UNIVERSITÀ DEGLI STUDI DI PADOVA

UBO
Université de Bretagne Occidentale

Head Office: Università degli Studi di Padova & Université de Bretagne Occidentale

Department: Department of Comparative Biomedicine and Food Science & Laboratoire Universitaire de Biodiversité et Écologie Microbienne

---

Ph.D. course in: Veterinary Science & Ecole doctorale Ecologie, Géosciences, Agronomie et Alimentation
Specialitè Microbiologie, Virologie, Parasitologie
SERIES: XXXV

Innovative biomarker for authentication and qualitative assessment of dairy products

Coordinator: Ch.mo Prof. Mattia Cecchinato
Supervisor: Ch.mo Prof. Barbara Cardazzo & Emmanuel Coton
Co-Supervisor: Ch.mo Prof. Enrico Novelli & Jérôme Mounier

Ph.D. student: Marco Cardin

рукописи не горят…

Acknowledgements

I am extremely grateful to the people who have made this journey possible and accompanied me along the way. From my supervisors Emmanuel Coton and Enrico Novelli, who have guided me along the way, to my colleagues in Italy, and the warm people of LUBEM (especially Beatriz, Melanie, and Marie) who became my family in France, I extend my heartfelt gratitude.

A special acknowledgment goes to my long-lasting friends. I don't know half of you half as well as I should like, and I like less than half of you half as well as you deserve.

Nevertheless, without my family, Adriano, Luciana, Luca, and Ksenia, I probably wouldn't have endured these dark months.

You will always have my heart and my thoughts.

Lastly, I would like to greet you, Ksenia. I am proud of your work and courage.

We will beat it.

General summary

Traditional cheeses, especially those with protected geographical indications, are highly susceptible to food fraud due to consumers' willingness to pay a premium for authentic and traditional products. To combat this issue, reliable food authentication methods are essential. This PhD thesis focuses on the study and performance evaluation of the authentication methods used for cheese geographical origin verification, including chemical, physical, and DNA-based approaches. The first study, a literature review, highlighted that isotope and elemental fingerprinting methods had consistent accuracy in cheese origin authentication and emphasized the need for more research to assess the discriminative power of other methods, such as near-infrared spectroscopy, nuclear magnetic resonance, and DNA-based techniques. DNA-based techniques, particularly metabarcoding, showed promising results for origin authentication, but metagenomics, which offers a detailed view of the cheese microbiota down to the strain level, presented greater potential for enhancing authentication capabilities.

The second study focused on mountain Caciotta cheese, employing shotgun metagenomics and volatilomics to understand the factors contributing to cheese typicity and diversity. Geographical origin, alongside factors like curd cooking temperature, pH, salt concentration, and water activity, played a significant role in shaping the cheese's distinctive characteristics. Notably, viral communities exhibited higher biodiversity and effectively discriminated cheese origins. Among the dominant bacteria, *Streptococcus thermophilus* displayed higher intraspecific diversity and stronger associations with cheese origin compared to *Lactobacillus delbrueckii*. The study also identified non-starter lactic acid bacteria and phages specific to each origin, providing valuable insights into the cheese's unique microbial composition. The volatilome of mountain Caciotta cheese showcased prominent levels of alcohols and ketones, with distinct differences in the relative abundances of enzymes linked to flavor development, further contributing to its typicity.

Lastly, a multi-omics approach for cheese origin authentication was explored, incorporating shotgun metagenomics, volatilome analysis, near-infrared spectroscopy, stable isotopes, and elemental analysis. DNA-based analysis, particularly viral communities, achieved high

classification accuracy rates compared to bacterial communities. Volatile organic compounds demonstrated potential for clustering cheese according to its origin, while near-infrared spectroscopy showed moderate classification accuracy. Elemental composition analysis revealed significant variations in elements related to dairy equipment, macronutrients, and rare earth elements across different origins, leading to the highest performances in origin authentication.

This work does have some limitations that could spur future research in cheese science and origin authentication, including the absence of culture-based microbiological analysis and the need for further investigations employing different classification models. Additionally, future research is encouraged to explore diverse traditional cheese varieties and geographical locations to validate the findings extensively. The integration of innovative tools, such as miniaturized vibrational spectroscopy devices and blockchain systems, holds promise for enhancing food authentication practices.

In conclusion, this study underscores the potential of multi-omics techniques, particularly metagenomics and volatilomics, in cheese origin authentication. The case study focusing on mountain Caciotta cheese provides valuable insights into the role of geographical factors in shaping the characteristics of artisanal cheeses. Overall, this research can contribute to the preservation of cultural heritage and economic sustainability in mountainous regions through robust authentication systems for traditional cheese products.

Résumé général

Les fromages traditionnels, en particulier ceux bénéficiant d'indications géographiques protégées, sont très vulnérables à la fraude alimentaire en raison de la volonté des consommateurs de payer un supplément pour des produits authentiques et traditionnels. Pour lutter contre ce problème, des méthodes fiables d'authentification alimentaire sont essentielles. Cette thèse se concentre sur la valutation dell'état actuel des méthodes d'authentification utilisées pour vérifier l'origine géographique des fromages, notamment les approches chimiques, physiques et basées sur l'ADN. La première étude, une revue de la littérature, a mis en évidence que les méthodes d'analise isotopique et des éléments présentaient une précision constante dans l'authentification de l'origine des fromages et a souligné la nécessité de davantage de recherches pour évaluer le pouvoir discriminant d'autres méthodes, telles que la spectroscopie proche infrarouge, la résonance magnétique nucléaire et les techniques basées sur l'ADN. Les techniques basées sur l'ADN, comme le métabarcodage, ont montré des résultats prometteurs pour l'authentification de l'origine du fromage, mais la métagénomique, qui offre une vue détaillée de la microbiote du fromage jusqu'au niveau des souches, présente un potentiel plus élevé pour améliorer les capacités d'authentification.

La deuxième étude s'est concentrée sur le fromage Caciotta de montagne, en utilisant la métagénomique à la methode shotgun et la volatilomique pour comprendre les facteurs contribuant à la typicité et à la diversité du fromage. L'origine géographique, ainsi que des facteurs tels que la température de coagulation du caillé, le pH, la concentration en sel et l'activité de l'eau, ont joué un rôle significatif dans la formation des caractéristiques distinctives de la microbiote du fromage. Notamment, les communautés virales ont montré une biodiversité plus élevée et ont efficacement discriminer les origines des fromages. Parmi les bactéries dominantes, *Streptococcus thermophilus* a montré une plus grande diversité intraspécifique et des associations plus fortes avec l'origine du fromage par rapport à *Lactobacillus delbrueckii*. L'étude a également identifié des bactéries lactiques non démarreurs et des phages spécifiques à chaque origine, fournissant des informations précieuses sur la composition microbienne unique du fromage. Le volatilome du fromage Caciotta de

montagne a présenté des niveaux élevés d'alcools et de cétones, avec des différences distinctes dans les abondances relatives des enzymes liées au développement du goût, contribuant ainsi davantage à sa typicité.

Enfin, une approche multi-omique pour l'authentification de l'origine du fromage a été explorée, en intégrant la métagénomique, l'analyse du volatilome, la spectroscopie proche infrarouge, les isotopes stables et l'analyse élémentaire. L'analyse basée sur l'ADN, en particulier les communautés virales, a atteint des taux de précision de classification élevés par rapport aux communautés bactériennes. Les composés organiques volatils ont démontré le potentiel de regrouper le fromage en fonction de son origine, tandis que la spectroscopie proche infrarouge a montré une précision de classification modérée. L'analyse de la composition élémentaire a révélé des variations significatives dans les éléments liés à l'équipement laitier, aux macronutriments et aux terres rares entre différentes origines, conduisant aux meilleures performances en termes d'authentification de l'origine.

Ce travail présente certaines limites qui pourraient stimuler la recherche future dans le domaine de la science du fromage et de l'authentification de l'origine, notamment l'absence d'analyses microbiologiques basées sur la culture et la nécessité de mener d'autres enquêtes utilisant différents modèles de classification. De plus, des recherches futures sont encouragées pour explorer diverses variétés traditionnelles de fromages et des lieux géographiques différents afin de valider les résultats de manière approfondie. L'intégration d'outils innovants, tels que des dispositifs de spectroscopie vibrationnelle miniaturisés et des systèmes de blockchain, promet d'améliorer les pratiques d'authentification alimentaire.

En conclusion, cette étude met en évidence le potentiel des techniques multi-omiques, en particulier la métagénomique et la volatilomique, dans l'authentification de l'origine du fromage. L'étude de cas axée sur le fromage Caciotta de montagne fournit des informations précieuses sur le rôle des facteurs géographiques dans la formation des caractéristiques des fromages artisanaux. Dans l'ensemble, cette recherche contribue à la préservation du patrimoine culturel et à la durabilité économique dans les régions montagneuses grâce à des systèmes d'authentification solides pour les produits de fromage traditionnels.

Riassunto generale

La disponibilità dei consumatori a pagare un premium price per prodotti come i formaggi tradizionali, in particolare quelli con indicazioni geografiche protette, li rende altamente suscettibili alle frodi alimentari. Per combattere questo problema, sono essenziali metodi affidabili di autenticazione degli alimenti. Questa tesi di dottorato si concentra sullo studio e valutazione delle performance dei metodi di autenticazione utilizzati per la verifica dell'origine geografica del formaggio, inclusi approcci chimici, fisici e basati sul DNA. Il primo studio, una review della letteratura scientifica, ha evidenziato che l'analisi degli elementi e degli isotopi avevano un'accuratezza costante nell'autenticazione dell'origine del formaggio e ha sottolineato la necessità di ulteriori ricerche per valutare il potere discriminante di altri metodi, come la spettroscopia nel vicino infrarosso, la risonanza magnetica nucleare e tecniche basate sul DNA. Le tecniche basate sul DNA, in particolare il metabarcoding, hanno mostrato risultati promettenti per l'autenticazione dell'origine, ma la metagenomica, che offre una visione dettagliata del microbioma del formaggio fino al livello del ceppo, potrebbe avere un potenziale maggiore per migliorare le capacità di autenticazione.

Il secondo studio si è concentrato sulla Caciotta di montagna, impiegando metagenomica e volatiloma per comprendere i fattori che contribuiscono alla tipicità e alla diversità microbiologica del formaggio. L'origine geografica, insieme a fattori come la temperatura di cottura della cagliata, il pH, la concentrazione di sale e l'attività dell'acqua, hanno giocato un ruolo significativo nel plasmare le caratteristiche distintive del microbioma del formaggio. In particolare, le comunità virali hanno mostrato una maggiore biodiversità e hanno discriminato efficacemente le origini del formaggio. Tra i batteri dominanti, *Streptococcus thermophilus* ha mostrato una maggiore diversità intraspecifica e associazioni più forti con l'origine del formaggio rispetto a *Lactobacillus delbrueckii*. Lo studio ha anche identificato batteri lattici non starter e fagi specifici per ciascuna origine, fornendo preziose informazioni sull'unicità della composizione microbica del formaggio. Il volatiloma della Caciotta di montagna presentava livelli elevati di alcoli e chetoni, con nette differenze nelle

abbondanze relative di enzimi legati allo sviluppo dei composti aromatici che contribuiscono alla formazione dei suoi caratteri tipici.

Infine, è stato esplorato un approccio multi-omico per l'autenticazione dell'origine del formaggio, che incorpora la metagenomica shotgun, l'analisi del volatiloma, la spettroscopia nel vicino infrarosso, gli isotopi stabili e l'analisi elementare. L'analisi basata sul DNA, in particolare le comunità virali, ha raggiunto tassi di accuratezza della classificazione elevati e migliori rispetto alle comunità batteriche. I composti organici volatili hanno dimostrato un buon potenziale per classificare il formaggio in base alla sua origine, mentre la spettroscopia nel vicino infrarosso ha mostrato una moderata precisione di classificazione. L'analisi della composizione elementare ha rivelato variazioni significative negli elementi relativi alle attrezzature casearie, ai macronutrienti e alle terre rare per le diverse origini, portando migliori prestazioni nell'autenticazione dell'origine.

Questo lavoro presenta alcune limitazioni che potrebbero stimolare la ricerca futura nella scienza del formaggio e nell'autenticazione dell'origine, tra cui l'assenza di analisi microbiologiche basate sui metodi classici di coltura microbiologica e la necessità di ulteriori indagini che utilizzino diversi modelli di classificazione. Inoltre, si incoraggia la futura ricerca ad esplorare diverse varietà di formaggi tradizionali e provenienti da più aree geografiche per convalidare ulteriormente i risultati ottenuti. L'utilizzo di innovativi, come dispositivi di spettroscopia vibrazionale miniaturizzati e sistemi blockchain, dovrebbero essere altresì considerate al fine di migliorare le pratiche di autenticazione degli alimenti e combattere le frodi.

In conclusione, questo studio sottolinea il potenziale delle tecniche multi-omiche, in particolare la metagenomica e la volatiliomica, nell'autenticazione dell'origine del formaggio. La Caciotta di montagna, oggetto di questo caso studio, fornisce preziosi spunti sul ruolo dei fattori geografici nel plasmare le caratteristiche dei formaggi tradizionali. Nel complesso, questa ricerca può contribuire alla conservazione del patrimonio culturale e alla sostenibilità economica nelle regioni montuose attraverso robusti sistemi di autenticazione per i prodotti caseari tradizionali.

