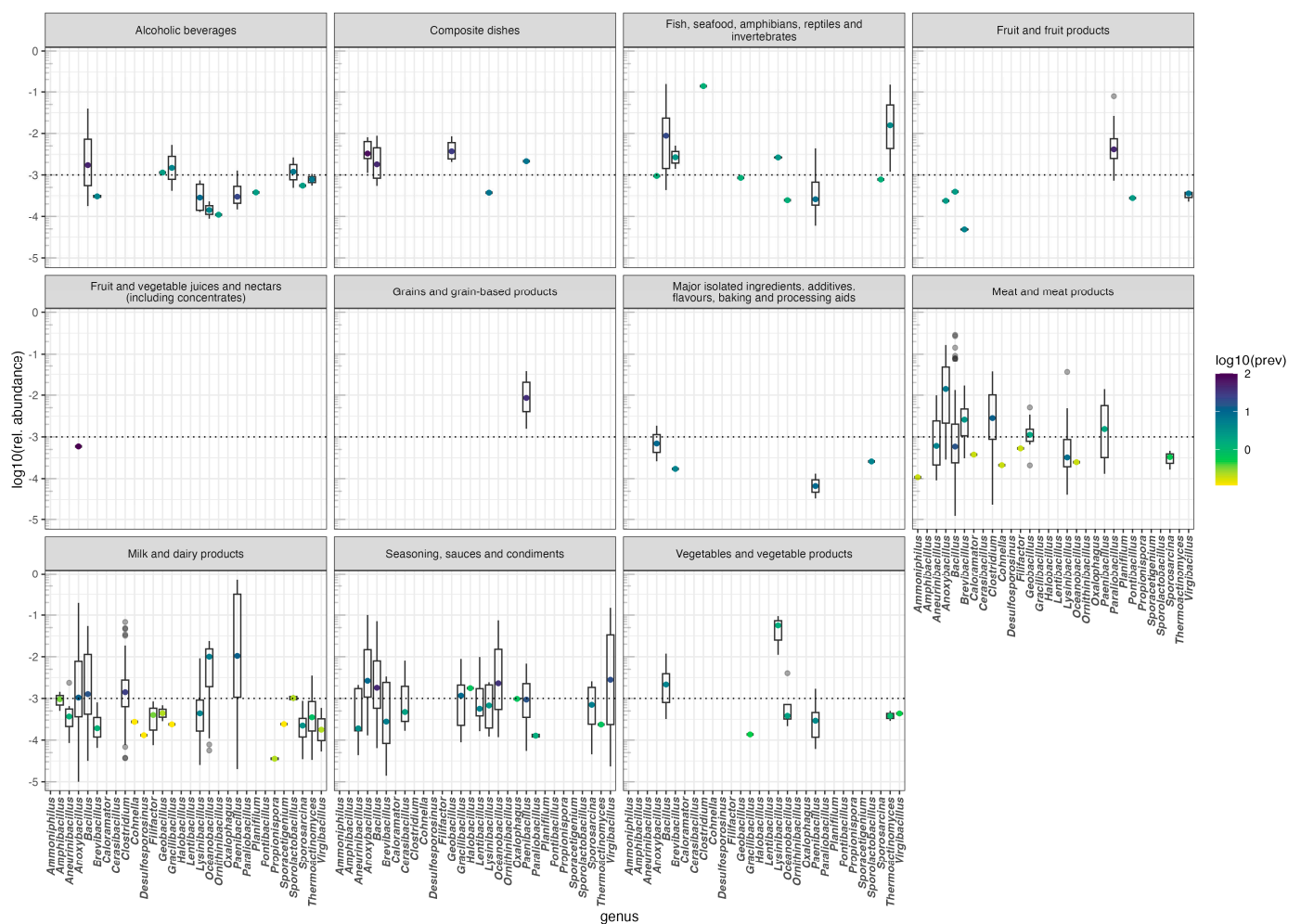


**Figure 3.** Relative abundance of fungal families in 5 alcoholic beverages. Data extracted from FoodMicrobionet 5.0.

### 3.3.2. Using FoodMicrobionet to Explore the Distribution of Microbial Groups in Foods: The Case of Endospore-Forming Bacillota

