Investigation of spontaneously fermented food matrices using high throughput DNA sequence-based analytical methods to explore microbial ecology

Francesca Cristetti (francesca.cristetti@unito.it)

Department of Agriculture, Forest and Food Sciences, University of Turin, Grugliasco, Italy

Tutor: Prof.ssa Kalliopi Rantsiou

This PhD thesis research project is aimed at investigating the drivers of a spontaneously fermented food matrix (e.g., coffee, cocoa, and table olives) through a multi-omics approach. The use of high throughput DNA sequence-based analytical methods will contribute to exploring microbial ecology and will be used as an indicator of origin to study the authenticity of raw materials.

Indagine su matrici alimentari fermentate spontaneamente utilizzando metodi analitici ad alto rendimento per il sequenziamento del DNA per esplorare l'ecologia microbica

Questo progetto di tesi di dottorato mira a studiare diverse matrici alimentari fermentate spontaneamente (es. caffè, cacao e olive da tavola) per mezzo di un approccio multi-omico. L’utilizzo di metodi di sequenziamento ad alto rendimento contribuirà ad esplorare l’ecologia microbica in modo da utilizzarla come indicatore di origine per studiare l’autenticità delle materie prime.

# **1. State-of-the-Art**

The microbial ecology of food concerns the study of the type of microorganisms present (diversity and population number), their rate of occurrence, activities (functionality), and interactions with each other (microbial communities) in the environment. In recent years, the study of the microbiome has become increasingly important to understand the interactions between microbial populations, especially in fermented foods as they are a complex matrix (De Filippis et al., 2018). Both culture-dependent and independent methods are essential for investigating food microbial ecology. However, it is estimated that 99% of microorganisms present in nature are typically not cultivated using standard laboratory techniques (Amann et al., 1995). The application of molecular methods has allowed the detection of populations that were largely undetected previously, thus generating a more accurate definition of the drivers involved in the fermentation process and their functions (Cocolin et al., 2013).

Cocoa and coffee beans have a complex microbial ecology that has a crucial role in the sensory characteristics respectively of chocolate and coffee because aroma and flavor precursor molecules are synthesized during the fermentation process. Thus, diversity could derive from yeast and bacterial populations present in the close environment during this process. When the fermentation occurs in a spontaneous or uncontrolled manner, the chemical and sensory quality of beans varies from location to location and can be compromised because of the non-controlled process. Hence, the variability of the fermenting microbial species from different regions remains one of the main factors influencing the variability of products’ quality (Figueroa-Hernandez et al., 2019).

Table olives, instead, are produced from the raw drupe fruits deriving from varieties of the cultivated olive tree. The raw fruit is inedible and highly bitter due to the presence of oleuropein, which needs to be degraded. Treatments to reduce the bitterness of the fruit involve brining, acidification, enzymatic hydrolysis, and spontaneous fermentation (Perpetuini et al., 2020). During these processes, the microbial ecology varies and may originate from the raw materials and the environment, in which it is possible to detect mixed yeasts, lactic acid bacteria (LAB), acetic acid bacteria (AAB), filamentous fungi, spoilage bacteria, and others (Serra et al., 2019).

The spontaneous process for cocoa and coffee involves yeasts that belong to the Saccharomycetaceae family including Hanseniaspora, Saccharomyces, Kluyveromyces, and Pichia that can produce ethanol from the fermentation of carbohydrates and can be responsible for flavour precursor production, with a concomitant increase in temperature. Lactic and acetic bacteria groups (e.g., Limosilactobacillus fermentum, Lactiplantibacillus plantarum, and Acetobacter pasterianus) are responsible for lactic acid and acetic acid production (Figueroa-Hernandez et al., 2019). In the case of table olives LAB (e.g., Lactiplantibacillus plantarum, Lactiplantibacillus pentosus) are responsible for oleuropein degradation thanks to their enzymatic activities and cause the acidification of the final product, providing microbial stability and elongating the shelf-life of the olives while yeasts (e.g., Saccharomyces cerevisiae, Candida boidinii) are involved in the production of volatile compounds and metabolites that increase the overall quality and preserve their sensory features (Perpetuini et al., 2020).