Table of contents

General introduction

Cheese is a widely consumed dairy product. In 2021, the global market value of cheese stood at roughly 77.6 billion U.S. dollars, a number which is projected to exceed the 113 billion U.S. dollar mark by 2027 (Imarc, 2022). Cheese production has been increasing slightly with each consecutive year since 2015 and about half of all the cheese made in 2021 was produced by the European Union (USDA, 2022). However, as a wide variety of cheese is known, linked to the use of different milk types, process and know-how, differences in market value of different cheese types exist. In agreement with a societal demand for less processed and traditional foods, cheeses benefiting a traditional image are well considered and can have a higher economic value. As example, in Poland 68% of interviewed consumers declared that they would always choose traditional cheese (Roślinnie Jemy, 2021) while in Italy 70.1% of interviewed consumers declare that the place of origin is the most important element that would characterize their choice (Pilone et al., 2015).

The European Union is committed to protecting geographical indications to preserve the unique qualities of products linked to their geographical origin and traditional methods of production. Since the introduction of Council Regulations in 1992 (EEC No. 2081/92 and No. 2082/92), the European Union has introduced a range of quality schemes including Protected Designation of Origin (PDO), Protected Geographical Indication (PGI), and Traditional Speciality Guaranteed (TSG), as well as the optional quality term "Mountain Product" (EU Parliament and Council Regulation No. 1151/2012, art. 31, and Regulation No. 665/2014). Additionally, there are international and national voluntary certifications, such as the EU organic certification (Council Regulation EEC No. 2092/91, currently under EU Regulation 2018/848 of the European Parliament and of the Council) and national organic labels in some EU countries (*e.g.*, Agriculture Biologique in France), that aim to improve environmental sustainability and animal welfare.

These quality schemes offer a range of benefits, including premium pricing directly benefiting the producers, positive socio-economic impacts for territories and a variety of ecosystem services such landscape and environmental preservation and safeguarding of

biodiversity and cultural knowledge (Mancini and Donati, 2013; Deselnicu et al., 2013; Arfini et al., 2019). In light of these benefits, recent projects have highlighted the potential for food quality schemes to promote more sustainable production patterns (Arfini and Bellassen, 2019). At the same time, consumers benefit from food quality schemes in terms of assurance of high-quality standards, informed choices, fair competition, and reliable information. Studies have explored consumers' preferences for food quality schemes, including their willingness to pay for products with food quality labels compared to conventional ones (Menapace et al., 2011; Aprile et al., 2012; Garavaglia and Mariani, 2017; Santeramo and Lamonaca 2020). For instance, Spanish consumers are more willing to pay similar price premiums for PDO and organic cheese than for reduced-fat-content cheese. Another area of study has been consumer interest in geographical indications combined with the optional quality term "Mountain Product", with consumers showing positive attitudes towards this label and its combination with PDO (Sanjuan and Khliji, 2016; Menozzi et al., 2022). This reflects a growing focus on protecting natural resources and supporting small farmers and local traditions. As geographical indication quality is recognised and appreciated by consumers, their willingness to pay higher prices for this type of product (Figure 1) is often an attraction for fraudulent action such as origin mislabelling (García-Hernández et al., 2022).

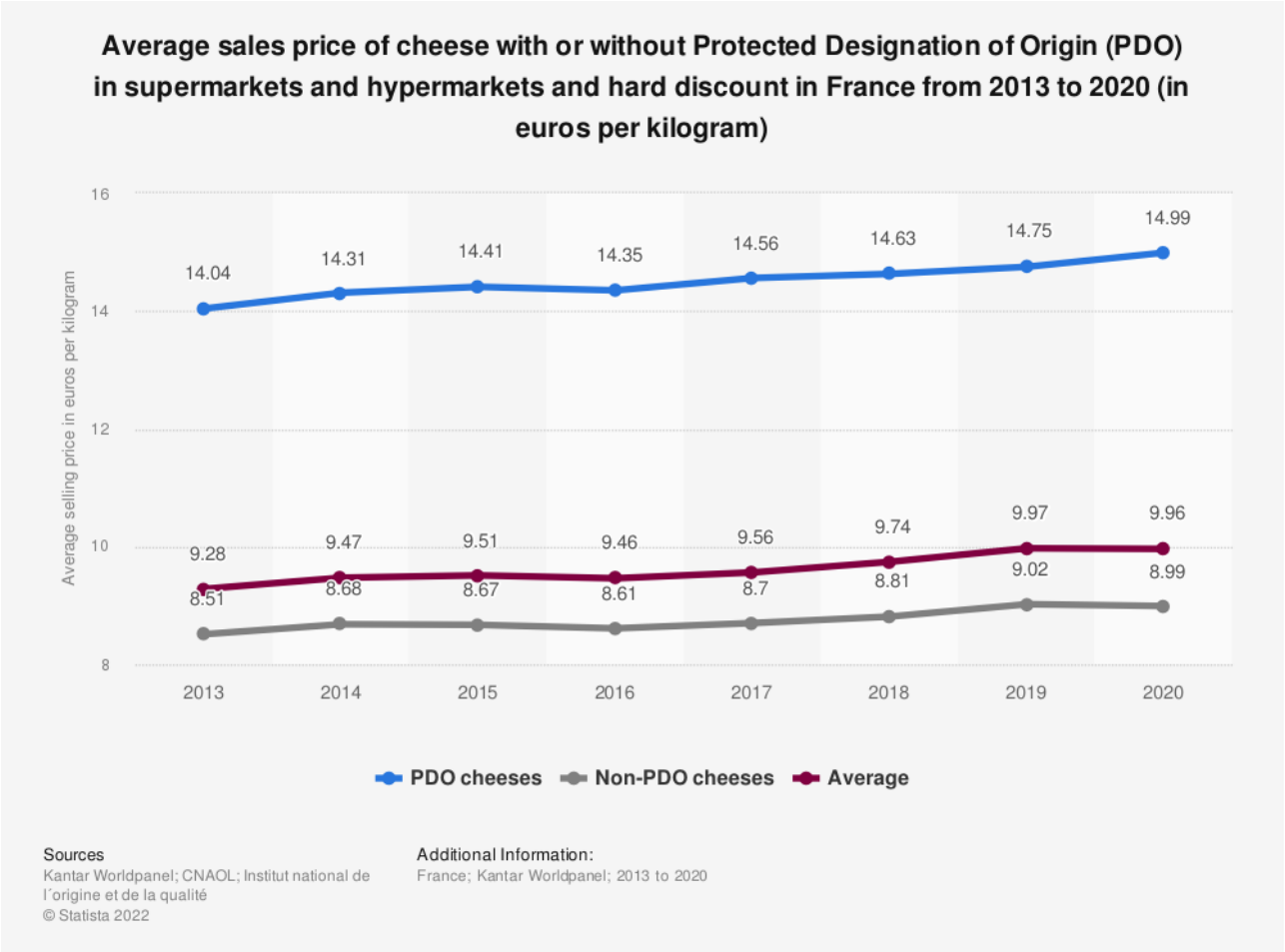


Figure 1. Differences in average price of PDO cheese in comparison to conventional cheese.

In particular, when food frauds are related to food categories such as protected designation of origin (PDO) or mountain products, they affect crucial ecosystems, decreasing the products' value, and compromising the economic and social sustainability of the production. The European food legislator has implemented a strict traceability system based on product labelling and protected the value of these products. However, the implementation of these systems needs appropriate testing methods to protect the consumer against frauds and to ensure their trust (Barcaccia et al., 2015). In this context, implementing a system to authenticate food origin and evaluate their quality may provide additional tools to counteract the problem.

Food authentication is the process that verifies that a food is in compliance with its label description (Danezis et al., 2016). This may include production methods (*e.g.* conventional *vs*. organic), processing technologies (e.g. use of raw *vs*. pasteurised milk) and origin. The

origin of the cheeses may have its identity related to the edaphoclimatic conditions established by the environment, in which the climate influences the composition and quality of the raw material used (Silva et al., 2022). The location where the cheese is produced and the herd management can also affect the composition of the milk microbiota, leading to unique biochemical changes in the cheese (Reuben et al., 2023). Furthermore, the use of local ingredients and traditional know-how in the cheesemaking processes can contribute to a set of characteristics that are specific to the region, associated with the "terroir" notion that reflects the essence of the area where it was made (Karoui et al., 2005a; Karoui et al., 2005b). Various methods have been employed to authenticate cheese origin. These can be classified in chemical, physical and DNA-based. Chemical methodologies frequently employed in cheese origin authentication include the analysis of stable isotope ratio, trace elements, fatty acids and volatile organic compounds. On the other hand, physical methods include infrared spectroscopy and nuclear magnetic resonance. While these methods have largely been used for food authentication, other methods based on the food-associated microbiota and its metabolism are developing.

In this context, DNA-based methods refer to the possibility of creating a molecular fingerprint based on the cheese microbiome. Cheese microbiome can be seen as a complex assembly of microorganisms that originate from farmland and are shaped through cheesemaking technology and storage/ripening conditions until consumer consumption (Cardin et al., 2022). Metabarcoding and shotgun metagenomics sequencing are two DNA based techniques able to produce consistent amounts of data using low sample quantities. These analyses have seen a continuous decrease of cost per sample (Komarova et al., 2020), making them a useful tool to characterize microbial community and identify microbial signatures. Many authors have reported a greater potential in microbial community description using metabarcoding and metagenomics than phenotypical analysis, since multiple microorganisms as bacteria, virus, fungi and archaea can be detected in the same analysis (Sattin et al., 2016; Kamilari et al., 2019; Afshari et al., 2020). The application of metabarcoding or shotgun metagenomics sequencing could offer other interesting benefits

allowing to authenticate dairy products (*i.e.* specific "house" signature) and to investigate their quality at the same time.

Cheese microbiota diversity and the possibility of establishing a microbial signature from different processing environments, named "house microbiota", have been described by multiple authors (Bokulich and Mills, 2013; Calasso et al., 2016). However, despite detailed bacterial community characterization, most of the researchers have focused only on 16S rRNA marker and only few of them have combined DNA and metabolomic analysis. Moreover, it is well known that the 16S rRNA marker can discriminate bacteria up to the family/genus level, while cheese has complex microbiota in which strains from bacteria, fungi and viruses might play a key role in microbiota dynamics. These disadvantages can be overcome by shotgun metagenomics (Quince et al., 2017). This method, via the analysis of all the genetic material available in the sample, can assess the taxonomic classification of bacteria, fungi and virus (Ranjan et al., 2016). Moreover, shotgun metagenomics offers multiple advantages in comparison with metabarcoding, since no amplification step is required and all the genetic material in the sample is used for genome reconstruction to reach a deeper taxonomic assignation, potentially to the strain level (Ranjan et al., 2016; Quince et al., 2017). Moreover, beyond the diversity information, microbial functionalities can be accessed.

Nevertheless, few studies have employed shotgun metagenomics to characterize and/or investigate the possibility to authenticate cheese origin. Parente et al. (2020) highlighted the need for more research to understand dairy microbiota, since 50% of the published works have used 24 samples or less. Hence, the use of shotgun metagenomics, combined with a systematic sampling for at least two years is needed to gain a deeper understanding of the microbial community and achieve the goal of food origin authentication (Montel et al., 2014; Medina et al., 2019).

In the framework of this PhD work, we used typical mountain Caciotta cheese to investigate the performances of different origin authentication methods including shotgun metagenomics as well as volatilome which is an aspect of the microbiota expression.

Caciotta is a type of Italian cheese that is made from raw cow's milk. It is a traditional, short-medium ripened cheese characterized by dark ivory rind color and yellow-pale core (Figure 2A)(Bancalari et al., 2020). The production of Caciotta from northern Italy mountain area is often based on small-scale dairy employing traditional methods. Indeed, mountain areas (Figure 2B) such as Alti Pascoli della Lessina and Trento province are known and recognized for their cheesemaking history (Apolito 2018).

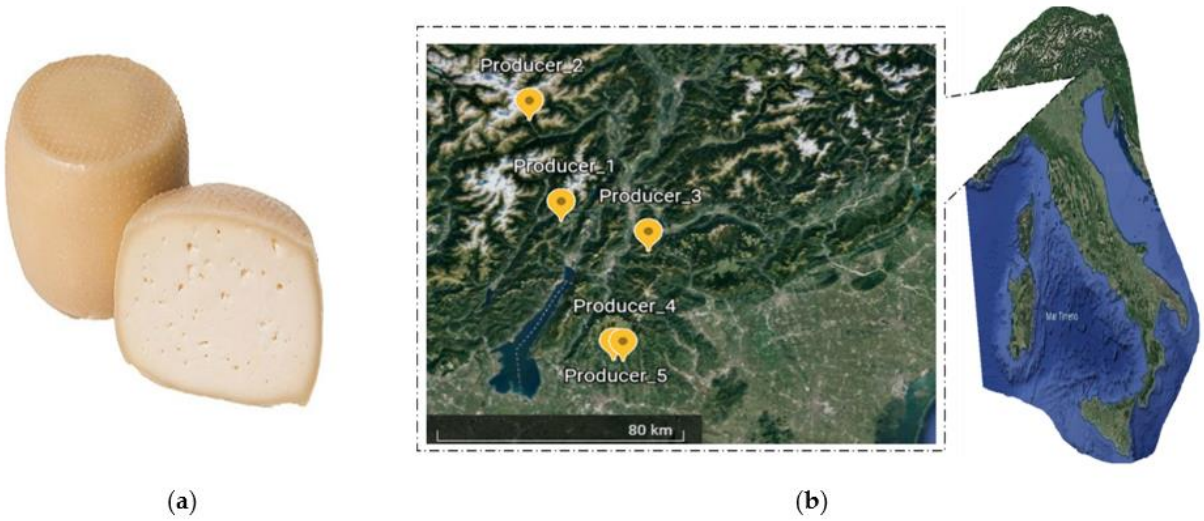


Figure 2. (**a**) Picture of typical Caciotta cheese. (**b**) Map of the origin of mountain Caciotta producers that was investigated in this study. Producer 1 is located in the Giudicarie Esteriori area while producers 2 and 3 are located in the Trento province. Producers 4 and 5 are located in the Alti Pascoli della Lessinia area.