We have recently used FoodMicrobionet to explore the distribution of *Lactobacillaceae* in foods [34]. The same approach can be easily used to explore the distribution of other microbial groups which play important roles in the safety and spoilage of foods. Bacteria producing endospores are important in some food fermentations [58] but also contribute to the spoilage of several foods (see, for example [59–64]) and some pathogenic species (*Clostridium botulinum*, *C. perfringens*, *Bacillus cereus*) are causative agents of foodborne outbreaks [65,66]. Using the approach used in [34], we extracted data on the relative abundance of endospore-forming Bacillota and used them to build Figure 4. In this figure,

the panes (facets) show the distribution of the relative abundance of the main genera in different categories of the L1 level of the FoodEx2 classification. Visual cues on prevalence are also included: The median relative abundance in a given food group is marked by a dot whose color reflects the prevalence. The food groups with the highest diversity of endospore-forming bacteria are milk and dairy products (19 genera), meat and meat products (14) and seasoning and spices (16). While most genera appear in foods with low relative abundance and prevalence, and are possibly contaminants from raw materials or food contact environments and surfaces, others, like *Anoxybacillus*, *Bacillus*, and *Paenibacillus* have a much wider distribution and sometimes their relative abundance exceeds 0.10. Other genera, like *Clostridium* and *Virgibacillus*, are found at relatively high relative abundances and are prevalent only in some foods. The comprehensive annotation of samples in FoodMicrobionet allows the further exploration of these types of data, down to the lowest level of the FoodEx2 classification, providing scientists with a powerful tool to formulate or disprove scientific hypotheses.



**Figure 4.** Relative abundance and prevalence of genera of the phylum *Bacillota* which include endospore-forming bacteria in major food groups. Data extracted from FoodMicrobionet 5.0. Median abundance is shown as a dot whose color scale represents the base 10 logarithm of the prevalence (as %) of the genus in a given food group.

#### 4. Conclusions

This iteration of FoodMicrobionet provides for the first time both data on bacterial and fungal communities, thus allowing the obtainment of a comprehensive picture of the microbial ecology of foods. With very few changes, the database is retro-compatible with

most of the analysis scripts we have developed in the last few years, providing a vast array of statistical and graphic tools for the analysis and representation of the structure of microbial communities. We plan to continue expanding the database by adding data for fungi for existing studies and adding more studies. In addition, we plan to make the ShinyFMBN app available online by using shinyapps.io to eliminate the need for the installation of R.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/foods13111689/s1>, Supplementary File S1: FoodMicrobionet table specifications; Table S1: Distribution of target regions and sequencing platforms for studies in FoodMicrobionet 5.0.

**Author Contributions:** Conceptualization, E.P. and A.R.; methodology, E.P.; software, E.P.; validation, E.P. and A.R.; data curation, E.P. and A.R.; writing—original draft preparation, E.P. and A.R.; writing—review and editing, E.P.; supervision, E.P.; project administration, E.P.; funding acquisition, E.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was carried out within the PRIN 2022 PNRR Project NYC diversity P20229JMMH and received funding from the European Union Next-GenerationEU, CUP C53D23007560001 (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR)—MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4—D.D. 1364/2023). This article and its contents reflect only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