Next-generation sequencing (NGS) is fundamental to obtain a deeper and more detailed characterization of the microbiome. In amplicon sequencing, template DNA is fragmented, bound to a substrate, and amplified by PCR to produce clonal representations of the original spatially separated fragments for subsequent sequencing, while shotgun sequencing requires genomes to be broken into small fragments, sequenced, and reassembled using overlapping sequences (Bedia, 2018). These techniques are used to describe in great detail microbial communities' structure and function and the information deduced is used as an indicator of origin to study the authenticity of raw materials.

# **2. PhD Thesis Objectives and Milestones**

Within the overall objective mentioned above, this PhD thesis project can be subdivided into the following activities according to the Gantt diagram given in Table 1.

A1) **Literature review and experimental design definition**

A2) **Optimization of sample preparation and DNA extraction protocols compatible with high-throughput sequencing** to study the entire microbial ecology (yeasts and bacteria) of the different matrices. The protocols will be tested to understand which method will be useful for subsequent Next Generation Sequencing to obtain a deeper and more detailed characterization of the microbial ecology.

A3) **Sequencing and bioinformatics analysis** to investigate16S, and 26S rRNA genes amplicon (metataxonomy) on a large number of samples. The data will be analyzed using bioinformatics tools to obtain the entire microbial ecology of the different matrices and to understand the differences between the origins.

A4) **Metabolome analysis** using gas chromatography-mass spectrometry (GC–MS), and liquid chromatography-mass spectrometry (LC-MS) to understand if the different origins of the matrices will also influence the production of secondary compounds and thus the final quality of the product.

A5) **Marker analysis for safety, quality, and authenticity,** correlating microbiological and metabolite trends with the origin of the different matrices. Finally, links between microbial ecology, metabolites, and abiotic factors (e.g., climatic parameters, interventions in primary production, and local practices) will be investigated to identify markers of safety, quality, and authenticity.

A6) **Writing and Editing** of the PhD thesis, scientific papers and oral and/or poster communications.

***Table 2***Gantt diagram for this PhD thesis project.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Activity Months | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** | **21** | **22** | **23** | **24** |
| A1) | ***Literature review and experimental design definition*** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| A2) | ***Optimization of sample preparation and DNA extraction protocols compatible with high-throughput sequencing*** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| A3) | ***Sequencing and bioinformatics analysis*** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| A4) | ***Metabolome analysis*** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| A5) | ***Marker analysis for safety, quality, and authenticity*** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| A6) | ***Thesis and Paper Preparation*** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

# **3. Selected References**

Amann R.I., Ludwig W., Schleifer K.H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation, Microbiol Rev, 43–169.

Bedia. (2018) Experimental Approaches in Omic Sciences, Comprehensive Analytical Chemistry, 13-32.

Cocolin L., Alessandria V., Dolci P., Gorra R., Rantsiou K. (2013) Culture independent methods to assess the diversity and dynamics of microbiota during food fermentation, International Journal of Food Microbiology, 29-43.

De Filippis F., Parente E., Ercolini D. (2018) Recent Past, Present, and Future of the Food Microbiome, Annual Review of Food Science and Technology, 589–608.

Figueroa-Hernandez C., Mota-Gutierrez J., Ferrocino I., Hernandez-Estrada Z.J., Gonzalez-Ríos O., Cocolin L., Suarez-Quiroz, M.L. (2019) The challenges and perspectives of the selection of starter cultures for fermented cocoa beans, International Journal of Food Microbiology, 41-50.

Perpetuini, G., Prete, R., Garcia-Gonzalez, N., Alam, M. K. & Corsetti, A. (2020) Table olives more than a fermented Food, Foods 9, 1–16

Serra, J. L., Moura, F. G., de Melo Pereira, G. V., Soccol, C. R., Rogez, H., & Darnet, S. (2019) Determination of the microbial community in Amazonian cocoa bean fermentation by Illumina-based metagenomic sequencing, Lwt, 106, 229-239.