Overall, the scientific questions addressed by this work were as follow:

- What are the main ecological drivers of typical Caciotta cheese produced in different regions with similar cheesemaking know-how?
- Is the cheese microbiome stable along time? Is it possible to identify considerable changes in cheese microbiome according to season or year?
- Can cheese microbiome be used as an authentication marker?
- What are the performances of shotgun metagenomics, volatilome, near infrared spectroscopy, stable isotope ratio and trace element analysis in authenticating the origin of typical mountain cheese?

Herein, the first chapter of this document is focused on a comprehensive literature review about geographical origin authentication methods, considering dairy frauds, European legislature and typicity definition. The second chapter is focused on the characterization of viral and bacterial communities of typical mountain Caciotta cheese and its associated volatilome. The following chapter reports a preliminary investigation on the ability of shotgun metagenomics data and volatilome analysis in authenticating cheese origin by using a cross-validated spatial partial least square discriminant analysis model. In conclusion, a general comparison of physical, chemical, and DNA-based methods for origin authentication will be provided.

Chapter 1: What methods can be used to authenticate cheese origin?

In this chapter the available origin authentication methods are discussed. The choice of origin authentication method is influenced by economic and legal considerations, as well as reported cases of frauds. While the chapter aims to provide a general introduction of the frauds, geographic indications and typicity terminology, the main body reports a critical view of the recent 179 papers dealing with origin authentication. Among the reported definition "Mountain product", which identifies "products for which the raw material and the feedstuffs for animals come essentially from mountain areas and, in the case of processed products, the processing also takes place in mountain areas" (EU Parliament and Council Regulation No. 1151/2012) will be further considered in the characterization of typical Caciotta cheese (Chapter 2) and in the authentication of its origin (Chapter 3). Chemical, physic, and DNA-based methods are described by providing the main analytical principles as well as authentication performances of cheese origin. Since DNA-based methods are emerging in this field, pivotal considerations on the factors affecting cheese microbiota (*i.e*. the consortium of prokaryotic, eukaryotic and viral populations inhabiting cheese) are provided. In the end, an overall consideration of the described methods and future perspectives aims to orient the reader toward the most suitable way to authenticate cheese origin.

The following is a review published in the journal *foods*.

Cardin, M., Cardazzo, B., Mounier, J., Novelli, E., Coton, M., & Coton, E. (2022). Authenticity and typicity of traditional cheeses: A review on geographical origin authentication methods. *Foods*, *11*(21), 3379. https://doi.org/10.3390/foods11213379

foods

MDPI

Review

Authenticity and Typicity of Traditional Cheeses: A Review on Geographical Origin Authentication Methods

Marco Cardin [1,2], Barbara Cardazzo [1,*], Jérôme Mounier [2], Enrico Novelli [1], Monika Coton [2] and Emmanuel Coton [2]

[1] Department of Comparative Biomedicine and Food Science, University of Padova, Viale Università 16, 35020 Legnaro, PD, Italy
[2] Univ Brest, INRAE, Laboratoire Universitaire de Biodiversité et Écologie Microbienne, F-29280 Plouzané, France
* Correspondence: barbara.cardazzo@unipd.it

Abstract: Food fraud, corresponding to any intentional action to deceive purchasers and gain an undue economical advantage, is estimated to result in a 10 to 65 billion US dollars/year economical cost worldwide. Dairy products, such as cheese, in particular cheeses with protected land- and tradition-related labels, have been listed as among the most impacted as consumers are ready to pay a premium price for traditional and typical products. In this context, efficient food authentication methods are needed to counteract current and emerging frauds. This review reports the available authentication methods, either chemical, physical, or DNA-based methods, currently used for origin authentication, highlighting their principle, reported application to cheese geographical origin authentication, performance, and respective advantages and limits. Isotope and elemental fingerprinting showed consistent accuracy in origin authentication. Other chemical and physical methods, such as near-infrared spectroscopy and nuclear magnetic resonance, require more studies and larger sampling to assess their discriminative power. Emerging DNA-based methods, such as metabarcoding, showed good potential for origin authentication. However, metagenomics, providing a more in-depth view of the cheese microbiota (up to the strain level), but also the combination of methods relying on different targets, can be of interest for this field.

Keywords: cheese; geographical origin; authentication; next-generation sequencing; volatilome; isotopic analysis; trace element analysis; infrared fingerprinting

Citation: Cardin, M.; Cardazzo, B.; Mounier, J.; Novelli, E.; Coton, M.; Coton, E. Authenticity and Typicity of Traditional Cheeses: A Review on Geographical Origin Authentication Methods. *Foods* **2022**, *11*, 3379. https://doi.org/10.3390/foods11213379

Academic Editors: Andrea Marchetti and Simona Sighinolfi

Received: 3 October 2022
Accepted: 22 October 2022
Published: 26 October 2022

1. Introduction

The shared definition of food fraud relates to intentional illegal acts performed by food value chain operators for economic gain [1]. More specifically, in the framework of the European agri-food chain legislation, food fraud is defined as "any suspected intentional action by businesses or individuals for the purpose of deceiving purchasers and gaining undue advantage therefrom, in violation of the rules referred to in Article 1(2) of Regulation (EU) 2017/625" [2]. Behind the term "food fraud", multiple practices designed to deceive purchasers, which are categorized under the following denominations: (i) substitution, (ii) concealment, (iii) dilution, (iv) unapproved enhancement, (v) counterfeit, (vi) grey market/forgery, and finally, (vii) mislabeling, exist. Substitution corresponds to total or partial replacement of a food, including ingredients or nutrients, with one of lower value. Concealment hides the low quality of food ingredients or food products. Dilution is self-explanatory and corresponds to the action of mixing a high-value ingredient with a lower one, while unapproved enhancement improves food quality by adding undeclared or unknown ingredients. These four food fraud types are grouped under the term "adulteration". Counterfeit refers to the infringement of Intellectual Property Rights via replication of a product or its packaging, while grey market or forgery corresponds to production, theft,

and diversion involving unauthorized sales of foodstuffs. The latter generally concerns products for which production agreements or quotas exist or geographical restrictions apply. Finally, mislabeling corresponds to distorted information or false claims on packaging or labels [3].

The global estimated value of food frauds each year ranges from 10 to 65 billion US dollars, without considering potentially related losses [1]. Indeed, unfair competition may not only result in economic losses for honest producers and retailers but could also impact food safety and quality, public health, and society at a large scale, thus impeding a perfectly accurate estimation of food fraud socio-economic impacts [4,5].

In recent years, dairy products have been systematically listed among the most common food frauds [6–11], with cheese being the most prevalent [11,12]. In this case, fraudulent documentation and adulteration/substitution were the most frequent events [12,13]. Between 2000 and 2018, the HorizonScan program (a subscription-based service monitoring global food integrity issues, including brand identity) reported 245 cases of dairy frauds, from which 51% were characterized by fraudulent documentation [12]. Similar findings were reported by Montgomery et al. [13], who stated that cheese fraudulent documentation accounted for 74% of total fraud cases (n = 98) between 2015 and 2019 [10,11]. In this context, actions against food fraud are taken by inspecting agencies, producers, and retailers [14–17]. Nevertheless, the increased complexity of a globalized supply chain can impair fraud incidents from being detected.

The term "authenticity" for food products is associated with the fact that there is a "match between the food product characteristics and the corresponding food product claims" [18]. In this context, cheeses are defined as ripened or unripened soft, semi-hard, hard, or extra-hard products, obtained by milk protein coagulation using rennet, other suitable coagulating agents, or processing technologies, and have whey protein/casein ratios that do not exceed that of milk [19]. Cheese quality is often linked to value descriptors such as environmental welfare standards, production methods, and safety claims, but also geographical origin [20].

As for the term "typicity", it is defined by the unique combination of natural and human factors associated with a specific terroir [21]. Cheese typicity (i.e., the recognizable organoleptic traits associated with a given cheese) is acquired from the specific raw materials used, traditional tools, and the encountered environmental and production conditions, cheese-making process, and geographical area [22]. At the European level, linked to this typicity, certain cheeses can be recognized with distinctive labels, such as Protected Designation of Origin (PDO), which indicates that the products were entirely manufactured in a defined geographical area, Protected Geographical Indication (PGI), which correlates a geographical area with at least one of the product transformation steps, and Traditional Specialties Guaranteed (TGS), that highlights a traditional aspect for food products without any link to a specific geographical area. Other labels, such as "product of island farming" or "mountain product", can also be found [23–25]. A strong societal demand currently exists for natural, local, traditional, and authentic foods. Authentic food quality is recognized by the consumer to have a higher added value, but this higher value increases the risk for fraud, in particular fraudulent documentation including omission or irregular use of geographical origin and failure to adopt suitable traceability systems (i.e., corresponding to mislabeling) [13,26]. At the European level, Regulation [27] established the implementation of a comprehensive traceability system within food businesses and required a suitable documentation system to identify the product along the food chain, while Regulation [28] (Art. 26) requires labeling "country of origin" for products such as Geographical Indication and meat or products, for which mislabeling would mislead the consumer. Moreover, the increasing consumer demand for natural, local, and traditional foods has led to national laws, such as in Italy and Spain, regulating how to label the geographical origin of milk and milk derivatives [29,30]. Accordingly, geographical origin, described as a "specific location" that serves to designate a product origin such as territory of a member, region, or locality in a given territory [31] (Art. 22), has crucial relevance in dairy products. The need for food

authentication methods is driven by different actors. This includes producers and retailers for whom food fraud induces economical losses, public authorities that verify compliance with agri-food chain legislation, and finally, the consumer, to ensure trust when buying a product. While, in the past, authentication analyses focused on evaluating a single molecule or single parameters, nowadays, these methods are evolving from targeted to untargeted approaches. This enables the description of multiple product features and characteristics to provide a way to develop fingerprints for cheese geographical origin authentication. This is particularly relevant for protected land- and tradition-related labeled cheeses that are often incriminated in food frauds due to their high economical value.

Two different strategies can be employed to authenticate cheese origin. The first one involves exploring the relationship between the biological and/or chemical components, assuming that their proportions are constant for a particular cheese at a specific time during production or shelf-life. In this context, it seems clear that pattern recognition methods (e.g., such as principal component analysis (PCA) and orthogonal projections to latent structures discriminant analysis (OPLS-DA)) create unique classes, potentially differentiating typical and fraudulent products [32]. The second strategy aims to find specific chemical or biological components which can be used as markers for traditional cheese authentication (e.g., mass spectrometry analysis, such as stable isotope ratio and trace elements) that can all reflect cheese chemical composition. DNA-based methods for authentication are also emerging in the dairy sector, as shown recently by the work of Kamilari et al. [33], who used cheese microbiota metabarcoding for this purpose. However, for authentication of geographical origin, methods such as isotopic profiles [34] are generally preferred. As food labeling systems are constantly evolving, in parallel with legislation (e.g., Regulation [2]), analytical tests and technical control measures need to be improved and updated to counteract present and emerging fraud systems [35].

In this context, this review (based on 167 articles published over the last 27 years) aims to present methods, either physical and chemical or DNA-based, that are currently used for cheese geographical origin authentication, highlighting their principle, application, discriminative power, and advantages and limits, as well as to present future perspectives in this analytical field.

2. Chemical and Physical Methods for Cheese Origin Authentication

Polyphasic chemical and physical analysis approaches are now more common than single-parameter descriptions (i.e., dry matter, total protein, salt content) [36] to decipher cheese composition profiles. This is reinforced by the fact that cheese characterization, based on general chemical parameters, does not efficiently discriminate cheese geographical origin, as recently shown for water buffalo mozzarella by Salzano and colleagues [37]. That is why multiple signals are analyzed to acquire specific insight into typical cheese characteristics connected with its origin. Isotope and trace element fingerprinting methods have been considered as reference methods; however, other chemical and physical analyses can be used for geographical origin authentication.

2.1. Stable Isotope Ratio Mass Spectrometry

Stable isotope ratio mass spectrometry (IRMS) is among the most common methods for food geographical origin authentication [3]. It detects natural isotopic abundance of light and heavy stable isotopes, which mostly depend on climatic or geographical conditions (mainly latitude and altitude). The stable isotope ratio is also affected by biological and environmental interactions, and thus geographical product origins can be differentiated even if these have a high degree of similarity. Elements are called isotopes when their atoms are made by the same number of protons but a different number of neutrons, yielding a different atomic mass than the normal element [38]. Stable isotopes are non-radioactive isotopes and do not decay rapidly to form other elements.