**Data Availability Statement:** The database and related scripts are available on GitHub (<https://github.com/ep142/FoodMicrobionet>, last accessed on 23 May 2024). Additional data (phyloseq objects with ASVs, further metadata) are available on Zenodo (<https://zenodo.org/doi/10.5281/zenodo.6954039>, last accessed on 23 May 2024). The Shiny app is available on Mendeley data (<https://data.mendeley.com/datasets/8fwwjpm79y/9>, last accessed on 23 May 2024).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Humblot, C.; Guyot, J.P. Pyrosequencing of tagged 16S rRNA gene amplicons for rapid deciphering of the microbiomes of fermented foods such as pearl millet slurries. *Appl. Environ. Microbiol.* **2009**, *75*, 4354–4361. [[CrossRef](#)] [[PubMed](#)]
2. Pollock, J.; Glendinning, L.; Wisedchanwet, T.; Watson, M. The madness of microbiome: Attempting to find consensus “best practice” for 16S microbiome studies. *Appl. Environ. Microbiol.* **2018**, *84*, 3225. [[CrossRef](#)] [[PubMed](#)]
3. Vos, M.; Quince, C.; Pijl, A.S.; de Hollander, M.; Kowalchuk, G.A. A comparison of *rpoB* and 16S rRNA as markers in pyrosequencing studies of bacterial diversity. *PLoS ONE* **2012**, *7*, e30600. [[CrossRef](#)] [[PubMed](#)]
4. Fouhy, F.; Clooney, A.G.; Stanton, C.; Claesson, M.J.; Cotter, P.D. 16S rRNA gene sequencing of mock microbial populations—impact of DNA extraction method, primer choice and sequencing platform. *BMC Microbiol.* **2016**, *16*, 123. [[CrossRef](#)] [[PubMed](#)]
5. Yang, R.-H.; Su, J.-H.; Shang, J.-J.; Wu, Y.-Y.; Li, Y.; Bao, D.-P.; Yao, Y.-J. Evaluation of the ribosomal DNA internal transcribed spacer (ITS), specifically ITS1 and ITS2, for the analysis of fungal diversity by deep sequencing. *PLoS ONE* **2018**, *13*, e0206428. [[CrossRef](#)] [[PubMed](#)]
6. Karkouri, K.E.; Murat, C.; Zampieri, E.; Bonfante, P. Identification of internal transcribed spacer sequence motifs in truffles: A first step toward their DNA bar coding. *Appl. Environ. Microbiol.* **2007**, *73*, 5320–5330. [[CrossRef](#)] [[PubMed](#)]
7. Bokulich, N.A.; Mills, D.A. Improved selection of internal transcribed spacer-specific primers enables quantitative, ultra-high-throughput profiling of fungal communities. *Appl. Environ. Microbiol.* **2013**, *79*, 2519–2526. [[CrossRef](#)]
8. Filippis, F.D.; Parente, E.; Ercolini, D. Metagenomics insights into food fermentations. *Microb. Biotechnol.* **2017**, *10*, 91–102. [[CrossRef](#)] [[PubMed](#)]
9. Filippis, F.D.; Parente, E.; Ercolini, D. Recent past, present, and future of the food microbiome. *Annu. Rev. Food Sci. T* **2018**, *9*, 589–608. [[CrossRef](#)]
10. Ferracino, I.; Rantsiou, K.; Cocolin, L. Microbiome and -omics application in food industry. *Int. J. Food Microbiol.* **2022**, *377*, 109781. [[CrossRef](#)]
11. Yeluri Jonnala, B.R.; McSweeney, P.L.H.; Sheehan, J.J.; Cotter, P.D. Sequencing of the cheese microbiome and its relevance to industry. *Front. Microbiol.* **2018**, *9*, 1020. [[CrossRef](#)] [[PubMed](#)]
12. Bokulich, N.A.; Ziemski, M.; Robeson, M.S.; Kaehler, B.D. Measuring the microbiome: Best practices for developing and benchmarking microbiomics methods. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 4048–4062. [[CrossRef](#)] [[PubMed](#)]



13. Hakimzadeh, A.; Asbun, A.A.; Albanese, D.; Bernard, M.; Buchner, D.; Callahan, B.; Caporaso, J.G.; Curd, E.; Djemiel, C.; Durling, M.B.; et al. A pile of pipelines: An overview of the bioinformatics software for metabarcoding data analyses. *Mol. Ecol. Resour.* **2023**, *early review*. [[CrossRef](#)]
14. Walsh, L.H.; Coakley, M.; Walsh, A.M.; O'Toole, P.W.; Cotter, P.D. Bioinformatic approaches for studying the microbiome of fermented food. *Crit. Rev. Microbiol.* **2022**, *49*, 693–725. [[CrossRef](#)] [[PubMed](#)]
15. Wen, T.; Niu, G.; Chen, T.; Shen, Q.; Yuan, J.; Liu, Y.-X. The best practice for microbiome analysis using R. *Protein Cell* **2023**, *14*, 713–725. [[CrossRef](#)] [[PubMed](#)]
16. Tedersoo, L.; Albertsen, M.; Anslan, S.; Callahan, B. Perspectives and benefits of high-throughput long-read sequencing in microbial ecology. *Appl. Environ. Microbiol.* **2021**, *87*, e00626–21. [[CrossRef](#)] [[PubMed](#)]
17. Matsuo, Y.; Komiya, S.; Yasumizu, Y.; Yasuoka, Y.; Mizushima, K.; Takagi, T.; Kryukov, K.; Fukuda, A.; Morimoto, Y.; Naito, Y.; et al. Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MINION™ Nanopore sequencing confers species-level resolution. *BMC Microbiol.* **2021**, *21*, 35. [[CrossRef](#)] [[PubMed](#)]
18. Rubiola, S.; Macori, G.; Civera, T.; Fanning, S.; Mitchell, M.; Chiesa, F. Comparison between full-length 16S rRNA metabarcoding and whole metagenome sequencing suggests the use of either is suitable for large-scale microbiome studies. *Foodborne Pathog. Dis.* **2022**, *19*, 495–504. [[CrossRef](#)] [[PubMed](#)]
19. Baloğlu, B.; Chen, Z.; Elbrecht, V.; Braukmann, T.; MacDonald, S.; Steinke, D. A Workflow for accurate metabarcoding using Nanopore MinION sequencing. *Methods Ecol. Evol.* **2021**, *12*, 794–804. [[CrossRef](#)]
20. Rué, O.; Coton, M.; Dugat-Bony, E.; Howell, K.; Irlinger, F.; Legras, J.-L.; Loux, V.; Michel, E.; Mounier, J.; Neuvéglise, C.; et al. Comparison of metabarcoding taxonomic markers to describe fungal communities in fermented foods. *Peer Community J.* **2023**, *3*, e97. [[CrossRef](#)]
21. Callahan, B.J.; Wong, J.; Heiner, C.; Oh, S.; Theriot, C.M.; Gulati, A.S.; McGill, S.K.; Dougherty, M.K. High-Throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Res.* **2019**, *47*, e103. [[CrossRef](#)]
22. Lobanov, V.; Gobet, A.; Joyce, A. Ecosystem-specific microbiota and microbiome databases in the era of big data. *Environ. Microbiome* **2022**, *17*, 37. [[CrossRef](#)] [[PubMed](#)]
23. Gonzalez, A.; Navas-Molina, J.A.; Kosciolk, T.; McDonald, D.; Vázquez-Baeza, Y.; Ackermann, G.; DeReus, J.; Janssen, S.; Swafford, A.D.; Orchanian, S.B.; et al. Qiita: Rapid, web-enabled microbiome meta-analysis. *Nat. Methods* **2018**, *15*, 796–798. [[CrossRef](#)] [[PubMed](#)]
24. Mitchell, A.L.; Almeida, A.; Beracochea, M.; Boland, M.; Burgin, J.; Cochrane, G.; Crusoe, M.R.; Kale, V.; Potter, S.C.; Richardson, L.J.; et al. MGnify: The microbiome analysis resource in 2020. *Nucleic Acids Res.* **2019**, *48*, D570–D578. [[CrossRef](#)] [[PubMed](#)]