Usually, stable isotope analysis is expressed as a ratio using an international standard to calculate it (Equation (1)):

$$\delta‰ = (R_{Sample} - R_{Standard})/R_{Standard} * 1000 \quad (1)$$

where R is the ratio between heavy and light isotopes [39]. The results of stable isotope ratios are always expressed as a percentage (‰—per mille unit) of international standard samples received from international organizations such as Vienna-Pee Dee Belemnite (V-PDB) for $\delta^{13}C$, Aria (AIR) for $\delta^{15}N$, Vienna—Standard Mean Ocean Water (V-SMOW) for $\delta^{2}H$ and $\delta^{18}O$, and Vienna—Canyon DiabloTriolite (V-CDT) for $\delta^{34}S$ [40].

IRMS has been widely used for cheese origin authentication [41–43] and many stable and unstable isotope ratios have been investigated. Among them, $^{13}C/^{12}C$, 15N/14N, $^{2}H/H$, $^{18}O/^{16}O$, and $^{34}S/^{32}S$ elements are the most commonly used, while $^{87}Sr/^{86}Sr$, $^{44}Ca/^{40}Ca$, $^{44}Ca/^{42}Ca$, and $^{206}Pb/^{204}Pb$ are occasionally reported [38,44]. For cheese authentication, IRMS is based on predictable and reproducible responses of stable isotopes to typical factors such as geographical origin, animal origin, seasonality, and manufacturing processes [38]. Animal feed was shown to have the highest impact on $\delta^{13}C$ and $\delta^{15}N$, while $\delta^{2}H$ is heavily influenced by the animal diet, and its combination with $\delta^{18}O$ is mainly impacted by geographical origin and seasonality [38,41]. $\delta^{34}S$ is mostly linked to geographical origin, i.e., soil geology, and it is not correlated with other stable isotope ratios [45]. On the contrary, further studies are still needed to gauge the effect of cheese-making on stable isotope ratios [38] as different results on these ratios in milk and cheese have been reported. They suggested either no major difference between milk and cheese obtained after milk processing [41,43] or a partial impact [34,42]. Considering isotope abundances in different organic macromolecules, the casein fraction has been reported to be the most reliable for origin authentication [39].

In general, possible effects of the cheese-making process on stable isotope ratios could be related to fat removal (in particular, the glycerol fraction), curd acidification, curd clotting (e.g., use of a commercial starter vs. natural milk cultures), curd washing, curd heat treatment (e.g., 50 °C), and salt washing/brining, while ripening time has not yet been reported to impact their composition [38,46]. For the casein stable isotope ratio of milk, the corresponding cheese did not show any significant differences for $\delta^{13}C$, $\delta^{15}N$, and $\delta^{2}H$, but an unexplained and significant difference ($p < 0.001$) was reported for $\delta^{18}O$ in typical pressed-cooked cheese [41]. The authors suggested a relevant fractionation in the animal in comparison with the feed, but a lack of isotopic fractionation during cheese-making [41]. However, Bontempo et al. [34] obtained a different isotopic ratio comparing milk and corresponding Mozzarella di Bufala Campana PDO for $\delta^{2}H$ and $\delta^{18}O$. Further studies on milk and corresponding cheeses obtained through different processes and technologies may provide additional insight into the eventual changes of stable isotope ratios during cheese-making. Table 1 reports the advantages and limitations of stable isotope ratio mass spectrometry for cheese origin authentication.

Table 1. Advantages and limitations of stable isotope ratio mass spectrometry for cheese origin authentication.

Advantages	Reference	Limitations	Reference
Isotope ratio such as ^{18}O, ^{2}H, and ^{34}S have a predictable and reproducible response toward geographical origin	[38]	Unclear effect of cheese-making on stable isotope ratio	[34,43]
Elevate correct classification rate	[42]	High operating cost	[47]
Consistent accuracy in origin authentication	[34,41,42]	^{13}C and ^{15}N are highly affected by animal feed	[41]

From an applied point of view, IRMS was used to discriminate the origin of two typical mountain cheeses from Italy [41]. The authors combined $\delta^{13}C$, $\delta^{15}N$, $\delta^{2}H$, and $\delta^{18}O$ from the casein fraction to build a canonical discriminant model that, after cross-validation, was able to correctly classify 96% of the milk and cheese samples. In Brazil, Silva and

colleagues [42] analyzed $\delta^{13}C$, $\delta^{15}N$, $\delta^{2}H$, and $\delta^{18}O$ from the water fraction of milk and cheese. After using the $\delta^{2}H$ and $\delta^{18}O$ in a linear regression model, they were able to discriminate milk and cheese samples according to their production area. In another study, Pillonel and colleagues [48] showed that the combined measurement of $\delta^{13}C$, $\delta^{15}N$, $\delta^{18}O$, and $\delta\ ^{34}S/^{32}S$ of the casein fraction authenticated Swiss vs. French Raclette cheeses.

In summary, isotope fingerprinting has proven to be reliable for the geographical origin authentication of typical cheeses. It is worth noting that stable isotope ratio analyses are already used as a traceability tool for some PDO cheeses such as Grana Padano and Parmigiano Reggiano (Regulation (EU) No. 1151/2012, amendment 2017/C 358/10 and 2018/C 132/07). This method is also often combined with trace element determination using inductively coupled plasma-mass spectrometry (ICP-MS) or inductively coupled plasma-atomic emission spectroscopy (ICP-AES).

2.2. Inductively Coupled Plasma

Cheese elemental fingerprinting is also currently used for origin authentication using inductively coupled plasma methods (ICP-MS or ICP-AES). These analyses rely on electromagnetic plasma to induce atom ionization (MS) or excitation (AES), to detect the different elements. Four main steps, including sample introduction and aerosol generation, plasma ionization/excitation, signal discrimination, and detection are used in ICP-MS and ICP-AES [44]. While MS analyzes ionized elements, AES detects light emitted by excited atoms. For cheese geographical origin authentication, ICP-MS is the preferred choice since it is a rapid, multi-element analysis able to quantify trace (ppm–ppb) or ultra-trace (ppb–ppq) elemental concentrations [49]. As elemental composition is mainly affected by geological and pedological traits, multiple factors can affect the element content in cheese, such as animal breed, feed vegetation, drinking water, and mineral supplementation [49,50]. Possible effects of cheese-making on cheese elements could derive from the clotting agent, curd acidification, manufacturing equipment, curd washing, and salt washing/brining [34,51]. Indeed, some authors reported that exclusion of Cu^{2+} and Zn^{2+} in multivariate analysis was necessary since a high transfer rate of these elements from dairy equipment to cheese was expected [52]. Table 2 reports the advantages and limitations of inductively coupled plasma for cheese origin authentication.

Table 2. Advantages and limitations of inductively coupled plasma-mass spectrometry for cheese origin authentication.

Advantages	Reference	Limitations	Reference
Elemental composition is mainly affected by geological and pedological traits	[49]	Animal feed and mineral supplementation affect elements' composition	[50]
Low operation costs with good analytical performance	[47]	Some elements such as Cu^{2+} and Zn^{2+} are highly affected by dairy equipment	[52]
Fast and multi-element analysis	[44]	Requires careful sample preparation	[39]

Elemental fingerprinting data obtained by Danezis and colleagues [50] have provided useful insights for an understanding of multiple element variations, including rare earth and precious metals, occurring in Greek Graviera cheeses obtained from 9 different regions (n = 105 samples). These authors analyzed 61 different elements, including rare earth (Dy, Er, Eu, Nd, Pr, Sc, Sm, Y, Yb), precious metals (Au, Pd, Re, Ru), ultra-trace elements (Nb, Ta, Tl, W, Zr), and trace elements (Ag, Al, Bi, Cd, Cu, Mo, Ni, Pb), and found significant differences according to cheese origin. They were able to correctly classify 92.1% of the tested cheeses (21 traditional and 9 commercial cheeses) using discriminant analysis [53]. The model was built on 65 elements, but the most significant variables were Ce, Dy, Eu, Gd, Ho, La, Nd, Pr, Sm, Tb, Y, Yb, Pd, As, Ba, Co, Fe, Ga, Mo, Ni, Ti, Zr, Ca, and P. Even if the authors did not perform any cross-validation, their results showed that ICP-MS was useful for cheese origin authentication, achieving high correct classification rates. According to

the same study, rare earth elements seemed to significantly increase the discriminatory power of ICP-MS.

As previously stated, a combination of IRMS and ICP-based methods can also be used. For example, IRMS (for the casein fractions) and ICP-MS achieved high performance for cheese origin classifications. This was shown for semi-cooked typical Italian alpine cheeses, namely Asiago, Fontina, Toma, Vezzena Montasio, Spressa, and Puzzone (n = 109). The use of Ba, Ca, Cu, Ga, K, Mg, Rb, Zn, $\delta^{13}C_{casein}$, $\delta^{15}N_{casein}$, $\delta^{13}C_{glycerol}$, and $\delta^{18}O_{glycerol}$ as predictive variables in canonical discriminative analysis was able to correctly classify 94% of the tested samples [52]. Another example was for authentication of PDO Parmigiano Reggiano vs. 11 imitation cheeses from different origins [54]. In their process, Camin et al. [54] used variables selected by the Random Forest algorithm ($\delta^{13}C_{casein}$, $\delta^{2}H_{casein}$, $\delta^{15}N_{casein}$, $\delta^{34}S_{casein}$, and Sr, Cu, Mo, Re, Na, U, Bi, Ni, Fe, Mn, Ga, Se, and Li) to create a supervised two-class model that was able to correctly classify 98.3% of the 265 hard cheese samples. Lastly, Nečemer and colleagues [51] discriminated different Slovenian cheese origins by combining P, S, K, Cl, Ca, Zn, and $\delta^{13}C_{casein}$ and $\delta^{15}N_{casein}$ contents, and a high correct classification rate (97%) confirmed that the dual IRMS/ICP-MS approach provided robust data to authenticate cheese geographical origins.

2.3. Infrared Spectroscopy

Spectroscopic analyses, including near-infrared (NIR) and mid-infrared (MIR), are based on the selective interaction of infrared beams with food molecules [55]. The infrared region includes a wide energy range (800 nm–1 mm, 12,500–10 cm^{-1}) and spectrophotometers can only evaluate a fraction of the wavelength, such as NIR (800–2500 nm, 12,500–4000 cm^{-1}), MIR (2.5–25 µm, 4000–400 cm^{-1}), and far-infrared (25–1000 µm, 400–10 cm^{-1}) [56]. Although infrared spectroscopy mainly involves vibrational energy, NIR incorporates both electronic and vibrational spectroscopy, while MIR mainly monitors molecular vibrations and far-infrared contains rotatory and vibrational movements. NIR spectroscopy mainly reflects the absorption information of overtone and combination tone of chemical bond vibrations of hydrogen-containing groups (C–H, O–H, N–H, and S–H) that reflect the anharmonic constant and the high-frequency vibration of the fundamental stretching of a XH bond (i.e., second overtone transition of C–H and O–H in the 1050–1400 nm region) [55,56]. Bands in the NIR region are weak or very weak, making this region markedly different from the others, but at the same time, more difficult to analyze. However, compared to the MIR region, OH and NH stretching bands of monomeric and polymeric species are better separated and differentiate free terminal functional groups from those within the molecule [56]. On the other hand, MIR spectroscopy is a highly sensitive method in which polar functional groups such as C=O, OH, and C–S exhibit intense bands. These bands, combined with other specificities of this region, such as stronger bands from antisymmetric vs. symmetric stretching, make this analysis useful for molecular fingerprinting [56].

In food fingerprints, spectroscopic techniques have gained popularity since these are fast, solvent-free, automatic, non-destructive, non-invasive, inexpensive, and can be used as a multiparameter analysis [49]. In general, NIR spectroscopy (in reflectance or transmittance mode) is more often used than MIR spectroscopy for food analysis as it requires less sample preparation and can be easily used for in-field analysis [57].

Cheese has proven to be a challenging matrix for infrared spectroscopic analysis as it is non-homogeneous (e.g., crystalline structure, holes) and numerous cheese types exist [58]. Nevertheless, many studies on cheese characterized and correctly predicted the chemical composition, manufacturing technique, ripening time, seasonality, and feeding system of milk-producing animals [59–62]. On the other hand, only a limited number of studies have focused on cheese origin authentication, possibly connected to the initial inability to differentiate milk geographical origin using infrared spectroscopic techniques [63] and sample size needed to validate the analysis [61]. Table 3 reports the advantages and limitations of infrared spectroscopy for cheese origin authentication.

Table 3. Advantages and limitations of infrared spectroscopy for cheese origin authentication.

Advantages	Reference	Limitations	Reference
May identify peculiar spectral differences connected with cheese origin	[57]	Requires large sample size for calibration	[64]
Fast, solvent-free, automatic, non-destructive, non-invasive, inexpensive, and can be used as a multiparameter analysis	[49]	Fingerprints may not detect less concentrated molecules connected with geographical origin	[57]
Minimal sample preparation	[58]	Lack of studies on the analysis and quantification the main sources of absorbance variation at each wavelength for cheese	[65]

According to Niermöller and Holroyd [64], studies conducted on a reduced sample size ($n < 60$) reported weak calibration for NIR spectra since, according to the multivariate model employed for the chemometrics analysis, sample size should be higher; in fact, PLS may require hundreds of samples per category. In another study, NIR spectroscopy correctly classified 96% of cheese samples from pasture-fed and conserved-forage-fed (hay or grass silage) dairy cattle [66]. More recently, NIR was used (850 to 1048 nm wavelength region) for a cheese model obtained from different dairy systems, and 67.1% of samples were correctly classified after applying a cross-validation (LDA model) [62]; however, the aim was not to discriminate the geographical origin of the tested samples. In a study comparing MIR and NIR spectroscopy performances, Karoui et al. [67] correctly classified 86.6% and 85.7% of 91 Emmental PDO cheeses obtained from Switzerland, France, Finland, Germany, and Austria using factorial discriminant analysis, thus highlighting MIR and NIR performances to discriminate cheeses produced with a similar process but with different geographic origins. In another study, Karoui et al. [68] used two MIR regions (3000–2800 and 1500–900 cm^{-1}) to successfully authenticate PDO Gruyère and L'Etivaz cheese, and 90.5% and 90.9% of samples were correctly classified using a factorial discriminant analysis.