25. Lagkouvardos, I.; Joseph, D.; Kapfhammer, M.; Giritli, S.; Horn, M.; Haller, D.; Clavel, T. IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. *Sci. Rep.* **2016**, *6*, 33721. [[CrossRef](#)] [[PubMed](#)]
26. Amir, A.; Ozel, E.; Haberman, Y.; Shental, N. Achieving pan-microbiome biological insights via the DbBact knowledge base. *Nucleic Acids Res.* **2023**, *51*, 6593–6608. [[CrossRef](#)] [[PubMed](#)]
27. Rodrigues, J.F.M.; Schmidt, T.S.B.; Tackmann, J.; Mering, C. von MAPseq: Highly Efficient k-mer search with confidence estimates, for rRNA sequence analysis. *Bioinformatics* **2017**, *33*, 3808–3810. [[CrossRef](#)] [[PubMed](#)]
28. Yilmaz, P.; Kottmann, R.; Field, D.; Knight, R.; Cole, J.R.; Amaral-Zettler, L.; Gilbert, J.A.; Karsch-Mizrachi, I.; Johnston, A.; Cochrane, G.; et al. Minimum information about a marker gene sequence (MIMARKS) and Minimum Information about Any (x) Sequence (MIxS) specifications. *Nat. Biotechnol.* **2011**, *29*, 415–420. [[CrossRef](#)] [[PubMed](#)]
29. Dooley, D.M.; Griffiths, E.J.; Gosal, G.S.; Buttigieg, P.L.; Hoehndorf, R.; Lange, M.C.; Schriml, L.M.; Brinkman, F.S.L.; Hsiao, W.W.L. FoodOn: A harmonized food ontology to increase global food traceability, quality control and data integration. *NPJ Sci. Food* **2018**, *2*, 23. [[CrossRef](#)] [[PubMed](#)]
30. EFSA. *The Food Classification and Description System FoodEx 2 (Revision 2)*; EFSA Supporting Publications: Parma, Italy, 2015; pp. 1–90. [[CrossRef](#)]
31. Parente, E.; Coccolin, L.; De Filippis, F.; Zotta, T.; Ferrrocino, I.; O'Sullivan, O.; Neviani, E.; Angelis, M.D.; Cotter, P.D.; Ercolini, D. FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis. *Int. J. Food Microbiol.* **2016**, *219*, 28–37. [[CrossRef](#)]
32. Parente, E.; Zotta, T.; Ricciardi, A. FoodMicrobionet v4: A large, integrated, open and transparent database for food bacterial communities. *Int. J. Food Microbiol.* **2022**, *372*, 109696. [[CrossRef](#)]
33. Parente, E.; De Filippis, F.; Ercolini, D.; Ricciardi, A.; Zotta, T. Advancing integration of data on food microbiome studies: Foodmicrobionet 3.1, a major upgrade of the FoodMicrobionet database. *Int. J. Food Microbiol.* **2019**, *305*, 108249. [[CrossRef](#)] [[PubMed](#)]
34. Parente, E.; Zotta, T.; Giavalisco, M.; Ricciardi, A. Metataxonomic insights in the distribution of *Lactobacillaceae* in foods and food environments. *Int. J. Food Microbiol.* **2023**, *391*, 110124. [[CrossRef](#)] [[PubMed](#)]
35. Parente, E.; Zotta, T.; Ricciardi, A. Microbial association networks in cheese: A meta-analysis. *bioRxiv* **2021**. [[CrossRef](#)]
36. Parente, E.; Ricciardi, A.; Zotta, T. The microbiota of dairy milk: A review. *Int. Dairy J.* **2020**, *107*, 104714. [[CrossRef](#)]
37. Zotta, T.; Ricciardi, A.; Condelli, N.; Parente, E. Metataxonomic and metagenomic approaches for the study of undefined strain starters for cheese manufacture. *Crit. Rev. Food Sci.* **2022**, *62*, 3898–3912. [[CrossRef](#)] [[PubMed](#)]
38. Callahan, B.J.; Sankaran, K.; Fukuyama, J.A.; McMurdie, P.J.; Holmes, S.P. Bioconductor workflow for microbiome data analysis: From raw reads to community analyses. *F1000Research* **2016**, *5*, 1492. [[CrossRef](#)] [[PubMed](#)]

39. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [[CrossRef](#)] [[PubMed](#)]
40. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2011**, *17*, 10–12. [[CrossRef](#)]
41. Glöckner, F.O.; Yilmaz, P.; Quast, C.; Gerken, J.; Beccati, A.; Ciuprina, A.; Bruns, G.; Yarza, P.; Peplies, J.; Westram, R.; et al. 25 Years of serving the community with ribosomal RNA gene reference databases and tools. *J. Biotechnol.* **2017**, *261*, 169–176. [[CrossRef](#)]
42. Abarenkov, K.; Nilsson, R.H.; Larsson, K.-H.; Taylor, A.F.S.; May, T.W.; Frøslev, T.G.; Pawlowska, J.; Lindahl, B.; Pöldmaa, K.; Truong, C.; et al. The UNITE database for molecular identification and taxonomic communication of Fungi and other Eukaryotes: Sequences, taxa and classifications reconsidered. *Nucleic Acids Res.* **2023**, *52*, D791–D797. [[CrossRef](#)]
43. Parente, E.; Zotta, T.; Ricciardi, A. A review of methods for the inference and experimental confirmation of microbial association networks in cheese. *Int. J. Food Microbiol.* **2022**, *368*, 109618. [[CrossRef](#)]
44. Kauserud, H. ITS Alchemy: On the use of ITS as a DNA marker in fungal ecology. *Fungal Ecol.* **2023**, *65*, 101274. [[CrossRef](#)]
45. Bokulich, N.A.; Thorngate, J.H.; Richardson, P.M.; Mills, D.A. Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E139–E148. [[CrossRef](#)]
46. Bokulich, N.A.; Amiranashvili, L.; Amiranashvili, L.; Chitchyan, K.; Chitchyan, K.; Ghazanchyan, N.; Ghazanchyan, N.; Darbinyan, K.; Darbinyan, K.; Gagelidze, N.; et al. Microbial biogeography of the transnational fermented milk Matsoni. *Food Microbiol.* **2015**, *50*, 12–19. [[CrossRef](#)] [[PubMed](#)]
47. Bokulich, N.A.; Lewis, Z.T.; Boundy-Mills, K.; Mills, D.A. A new perspective on microbial landscapes within food production. *Curr. Opin. Biotechnol.* **2016**, *37*, 182–189. [[CrossRef](#)]
48. Filippis, F.D.; Valentino, V.; Alvarez-Ordóñez, A.; Cotter, P.D.; Ercolini, D. Environmental microbiome mapping as a strategy to improve quality and safety in the food industry. *Curr. Opin. Food Sci.* **2020**, *38*, 168–176. [[CrossRef](#)]
49. Parente, E.; Zotta, T.; Faust, K.; Filippis, F.D.; Ercolini, D. Structure of association networks in food bacterial communities. *Food Microbiol.* **2018**, *73*, 49–60. [[CrossRef](#)] [[PubMed](#)]
50. McMurdie, P.J.; Holmes, S. Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **2013**, *8*, e61217. [[CrossRef](#)] [[PubMed](#)]
51. McMurdie, P.J.; Holmes, S. Shiny-Phyloseq: Web application for interactive microbiome analysis with provenance tracking. *Bioinformatics* **2015**, *31*, 282–283. [[CrossRef](#)]
52. Tyakht, A.; Kopeliovich, A.; Klimenko, N.; Efimova, D.; Dovidchenko, N.; Odintsova, V.; Kleimenov, M.; Toshchakov, S.; Popova, A.; Khomyakova, M.; et al. Characteristics of bacterial and yeast microbiomes in spontaneous and mixed-fermentation beer and cider. *Food Microbiol.* **2021**, *94*, 103658. [[CrossRef](#)]
53. Rocha-Arriaga, C.; Espinal-Centeno, A.; Martínez-Sánchez, S.; Caballero-Pérez, J.; Alcaraz, L.D.; Cruz-Ramírez, A. Deep Microbial community profiling along the fermentation process of Pulque, a biocultural resource of Mexico. *Microbiol. Res.* **2020**, *241*, 126593. [[CrossRef](#)]