In conclusion, although some studies have tested NIR and MIR spectroscopy to authenticate cheese origins, it is still difficult to conclude to what extent these techniques can be applied given the relatively low number of samples used in these studies [61,65]. In this context, further work on larger sample sizes for diverse cheese categories will likely improve the discriminatory power of these methods.

2.4. Nuclear Magnetic Resonance

Another spectroscopic analysis based on selective interactions between electromagnetic radiation and sample molecules is nuclear magnetic resonance (NMR). Generally, NMR uses radio frequency pulses to induce magnetic resonance nuclei-oriented transitions in an external magnetic field. When electromagnetic radiation hits nuclei, oriented nuclei move from lower to higher energy status or resonant nuclei. Subsequently, nuclei emit energy to return to the lower energy status, producing free induction decay [69]. The detected energy produces an absorbance signal, expressed in ppm, obtained from the ratio of standard molecules (e.g., 3-(trimethylsilyl)-propionate-d4). NMR, based on nuclei magnetic angular momentum—spin—is characterized by the azimuthal quantum number (I). Only nuclei with an even number of neutrons and an odd number of protons can be detected by their magnetic angular momentum (e.g., ^{12}C and ^{16}O have I = 0, while ^{1}H and ^{13}C have I > 1). Different NMR methods have already been employed in food authentication. However, considering cheese geographical origin authentication, ^{1}H high-resolution magic angle spinning (HRMAS) NMR is the most commonly used technique [70,71]. For example, Shintu and Caldarelli [71] applied a discriminant analysis using unsaturated fatty acid, aspartic acid, serine, and olefinic proton signals to classify the geographical origin of 20 Emmental cheeses from Austria, Finland, France, Germany, and Switzerland. They correctly classified 89.5% of samples after cross-validation. Mazzei and Piccolo [70] successfully identified Mozzarella di Bufala Campana PDO cheese origin. The studied cheeses were obtained in the same region but from two different provinces, namely Salerno and Caserta.

These authors applied a discriminant analysis model based on four metabolites linked to milk processing (β-galactose, β-lactose, acetic acid, and glycerol) and, after cross-validation, 100% of samples were correctly classified. Similarly, Consonni and Cagliani [72] correctly differentiated 93.5% of Italian Parmigiano Reggiano PDO cheeses from foreign eastern European "Grana type" cheeses by applying a PLS-DA model on leucine, isoleucine, lactate, butanoate, and acetate. The same authors also discriminated Parmigiano Reggiano PDO cheeses based on their ripening times. However, no cross-validation was performed to test the developed model.

The simultaneous analysis of proteins, lipids, and other metabolite fractions by ^{1}H HRMAS NMR offers great opportunities for cheese geographical origin authentications [73]. Moreover, NMR offers multiple advantages, such as simple sample preparation, multiple metabolite quantifications, high experimental reproducibility, and it is also non-destructive. Nevertheless, some limits should be considered, such as high cost for acquisition and maintenance, and higher limits of detection (typically, 10 to 100 times) when compared to gas chromatography-mass spectrometry [73,74]. Table 4 reports the advantages and limitations of nuclear magnetic resonance for cheese origin authentication.

Table 4. Advantages and limitations of nuclear magnetic resonance for cheese origin authentication.

Advantages	Reference	Limitations	Reference
Allows to obtain detailed information on cheese metabolites	[44]	Extremely high cost of acquisition	[73]
May identify biomarkers related to geographical origin	[73]	High detection limits	[74]
Minimal sample preparation	[49]	Complex calibration	[73]

2.5. Gas Chromatography-Based Fatty Acid Analysis

Recently, fatty acid analysis has been investigated for cheese origin authentication purposes. Fatty acids are lipid components formed by carboxylic acids with saturated or unsaturated aliphatic carbon chains [75]. Generally, fatty acid analysis is based on four sequential steps: extraction, derivatization, chromatographic separation, and detection [75]. The analytical reference method for fatty acid analysis is gas chromatography (GC), which is usually combined with flame ionization detectors [76]. For GC, analytes are vaporized in a heated chamber and transported by high-pressure inert gas (e.g., N^2, He) through the stationary phase (i.e., column material), where selective interaction leads to compound separation [77]. Subsequently, based on their retention index, compounds are eluted into the hydrogen flame of the detector, creating an electrical signal [78].

Fatty acid analysis to authenticate milk origin has already been reported [79]. These authors were able to efficiently authenticate the geographical origin of milk as the combination of different feeding strategies, herd and farm management practices (leading to distinct feed fatty acid profiles), grazing, breeding, animals' genetics, animals' rumen microbiota, and difference in lactation days varied considerably according to geographic locations. In traditional cheese, milk fatty acid profiles are also impacted by manufacturing practices (e.g., use of *Cynara cardunculus* L. as a coagulating agent), cheese microbiota (see the section on DNA-based methods for cheese origin authentication), and ripening times [80,81]. For example, fatty acid profiles were used and correctly authenticated the producer origins of Serra da Estrela PDO cheeses from Portugal, even within a limited production area [76]. These authors used 12 fatty acids, namely caproic, caprylic, undecanoic, lauric, pentadecanoic, palmitic, palmitoleic, heptadecanoic, oleic, linoleic trans-isomer, heneicosanoic, and arachidonic acids, in a linear discriminant model to achieve an 88% correct classification rate after cross-validation. Higher classification rates (95% after cross-validation) were also reported by Margalho et al. [82] using 13 fatty acids for 11 artisanal cheeses produced in 5 major geographical regions in Brazil (sample size n = 402). Similarly, high classification percentages were also achieved by Danezis et al. [83], who analyzed 101 PDO and 11 non-PDO cheeses from Greece. These authors used pH, moisture, fat, NaCl, and linoleic acid contents to correctly discriminate the PDO from non-PDO cheeses (excluding

similar hard PDO cheese) and achieved 100% correct classifications. Although high correct classification rates were obtained by Margalho et al. [82] and Danezis et al. [83] for cheese origin authentication, the analyzed cheeses greatly differed in terms of physical aspect (i.e., spread, soft, semi-hard, and hard), fat content, and ripening time. In recent studies dealing with origin authentication of similar traditional cheeses, the correct classification rates reported by Reis Lima et al. [76], Gatzias et al. [84], and Vatavali et al. [85] were, respectively, 88%, 88.2%, and 91.1%.

2.6. Gas Chromatography-Based Volatilome Analysis

Among the different strategies to authenticate cheese origin, volatile organic compound (VOC) analysis has gained attention as volatile compounds, which result from cheese microbiota metabolic activities, are an important component of cheese typicity. Indeed, through glycolysis, proteolysis, and lipolysis, cheese microbiota produce a wide range of VOCs. These include aldehydes, ketones, alcohols, esters, lactones, hydrocarbons, free fatty acids ($n < 10$ carbon atoms), sulfur compounds, and amines that provide the typical cheese aroma [86–88]. In this context, GC-MS is most commonly used to analyze cheese volatilomes (see Medina et al. [89] for technical details on cheese VOC analysis using GC-MS).

Using volatile fingerprinting, Pillonel et al. [90] discriminated country of origin for both PDO and non-PDO Emmental cheeses based on butan-2-one, 3-hydroxybutanone, butan-2-ol, and octene concentrations by principal component analysis (PCA). The concentrations of 21 other volatile compounds also showed at least one significant difference connected with their origin. Cheese origin discrimination relied on the fact that volatile profiles varied both qualitatively and quantitatively according to their country or region of origin [90]. For example, PDO Emmental from Switzerland was differentiated from Polish and French Emmental cheeses based on free-fatty acid qualitative composition (2-methyl butanoic acid for PDO Emmental cheese compared with 3-methylbutanoic acid in French and Polish Emmental ones) and relative abundance (such as nonanoic acid), alcohol presence/absence (3 methylbut-2-en-1-ol presence only in Swiss Emmental), as well as other aliphatic hydrocarbons, ketones, aldehydes, and esters. More recently, Pluta-Kubica et al. [91] also differentiated Emmental cheese origin based on their VOC profiles.

Similarly, Salzano et al. [37] used GC-MS to authenticate water buffalo mozzarella PDO cheese from the non-PDO versions. Both milk and cheese samples were analyzed, and differences were highlighted for both matrices using partial least squares discriminant analysis (PLS-DA). Variable importance in projection (VIP) analysis selected the 15 highest scored variables. Among them, talopyranose, 2,3-dihydroxypropyl icosanoate, sorbose, 4-phenyl glutamic acid, oxalic acid, and galactose were the most prevalent in typical PDO mozzarella, while tagatose, lactic acid dimer, ribitol, dodecyl thioglycolate, n-acetyl glucosamine, valine, and diethylene glycol were more abundant in non-PDO mozzarella. These authors concluded that the combination of multiple practices, such as forage from the same region, natural milk starters with both LAB and yeast instead of citric acid, and different packaging, all impacted the volatilome of the final product. These differences could thus explain what distinguished the water buffalo mozzarella produced according to PDO rules vs. those not following such guidelines (i.e., non-PDO mozzarella).

Another study authenticated Pecorino cheese origins, namely Pecorino Romano PDO, Pecorino Sardo PDO, and Pecorino di Farindola (certified by the Slow Food Foundation) [92]. The authors compared VOC fingerprints by a linear discriminant analysis (LDA) and PLS-DA model. The most influential variables in the LDA model were 2-methyl butyl isovalerate, butan-2-one, butyl butanoate, ethyl acetate, nonan-2-one, and propan-2-one, while in the PLS-DA model, VIP analysis identified 14 relevant compounds, namely, butan-2-one, pentan-2-ol, ethyl acetate, dicaprylyl ether, propanoic acid, 3-methylbutan-1-ol, propan-2-ol, ethyl decanoate, heptan-2-ol, butan-2-ol, butyl butanoate, pentan-2-one, ethanol, and 2-methylpropanoic acid. Only six compounds were common to the two tested models.

Considering that both models yielded similar classification performances (total classification rate of 87.5% after cross-validation), differences in the most influential variables linked to origin authentication were related to the index score value applied in the VIP analysis. Indeed, while the authors used an index score of 1, a previous study by Salzano et al. [37] used a higher value (i.e., 1.5) to select variables of interest and obtain a high correct classification rate for water buffalo mozzarella. Noteworthy, Vatavali et al. [85] only classified 47.5% of Graviera cheese origins, and thus the discriminative power of VOC profiles for cheese authentication could vary considerably according to the considered cheese type.

GC-MS fingerprints can also be exploited to identify biomarkers connected with specific attributes of traditional products, such as animal feed requirements. Indeed, Caligiani et al. [93] validated a method to quantify cyclopropane fatty acids (e.g., dihydrosterculic acid) as biomarkers for cows fed with corn silage. PDO cheeses such as Parmigiano Reggiano, Fontina, Comté, and Gruyère do not permit silage to be used in cow feed, and thus the absence of cyclopropane fatty acids in such cheeses may confirm correct feeding management.

Another feed-associated fraction of cheese VOCs are terpenoids. Terpenoids are a highly diversified class of naturally occurring organic compounds or phytochemicals, also called isoprenoids. These are derived from isoprene units and produced by dicotyledon plants [88]. In Slovenian cheeses, VOC analysis discriminated cheeses into 4 clusters (average silhouette 0.764) according to their geographical origin and based on 9 monoterpenes, namely, α-pinene, camphene, α-phellandrene, β-pinene, 3-carene, 2-carene, limonene, tricyclene, and γ-terpinene [94]. In a similar way, Turri et al. [88] identified significant differences in 10 terpenes between pasture-producer of Historic Rebel cheese, and the results suggested that allo-ocimene, α-terpinolene, α-pinene, and δ-3-carene could be possible biomarkers to differentiate cheese origin. Overall, these studies highlighted that volatilome analysis can be an interesting tool for cheese origin authentication, although classification rate performances can vary among cheese varieties.

Another approach to directly analyze volatile compounds is the electronic nose (e-nose). This analytical technology, designed to mimic the human olfactory system [95], has gained interest in food authentication as it is highly correlated with consumer perception [96]. A typical e-nose comprises the sampling system, a set of non-selective sensors or mass spectrometer (MS), and a pattern-recognition system [97,98]. Nowadays, different sensors are used, such as metal-oxide semiconductors, conducting polymers, and piezoelectric crystal sensors [99]. In the case of cheese origin authentication, only a limited number of studies reported an e-nose strategy. For example, Pillonel et al. [90] obtained similar classification rates using a PCA model to evaluate Emmental cheese origins (90%). In the case of Pecorino cheeses, an e-nose and artificial neural network approach correctly classified 96.5% of Pecorino di Fossa PDO cheeses (n = 18) and Pecorino cheeses of other origins (n = 48) [100]. In conclusion, while different authors have reported e-noses to authenticate foods subjected to different frauds, including geographical origin [98,101,102], only a limited number of studies concerned cheese origin authentication. Table 5 reports the advantages and limitations of VOC analysis for cheese origin authentication.