54. Kamilari, E.; Mina, M.; Karallis, C.; Tsaltas, D. Metataxonomic analysis of grape microbiota during wine fermentation reveals the distinction of Cyprus regional terroirs. *Front. Microbiol.* **2021**, *12*, 726483. [[CrossRef](#)] [[PubMed](#)]
55. Kordowska-Wiater, M.; Pytko, M.; Stój, A.; Kubik-Komar, A.; Wyrostek, J.; Waśko, A. A Metagenetic insight into microbial diversity of spontaneously fermented Polish red wines and an analysis of selected physicochemical properties. *Appl. Sci.* **2022**, *12*, 4373. [[CrossRef](#)]
56. Djeni, T.N.; Keisam, S.; Kouame, K.H.; Assouhoun-Djeni, C.N.; Ake, F.D.M.; Amoikon, L.S.T.; Tuikhar, N.; Labala, R.K.; Dje, M.K.; Jeyaram, K. Dynamics of microbial populations and metabolites of fermenting saps throughout tapping process of ron and oil palm trees in Côte d’Ivoire. *Front. Microbiol.* **2022**, *13*, 954917. [[CrossRef](#)] [[PubMed](#)]
57. Martiniuk, J.T.; Hamilton, J.; Dodsworth, T.; Measday, V. Grape-associated fungal community patterns persist from berry to wine on a fine geographical scale. *FEMS Yeast Res.* **2023**, *23*, foac067. [[CrossRef](#)] [[PubMed](#)]
58. Li, Z.; Zheng, M.; Zheng, J.; Gänzle, M.G. *Bacillus* species in food fermentations: An underappreciated group of organisms for safe use in food fermentations. *Curr. Opin. Food Sci.* **2023**, *50*, 101007. [[CrossRef](#)]
59. Kakagianni, M.; Koutsoumanis, K.P. Mapping the risk of evaporated milk spoilage in the Mediterranean region based on the effect of temperature conditions on *Geobacillus stearothermophilus* growth. *Food Res. Int.* **2018**, *111*, 104–110. [[CrossRef](#)] [[PubMed](#)]
60. Burgess, S.A.; Lindsay, D.; Flint, S.H. Thermophilic bacilli and their importance in dairy processing. *Int. J. Food Microbiol.* **2010**, *144*, 215–225. [[CrossRef](#)]
61. Postollec, F.; Mathot, A.-G.; Bernard, M.; Divanac’h, M.-L.; Pavan, S.; Sohier, D. Tracking spore-forming bacteria in food: From natural biodiversity to selection by processes. *Int. J. Food Microbiol.* **2012**, *158*, 1–8. [[CrossRef](#)]
62. Chang, S.-S.; Kang, D.-H. *Alicyclobacillus* spp. in the fruit juice industry: History, characteristics, and current isolation/detection procedures. *Crit. Rev. Microbiol.* **2008**, *30*, 55–74. [[CrossRef](#)]
63. Olajide, A.M.; LaPointe, G. Detection of spore forming *Paenibacillus macerans* in raw milk. *J. Microbiol. Methods* **2020**, *177*, 106048. [[CrossRef](#)] [[PubMed](#)]
64. Langendries, S.; Goormachtig, S. *Paenibacillus polymyxa*, a jack of all trades. *Environ. Microbiol.* **2021**, *23*, 5659–5669. [[CrossRef](#)] [[PubMed](#)]

65. Markland, S.M.; Farkas, D.F.; Kniel, K.E.; Hoover, D.G. Pathogenic psychrotolerant sporeformers: An emerging challenge for low-temperature storage of minimally processed foods. *Foodborne Pathog. Dis.* **2013**, *10*, 413–419. [[CrossRef](#)]
66. Peck, M.W.; Stringer, S.C.; Carter, A.T. *Clostridium botulinum* in the post-genomic era. *Food Microbiol.* **2011**, *28*, 183–191. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.