Table 5. Advantages and limitations of volatilome analysis for cheese origin authentication.

Advantages	Reference	Limitations	Reference
High resolution, short separation time, high sensitivity, and low cost	[32]	High dependency on ripening time	[103,104]
Measures correlated with cheese quality	[87]	Highly impacted by extraction methods	[89]
Fingerprint profile with discrete correct classification rate	[92,105]	High variability of correct classification rate	[37,85,92]

Some important attributes that may impact the efficiency of this approach to discriminate cheese origin are provided. One important attribute to differentiate cheese is ripening time [106]. Typical cheeses are sold according to a minimum ripening time. However,

according to the initial extent of ripening, significant changes may occur during shelf-life. Contamination during portioning, inadequate temperature usage during transport, light exposure, and storage conditions may impact VOC profiles. In this context, specific VOC fractions, such as terpenoids, combined with VOC and microbiota correlations, may overcome possible changes in VOC profiles at the retailing stage. However, to our best knowledge, no studies have evaluated the effect of retail on cheese origin authentication.

3. DNA-Based Methods for Cheese Origin Authentication

As previously mentioned, the metabolic activities of cheese microbiota play a crucial role in the development of cheese typicity. For geographical origin authentication, microbiota fingerprinting is therefore of high interest as traditional and artisanal cheeses are produced with a more diversified microbiota associated with the cheese-making process (e.g., use of raw milk, starter, brine, equipment and materials, and ripening rooms). Distinct differences in the composition of this complex microbiota, composed of Gram-positive and -negative bacteria, fungi, archaea, and viruses, could be used for cheese origin authentication.

The first study on cheese microbiota using high-throughput sequencing (HTS) was performed by Quigley et al. [107] on traditional cheeses. Since then, numerous studies have been published on many aspects linked to cheese quality and typicity. While cheese microbial diversity was traditionally investigated using culture-dependent methods, hence overlooking unculturable or subdominant species, nowadays, culture-independent methods (HTS) have unraveled this diversity and provided further means to connect microbiota composition to cheese quality and typicity, but also origin. This success is due to the availability of new sequencing platforms, bioinformatic pipelines, and a continuous decrease in cost. Among high-throughput sequencing, amplicon sequencing—or selective amplification of polymorphic genes across their hypervariable regions—is the most widely reported in the scientific literature [108]. In this context, the use of DNA metabarcoding (also known as metagenetics) to study cheese microbiota was proposed as a tool for cheese origin authentication [33].

To perform DNA metabarcoding, cheese samples are first homogenized, then total DNA is most frequently extracted using commercial kits, ad hoc protocols, or a combination thereof [109–111]. Hypervariable regions of taxonomically relevant genes (e.g., 16S rDNA for bacteria and archaea, ITS, 18S rDNA, 26S rDNA for fungi) are amplified by PCR reactions, while a second amplification step tags amplicons with specific DNA fragments—barcodes—and dedicated adapters for the final sequencing step using next-generation technologies (e.g., Illumina, Pacbio, Iontorrent, or Nanopore). For further details on sampling, library preparation, and sequencing platforms, the reviews by Hugerth et al. [112] and Tilocca et al. [113] are suggested.

Typically, 16S rDNA and ITS (internal transcribed spacer) markers, targeting bacteria and fungi, respectively, are employed to generate compositional data describing microbial taxa and their relative abundance in cheese microbial communities. After sequencing, two complementary but different ways can be used for amplicon clustering from quality-checked data, namely, operational taxonomic units (OTUs) and amplicon sequencing variants (ASVs) [114]. On one hand, pipelines such as QIIME and IMNGS build sequence clusters based on their similarity (usually using a similarity cutoff of 97%) to obtain OTUs [115,116]. On the other hand, ASVs are obtained with the DADA2 pipeline (also available in QIIME2) by inferring biological sequences in a sample, and discerning sequence variant differences down to a single nucleotide [117,118]. Subsequently, taxonomic assignment is performed using a specific classifier tool (BLAST, RDP, UCLUST, SortMeRNA) against various reference databases, such as Greengenes, SILVA, and UNITE [119,120]. Generally, clustered OTUs/ASVs are analyzed from the phylum to the genus level since they can be less precise at the species level [121]. Identified taxa can be divided into dominant, subdominant, and rare sequences, representing 40% to 90%, 1% to 0.01%, or 0.01% to 0.0001% of reads per sample, respectively [122].

Using these DNA methods to determine cheese microbiota composition, literature data have shown that the core is usually dominated by lactic acid bacteria (LAB) [123], including starter and non-starter lactic acid bacteria (NSLAB), while in the case of cheese with edible rinds, rinds are usually dominated by salt-tolerant fungi and bacteria from the *Actinomycetota*, *Bacillota,* and *Pseudomonadota* phyla. Cheese microorganisms can be deliberately inoculated into the milk or on the cheese surface as starter or secondary cultures, but may also originate from multiple reservoirs, including raw milk (i.e., for raw milk cheeses), brine or salt, or dairy environments (cheese-making equipment and ripening shelves). A distinctive case corresponds to raw milk cheeses as cheese microbiota can be differentiated according to the amount of starter cultures employed during cheese-making and the origin of the raw milk used, at the farm level [124]. Since milk quality depends on many factors (e.g., animal health status, breed, lactation stage, teat skin, hides, feces management, farm dimension, feeding system, season, farm staff hygiene, and management), recent longitudinal studies have connected some of the main characteristics with farm origin [125–127]. Indeed, for raw milk cheeses, cheese microbiota can be directly impacted by raw milk microbiota [107], thus, both microbiota can be used for cheese origin authentication. Generally, most protected land- and tradition-related labeled cheeses are produced from raw milk, and the complex microbiota encountered in raw milk directly influences the unique cheese sensorial properties appreciated by consumers. As an example, in the EU, among the 284 cheeses recognized for their typicity, over 180 are raw milk cheeses ([128] https://ec.europa.eu/info/food-farming-fisheries/food-safety-and-quality/certification/quality-labels/geographical-indications-register/ accessed on 3 October 2022). Besides raw milk, traditional tools/equipment and the dairy environment are also shaping cheese microbiota at both species and strain levels. In fact, some key species may originate from the dairy environment [129]. In addition, as reported by Bokulich and Mills [130] and Calasso et al. [131] for LAB, different strains of a given species can colonize the dairy environment, and thus, the cheese. More recently, Sun et al. [132] determined that in-house microbiota were essential in shaping Bethlehem (PA, USA), a Saint-Nectaire-type cheese produced without starters by traditional methods. Using 16S rRNA amplicon sequencing and SourceTracker (a bioinformatic tool based on Bayesian inference that estimates the proportion of different sources contributing to a designated microbial community) [133], these authors identified wooden vats—used for overnight ripening—as a major source of desirable LAB that shaped cheese microbiota from acidification to ripening. Similar findings were reported by Montel et al. [134], who highlighted that traditional cheeses have complex and rich microbiota influenced by traditional equipment. Indeed, traditional cheeses are produced using cheese-making practices that tend to increase microbial diversity via contact with diverse microorganisms originating from dairy equipment [135]. Moreover, differences in cheese-making technologies (e.g., use of natural milk or whey culture rather than commercial starters, use of rennet or clotting agents, curd cooking, draining process, salting) and farm practices (e.g., type of housing, silage, grassland) between different production areas can also shape the cheese microbial community [136]. Considering these factors, we can question whether traditional cheeses can be differentiated based on their origin and if the main factors that affect cheese microbiota complexity and diversity play a significant role.

3.1. Main Factors Affecting Cheese Microbial Diversity

In this context, in a recent study by Kamimura et al. [137], the microbiota of 578 traditional Brazilian cheeses were analyzed with amplicon sequencing. Bacterial communities were distinctly clustered with PCA by cheese type and regional origins while, at the genus level, hierarchical cluster analysis separated production regions. These authors were thus able to identify specific origin-related microbiota. The core microbiota of Brazilian traditional cheeses displayed different relative abundances and oligotypes (i.e., closely related but distinct bacterial taxa) of LAB belonging to the *Enterococcus*, *Lactococcus*, *Streptococcus*, *Leuconostoc*, and *Lactobacillus sensu lato* genera, as well as other taxa belonging to the *Enter-*

obacteriaceae family and *Staphylococcus* genus. Within the same regional area, microbiota analysis differentiated the origin—Cerrado, Araxà, Canastra, Campos des Vertenes, and Serro—of traditional cheeses produced with a similar natural whey starter and ripening period (17 and 22 days). These findings were in agreement with those of another study that analyzed 97 samples of Minas artisanal cheese from 6 different producers located in the same region [138]. Starter cultures, consisting of *Streptococcus*, *Lactococcus*, and *Lactobacillus sensu lato* spp., constituted the core microbiota in all farms. However, significant differences in family- and genus-level bacterial community relative abundances were observed between the studied farms due to environmental factors such as geographical location. Even when dominant genera may be inferred to the natural whey cultures used, meta-analysis from amplicon sequencing data of traditional artisanal cheeses from Italy, Belgium, and Kalmykia indicated differences in bacterial structures between cheeses produced in different geographic areas (unweighted PCoA cluster MANOVA, $p < 0.001$ [139]). Those produced using natural milk cultures showed improved acidification without an effect on the typical cheese microbiota. Indeed, at the end of the ripening period, cheese origins clustered according to producer facilities (PCoA on Bray–Curtis) [140].

Another study was performed by Zago et al. [136]. In this case, 118 Grana Padano samples were analyzed after 7–8 months of ripening and a common core microbiota composed of *Lactobacillus*-, *Lactococcus*-, *Lacticaseibacillus*-, *Limosilactobacillus*-, and *Streptococcus*-dominant genera was observed. More precisely, differences in bacterial abundance, richness, and evenness were found for dominant and sub-dominant groups according to production region, a result also confirmed by PERMANOVA beta-diversity analysis. The authors also identified specific species that could be linked to several production areas; however, no species biomarkers were identified regardless of production area and non-metric multidimensional scaling did not show any clear clustering profile.

Some cheeses are produced in very small geographical zones by a limited number of producers. This is the case for Plaisentif and Historic Rebel cheeses from the mountainous regions in Italy, that are only produced during specific seasons (violet blooming season and grazing season) by 14 and 12 producers, respectively [88,122]. Both are raw milk cheeses produced without starter adjunction. Bacterial amplicon sequencing analysis (16S rDNA V4 region) for Plaisentif cheese identified dominant genera and, more importantly, differences in bacterial community profiles between producers thus detected fraudulent starter additions in some cheeses [122].

Based on a similar analysis for Historical Rebel cheese, the core microbiota was composed of 5 different genera—*Streptococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc,* and *Pediococcus*—with *Streptococcus* relative abundances ranging from 60% to 85% [88]. Richness and other alpha-diversity parameters differed among producers as well as in multivariate analysis (PCoA on unweighted Unifrac), and based on the observed significant differences, pasture area could be linked to the different Historic Rebel cheese producers.

3.2. Climatic and Environmental Condition

Another factor that can impact microbial communities of traditional cheeses are the climatic and dairy environment conditions that are directly associated with geographical origins. This was observed for traditional Chinese Rushan cheese produced using *Chaenomeles sinensis* boiled extract as a clotting agent in three different regions. Even if the same UHT milk and production equipment were used, geographical origins significantly impacted the relative abundance of 12 dominant genera, namely *Lactobacillus*, *Acinetobacter*, *Acetobacter*, *Lactococcus*, *Enterobacter*, *Moraxella*, *Enterococcus*, *Streptococcus*, *Kocuria*, *Staphylococcus*, *Chryseobacterium*, and *Exiguobacterium* [141], and is likely related to specific house microbiota and open-air drying. This result was also confirmed by PCoA clusters and Anosim analysis.

As typical cheese microbiota can be in part acquired from the specific raw materials used, traditional tools, environmental and production conditions, cheese-making process, and geographical area, a comparison between traditional and industrial cheeses may pro-

vide additional information to authenticate cheese origin. Noteworthy, some authors have reported that commercial starters, inoculated at ~10^6 CFU/mL, prevent resident microbiota from developing, especially during ripening [124,142]. Overall, milk pasteurization, use of similar commercial starters, similar industrial equipment, and standardized recipes for cheese production are crucial factors that decrease cheese microbial complexity and biodiversity and lead to highly standardized productions. These directly deplete the unicity of the matrix, and thus lower variance is detected (Figure 1). This hypothesis is in accordance with the study by Kamilari et al. [143], in which a significant decrease in bacterial diversity was observed in industrially produced Haloumi cheeses vs. artisanal products. However, the microbial diversity observed for artisanal Haloumi cheeses could not link them to their producers' geographical origins. In another study, aiming to authenticate cheese origin at the producer level, no distinction could again be made [144].

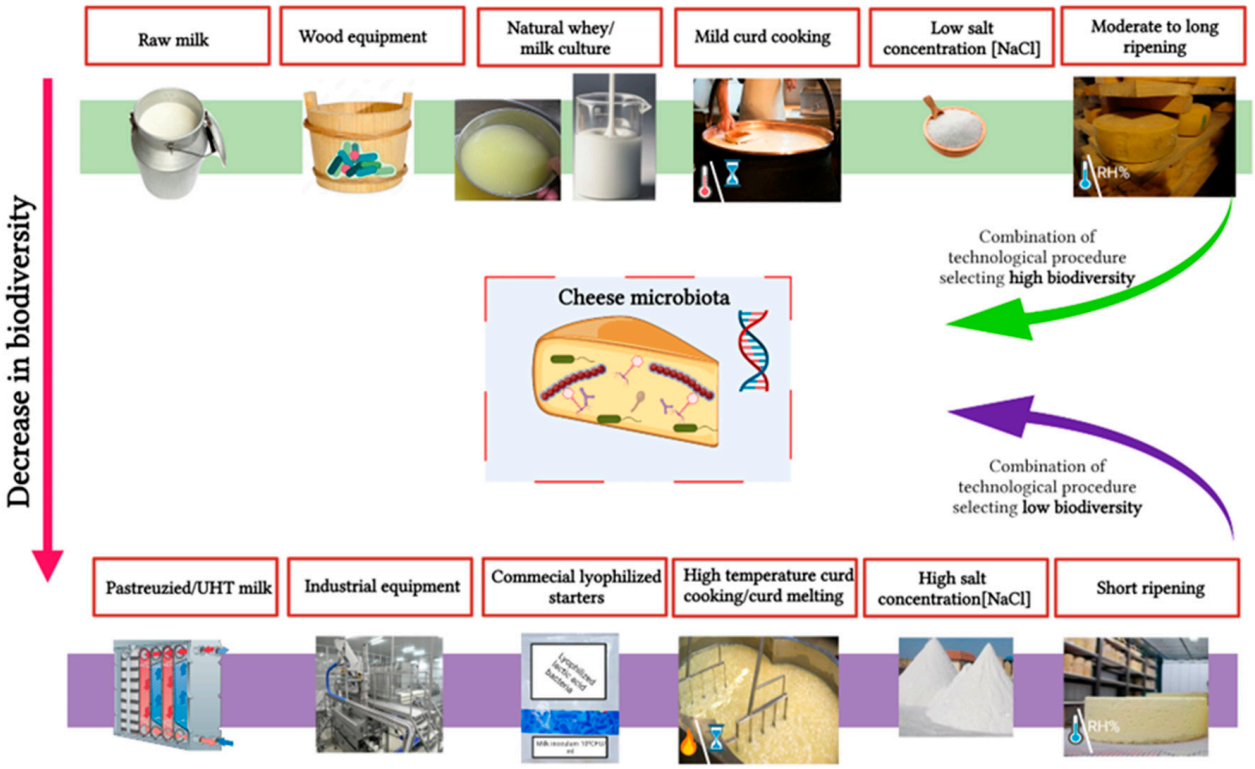


Figure 1. Technological factors affecting cheese microbiota biodiversity. Green and purple lines show combinations of technological factors during cheese-making that increase or decrease this biodiversity.

3.3. Cheese Ripening

Cheese ripening is another factor that affects microbial community diversity and cheese typicity. Ripening can be considered as a selection process that leads to cheese microbial composition changes. According to Gobetti et al. [123], intentionally added microorganisms used in cheese-making include primary starters (natural milk culture, natural whey culture, or lyophilized commercial starters), secondary or adjunct LAB starters, and milk autochthonous microbiota (NSLAB and others). These are the main ripening agents in intermediate to long ripening times, which mainly explain the observed diversity and typicity of the produced cheeses. The relationship between primary starters and NSLAB during maturation is well-known and involves a progressive reduction of the former in favor of the latter. The role of NSLAB is crucial in maturation and the development of the typical characteristics of traditional cheese. In this case, useful insights can be gained by comparing the genomic features of primary starters for the presence of genes involved in the metabolic pathways important for cheese maturation. While primary starters have important genetic features for the utilization of lactose—mainly connected with their acidification ability—

NSLAB possess many more genes coding for peptidases, peptide transporters within cells, and amino acid catabolism, that can represent an advantage during cheese maturation [145]. Moreover, compared to primary starters, NSLAB tend to adapt better to the hostile conditions of the cheese ripening environment, such as temperature, salt content, pH, and redox potential. In fact, NSLAB can adopt alternative metabolic pathways to produce energy from unconventional sources while resisting acid conditions. Therefore, NSLAB present in raw milk at a sufficient inoculum to colonize ripened raw milk cheeses, or acquired from the house microbiota, could be an indicator of geographical origin at the producer level. Beyond NSLAB, other microorganisms belonging to various groups can influence cheese ripening. This is the case of fungal communities in many traditional cheeses, such as Queijo de Azeitão in Portugal [146], Tomme d'Orchies in France [110], and Robiola di Roccaverano in Italy [147]. Indeed, fungal communities are well-known for their decisive role in flavor and texture of white and blue-veined mold-ripened cheeses due to lipolytic, proteolytic, and glycolytic activities, leading to high production of aromatic ketones and alcohols [134,148,149]. Generally, fungal species such as *Penicillium camemberti*, *Penicillium roqueforti*, *Debaryomyces hansenii*, *Kluyveromyces marxianus*, *Candida catenulata*, *Galactomyces geotrichum*, and *Mucor lanceolatus* are either deliberately added as technological adjunct cultures or present in the production environment [149–152]. Nevertheless, in traditional cheeses, fungal communities were reported to be more diverse than the used starter but, at the same time, not connected with geographical origin [109,146]. Table 6 reports advantages and limitations of amplicon sequencing for cheese origin authentication.

The mentioned studies showed that in traditional cheeses, the combination of artisanal cheese-making, specific raw materials, and characteristic environmental conditions shape microbial community diversity according to geographical origin. Most analyses conducted using 16S rDNA amplicon sequencing discriminated cheese origin, although taxonomic classification was still limited to genus/family-level descriptions and only a few cheese types per study were considered. To further assess the unicity of typical cheeses against food fraud, more in-depth studies, including meta-analyses on all available cheese data and increased depth of microbial population descriptions (e.g., metagenomics), are of interest.

Table 6. Advantages and limitations of HTS amplicon sequencing for cheese origin authentication.

Advantages	Reference	Limitations	Reference
Time- and cost-effective processing of large sample numbers	[153]	Analyses could be biased by sample processing, DNA extraction methods, and equimolar library preparation	[154]
Consolidated pipeline for data analysis	[155]	PCR amplification steps include errors, e.g., PCR specificity and variation of 16S rRNA copy number per genome	[108]
Identification of taxonomic groups associated with typical flavor and cheese-making technology	[156]	Under- or over-estimation of microbial community diversity	[112,155]
Allows improvement of cheese-making to ensure safety while preserving typicity	[137]	Lack of absolute abundance	[157,158]
Evaluation of core microbiome describing facility-associated microbial groups	[130,131]	Limited and uneven taxonomic resolutions	[159,160]
May pinpoint new biotypes	[143]	DNA amplicon sequencing typically does not discriminate between live and dead microorganisms (except if DNA stains such as propidium monoazide are used)	[87]

4. Conclusions and Perspectives

Geographical origin authentication is an important safeguard for food quality and safety but also from an economical point of view as it enables consumer protection and provides technical support to enforce national and international legislations [49]. Applying elemental and isotopic characterization, volatilome or microbiota analysis to typical products can protect them from food fraud and improve registration processes and marketing decision-making [161]. Nevertheless, our knowledge on authentication methods is far from complete. Indeed, most studies and methodologies employed to authenticate cheese origins only provide qualitative answers (e.g., does the method discriminate origin?) and often lack quantitative assessment (to what extent can different cheese origins be discriminated from

each other?). This is probably connected to the complexity of commonly used multivariate models such as PLS, PLS-DA, and LDA, and the use of specific algorithms (e.g., Random Forest, VIP) for variable selection and the need for internal or interlaboratory external and cross-validations.

Chemometrics approaches differ among analytical technologies. In this case, infrared methods—especially NIR spectroscopy—have consistently been applied together with chemometrics analysis to achieve good classification rates after cross-validation, but the number of samples employed for authentication purposes was often limited. Stable isotope ratios combined with trace element analysis have been shown to be the most accurate methods to authenticate cheese origin, since available studies reported consistent statistical analysis and modeling with high correct classification rates after cross-validation. Nevertheless, the actual discriminative power of the method for closely distant cheese producers remains unknown. Moreover, the impact of animal feeding and the high cost per sample must be considered. Considering that many production disciplinaries declare a minimum amount of local forage in animal feed, one possible strategy to increase cheese typicity and improve authentication using stable isotopic and elemental metabolomics would be to increase the use of local feed or use grazing.

Regarding DNA-based methods, amplicon sequencing can discriminate cheese geographical origin. However, some important considerations to assess traditional cheese origin are related to metabarcoding approach limitations. Indeed, microbial species and strains that originate from the dairy environment and that characterize traditional product origins could be key biomarkers for cheese authentication. In this context, shotgun metagenomics offers multiple advantages in comparison with amplicon metabarcoding, since no amplification step is required and all the genetic material in the sample is used for genome reconstruction to reach a deeper taxonomic assignation, potentially to the strain level [162,163]. For example, StrainPhlAn uses unique gene family markers and sample-specific consensus sequences to infer strain-level genotypes from different environments [164]. Considering that cheese is formed by microorganisms of different domains, appropriate sequencing depths would directly provide insight on bacteria, fungi, and viruses in the same analysis. These tools could thus be applied to milk, starters, and typical cheeses to obtain accurate DNA-based fingerprints to efficiently authenticate product origin [164]. This approach may help explore new traceability systems based on crucial components of fermented products, such as their virome [165]. Moreover, shotgun metagenomic analyses of gene richness, as a possible indicator of microbial community adaptation to different stress conditions, would be useful to reconstruct metabolic pathways connected with specific cheese traits, such as volatile compound production and metabolites that characterize typical cheeses. In this sense, integrated systems biology, combining metabolomics and metagenomics, could improve our knowledge on this subject, as only a few studies to date have combined both techniques [88,156,166]. While microbiota-based studies have compared typical and non-typical products, some omics approaches reported differences between typical and industrial products, thus making it difficult to clearly determine the factors that characterize typical food products. Artificial intelligence approaches, such as deep learning and machine learning, should be taken into consideration to improve classification rates and better-differentiate authentic and fraudulent products [167]. Overall, further research focused on comparing how well DNA-based analyses perform in comparison to the actual reference analyses (i.e., isotope fingerprinting and trace element analysis) used to authenticate cheese origin is needed. A combined approach, using isotope fingerprinting or trace element analysis and metagenomics, to obtain the highest discriminative power for cheese geographical origin authentication could also be of interest.

Author Contributions: M.C. (Marco Cardin), E.C. and J.M. conceptualized the layout of the review; M.C. (Marco Cardin) wrote the original draft, created the associated Tables and Figures, and edited the manuscript; B.C. contributed to the manuscript on cheese fraud and methodological aspects; E.N. contributed to physical and chemical aspects, and M.C. (Monika Coton) to volatilome aspects; J.M. contributed to DNA-based methods, and E.C. supervised the work and contributed to all aspects,

including writing—review and editing; B.C., E.N. and E.C. obtained the funding. All authors have read and agreed to the published version of the manuscript.

Funding: This work was completed in the framework of Marco Cardin's Ph.D. grant funded by the Università degli Studi di Padova (Italy) and the Université de Bretagne Occidentale (France). Funding number 20192413451/ISCR_D1.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Robson, K.; Dean, M.; Haughey, S.; Elliott, C. A Comprehensive Review of Food Fraud Terminologies and Food Fraud Mitigation Guides. *Food Control* **2021**, *120*, 107516. [CrossRef]
2. Regulation (EU) No. 625/2017 of the European Parliament and of the Council of 15 March 2017 on Official Controls and Other Official Activities Performed to Ensure the Application of Food and Feed Law, Rules on Animal Health and Welfare, Plant Health and Plant Protection Products, 95, 1–95. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0625&from=EN (accessed on 3 October 2022).
3. Eurofins Analytics France. *FoodIntegrity Handbook: A Guide to Food Authenticity Issues and Analytical Solutions*, 1st ed.; Morin, J.-F., Lees, M., Eds.; Eurofins Analytics France: Nantes, France, 2018. [CrossRef]
4. Spink, J.W. The Current State of Food Fraud Prevention: Overview and Requirements to Address 'How to Start?' And 'How Much Is Enough?'. *Curr. Opin. Food Sci.* **2019**, *27*, 130–138. [CrossRef]
5. Rocchi, B.; Romano, D.; Sadiddin, A.; Stefani, G. Assessing the economy-wide impact of food fraud: A SAM-based counterfactual approach. *Agribusiness* **2020**, *36*, 167–191. [CrossRef]
6. Moore, J.C.; Spink, J.; Lipp, M. Development and Application of a Database of Food Ingredient Fraud and Economically Motivated Adulteration from 1980 to 2010. *J. Food Sci.* **2012**, *77*, R118–R126. [CrossRef]
7. Everstine, K.; Spink, J.; Kennedy, S. Economically Motivated Adulteration (EMA) of Food: Common Characteristics of EMA Incidents. *J. Food Prot.* **2013**, *76*, 723–735. [CrossRef] [PubMed]
8. Hong, E.; Lee, S.Y.; Jeong, J.Y.; Park, J.M.; Kim, B.H.; Kwon, K.; Chun, H.S. Modern Analytical Methods for the Detection of Food Fraud and Adulteration by Food Category: Adulterated Food Categories and Their Analytical Methods. *J. Sci. Food Agric.* **2017**, *97*, 3877–3896. [CrossRef] [PubMed]
9. Breitenbach, R.; Rodrigues, H.; Brandão, J.B. Whose Fault Is It? Fraud Scandal in the Milk Industry and Its Impact on Product Image and Consumption—The Case of Brazil. *Food Res. Int.* **2018**, *108*, 475–481. [CrossRef]
10. Europol-Interpol Operation Results in Global Seizures of Fake and Illicit Food—OPSON VII [Media Release] 2019. Available online: https://www.europol.europa.eu/publications-events/publications/operation-opson-vii-analysis-report (accessed on 2 October 2022).
11. Europol-Interpol Operation Results in Global Seizures of Fake and Illicit Food—OPSON IX [Media Release] 2021. Available online: https://www.europol.europa.eu/cms/sites/default/files/documents/opson_ix_report_2021_0.pdf (accessed on 2 October 2022).
12. Gimonkar, S.; E. Van Fleet, E.; Boys, K.A. Dairy Product Fraud. In *Food Fraud*; Elsevier: Rockville, MD, USA, 2021; pp. 249–279. [CrossRef]
13. Montgomery, H.; Haughey, S.A.; Elliott, C.T. Recent Food Safety and Fraud Issues within the Dairy Supply Chain (2015–2019). *Glob. Food Secur.* **2020**, *26*, 100447. [CrossRef]
14. European Commission. The EU Food Fraud Network and the System for Administrative Assistance—Food Fraud. 2018. Available online: https://ec.europa.eu/food/system/files/2020-05/ff_ffn_annual-report_2018.pdf (accessed on 2 October 2022).
15. BRCGS, British Retail Consortium Global Standards Global Food Safety Standard (Issue 8) British Retail Consortium Global Standards. 2018. Available online: https://www.brcgs.com/search-results/global-standard-for-food-safety-issue-8/p-616/ (accessed on 2 October 2022).
16. FSSC 22000, Foundation Food Safety System Certification 22000 FSSC 22000 Scheme Version 5. Available online: https://www.fssc22000.com/ (accessed on 2 October 2022).
17. GFSI, Global Food Safety Initiative Tackling Food Fraud through Food Safety Management Systems. GFSI 2018. Available online: https://mygfsi.com/wp-content/uploads/2019/09/Food-Fraud-GFSI-Technical-Document.pdf (accessed on 2 October 2022).
18. European Committee for Standardization with the CEN Workshop Agreement (CEN WS/86) (Work Item Number WS086002)—Authenticity and Fraud in the Feed and Food Chain—Concepts, Terms, and Definitions (Reference CWA 17369:2019). Available online: https://www.nbn.be/shop/nl/zoeken/?k=%20CWA%2017369:2019 (accessed on 2 October 2022).
19. Codex Alimentarius International Food Standard. General Standard for Cheese. 1978. Available online: https://www.fao.org/fao-who-codexalimentarius/sh-proxy/es/?lnk=1&url=https%253A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252FStandards%252FCXS%2B283-1978%252FCXS_283e.pdf (accessed on 2 October 2022).
20. Valletta, M.; Ragucci, S.; Landi, N.; Di Maro, A.; Pedone, P.V.; Russo, R.; Chambery, A. Mass Spectrometry-Based Protein and Peptide Profiling for Food Frauds, Traceability and Authenticity Assessment. *Food Chem.* **2021**, *365*, 130456. [CrossRef]

21. Vandecandelaere, E.; Arfini, F.; Belletti, G.; Marescotti, A.; Allaire, G.; Cadilhon, J.; Casabianca, F.; Damary, P.H.G.; Estève, M.; Hilmi, M.; et al. Uniendo Personas, Territorios y Productos. Guía para Fomentar la Calidad Vinculada al Origen y las Indicaciones Geográficas Sostenibles. FAO 2010. Available online: https://agritrop.cirad.fr/596230/1/ID596230.pdf (accessed on 2 October 2022).
22. Picon, A. Cheese Microbial Ecology and Safety. In *Global Cheesemaking Technology*; Papademas, P., Bintsis, T., Eds.; John Wiley & Sons, Ltd: Chichester, UK, 2017; pp. 71–99. [CrossRef]
23. *Regulation (EU) No. 1151/2012 of the European Parliament and of the Council of 21 November 2012 on Quality Schemes for Agricultural Products and Foodstuffs*; Official Journal of the European Union: Brussels, European Union; Volume 343, pp. 1–29. Available online: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:343:0001:0029:en:PDF (accessed on 3 October 2022).
24. *Regulation (EU) No. 665/2014 of the European Parliament and of the Council of 11 March 2014 on Supplementing Regulation (EU) No 1151/2012 of the European Parliament and of the Council with Regard to Conditions of Use of the Optional Quality Term 'Mountain Product'*; Official Journal of the European Union: Brussels, European Union; Volume 179, pp. 23–25. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0665&rid=6 (accessed on 3 October 2022).
25. European Commission. *Joint Research Centre. Institute for Prospective Technological Studies. EU Island Farming and the Labelling of Its Products*; Publications Office: Luxembourg, European Union, 2013.
26. Sidali, K.L.; Capitello, R.; Manurung, A.J.T. Development and Validation of the Perceived Authenticity Scale for Cheese Specialties with Protected Designation of Origin. *Foods* **2021**, *10*, 248. [CrossRef] [PubMed]
27. *Regulation (EU) No. 178/2002 of the European Parliament and of the Council of 28 January 2002 Laying down the General Principles and Requirements of Food Law, Establishing the European Food Safety Authority and Laying down Procedures in Matters of Food Safety*; Official Journal of the European Communities: Brussels, European Union, 1 February 2002. Available online: https://www.fsai.ie/uploadedFiles/Legislation/Food_Legisation_Links/General_Principles_of_Food_Law/Consol_Reg178_2002.pdf (accessed on 3 October 2022).
28. *Regulation (EU) No. 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the Provision of Food Information to Consumers, Amending Regulations (EC) No 1924/2006 and (EC) No 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation (EC) No 608/2004 Text with EEA relevance*; Official Journal of the European Communities: Brussels, European Union; Volume 020. Available online: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:304:0018:0063:en:PDF (accessed on 3 October 2022).
29. Decreto del Ministero delle Politiche Agricole Alimentari e Forestali del 9 dicembre 2016 Indicazione Dell'origine in Etichetta della Materia Prima per il latte e i Prodotti Lattieri Caseari, in Attuazione del Regolamento (UE) n. 1169/2011, Relativo alla Fornitura di Informazioni sugli Alimenti ai Consumatori. GU Repubblica Italiana n. 15 del 19.01.2017. Available online: https://www.gazzettaufficiale.it/eli/id/2017/01/19/17A00291/sg (accessed on 2 October 2022).
30. Real Decreto 1181/2018, de 21 de Septiembre, Relativo a la Indicación del origen de la leche Utilizada como Ingrediente en el Etiquetado de la Leche y los Productos Lácteos. Available online: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2018-12837 (accessed on 4 October 2022).
31. WTO. Uruguay Round Agreement: TRIPS. In *Part II—Standards Concerning the Availability, Scope and Use of Intellectual Property Rights, Section 3: Geographical Indications*; World Trade Organization: Geneva, Switzerland, 1994. Available online: https://www.wto.org/english/docs_e/legal_e/trips_e.htm#part2 (accessed on 4 October 2022).
32. Esteki, M.; Shahsavari, Z.; Simal-Gandara, J. Gas Chromatographic Fingerprinting Coupled to Chemometrics for Food Authentication. *Food Rev. Int.* **2020**, *36*, 384–427. [CrossRef]
33. Kamilari, E.; Tomazou, M.; Antoniades, A.; Tsaltas, D. High Throughput Sequencing Technologies as a New Toolbox for Deep Analysis, Characterization and Potentially Authentication of Protection Designation of Origin Cheeses? *Int. J. Food Sci.* **2019**, *2019*, 1–15. [CrossRef] [PubMed]
34. Bontempo, L.; Barbero, A.; Bertoldi, D.; Camin, F.; Larcher, R.; Perini, M.; Sepulcri, A.; Zicarelli, L.; Piasentier, E. Isotopic and Elemental Profiles of Mediterranean Buffalo Milk and Cheese and Authentication of Mozzarella Di Bufala Campana PDO: An Initial Exploratory Study. *Food Chem.* **2019**, *285*, 316–323. [CrossRef] [PubMed]
35. Ruth, S.M.v.; Granato, D. Food Identity, Authenticity and Fraud: The Full Spectrum. *Foods* **2017**, *6*, 49. [CrossRef]
36. Kędzierska-Matysek, M.; Florek, M.; Skałecki, P.; Litwińczuk, A.; Chruścicki, A. A comparison of the physicochemical characteristics of the regional cheese Oscypek and the traditional cheese Gazdowski from the Polish Podhale. *Int. J. Dairy Technol.* **2014**, *67*, 283–289. [CrossRef]
37. Salzano, A.; Manganiello, G.; Neglia, G.; Vinale, F.; De Nicola, D.; D'Occhio, M.; Campanile, G. A Preliminary Study on Metabolome Profiles of Buffalo Milk and Corresponding Mozzarella Cheese: Safeguarding the Authenticity and Traceability of Protected Status Buffalo Dairy Products. *Molecules* **2020**, *25*, 304. [CrossRef]
38. O'Sullivan, R.; Schmidt, O.; Monahan, F.J. Stable Isotope Ratio Analysis for the Authentication of Milk and Dairy Ingredients: A Review. *Int. Dairy J.* **2022**, *126*, 105268. [CrossRef]
39. Camin, F.; Bertoldi, D.; Santato, A.; Bontempo, L.; Perini, M.; Ziller, L.; Stroppa, A.; Larcher, R. Validation of Methods for H, C, N and S Stable Isotopes and Elemental Analysis of Cheese: Results of an International Collaborative Study: Validation of Methods for Isotopic and Elemental Analysis of Cheese. *Rapid Commun. Mass Spectrom.* **2015**, *29*, 415–423. [CrossRef]

40. Novelli, E.; Balzan, S.; Perini, M.; Antonetti, P.; Fasolato, L.; Camin, F. ANALYTICAL EMPLOYMENT OF STABLE ISOTOPES OF CARBON, NITROGEN, OXYGEN AND HYDROGEN FOR FOOD AUTHENTICATION. *Ital. J. Food Saf.* **2011**, *1*, 71. [CrossRef]
41. Bontempo, L.; Lombardi, G.; Paoletti, R.; Ziller, L.; Camin, F. H, C, N and O Stable Isotope Characteristics of Alpine Forage, Milk and Cheese. *Int. Dairy J.* **2012**, *23*, 99–104. [CrossRef]
42. Silva, A.V.; Hélie, J.F.; Caxito, F.d.A.; Monardes, H.; Mustafa, A.F.; Stevenson, R. Multi-Stable Isotope Analysis as a Tool for Assessing the Geographic Provenance of Dairy Products: A Case Study Using Buffalo's Milk and Cheese Samples from the Amazon Basin, Brazil. *Int. Dairy J.* **2014**, *35*, 107–110. [CrossRef]
43. Altieri, S.; Saiano, K.; Biondi, M.; Ricci, P.; Lubritto, C. Traceability of 'Mozzarella Di Bufala Campana' Production Chain by Means of Carbon, Nitrogen and Oxygen Stable Isotope Ratios. *J. Sci. Food Agric.* **2020**, *100*, 995–1003. [CrossRef]
44. Luykx, D.M.A.M.; van Ruth, S.M. An Overview of Analytical Methods for Determining the Geographical Origin of Food Products. *Food Chem.* **2008**, *107*, 897–911. [CrossRef]
45. Pianezze, S.; Bontempo, L.; Perini, M.; Tonon, A.; Ziller, L.; Franceschi, P.; Camin, F. δ34S for tracing the origin of cheese and detecting its authenticity. *J. Mass Spectrom.* **2020**, *55*, e4451. [CrossRef] [PubMed]
46. Manca, G.; Camin, F.; Coloru, G.C.; Del Caro, A.; Depentori, D.; Franco, M.A.; Versini, G. Characterization of the Geographical Origin of Pecorino Sardo Cheese by Casein Stable Isotope (13C/12C and 15N/14N) Ratios and Free Amino Acid Ratios. *J. Agric. Food Chem.* **2001**, *49*, 1404–1409. [CrossRef]
47. Katerinopoulou, K.; Kontogeorgos, A.; Salmas, C.E.; Patakas, A.; Ladavos, A. Geographical Origin Authentication of Agri-Food Products: A Review. *Foods* **2020**, *9*, 489. [CrossRef]
48. Pillonel, L.; Tabacchi, R.; Bosset, J.O. Geographic authenticity of Swiss Cheeses: Selected results for the food control laboratories and perspective for the future. *Mitt. Lebensm. Hyg.* **2004**, *95*, 503–513.
49. Danezis, G.P.; Tsagkaris, A.S.; Camin, F.; Brusic, V.; Georgiou, C.A. Food Authentication: Techniques, Trends & Emerging Approaches. *TrAC Trends Anal. Chem.* **2016**, *85*, 123–132. [CrossRef]
50. Danezis, G.; Theodorou, C.; Massouras, T.; Zoidis, E.; Hadjigeorgiou, I.; Georgiou, C. Greek Graviera Cheese Assessment through Elemental Metabolomics—Implications for Authentication, Safety and Nutrition. *Molecules* **2019**, *24*, 670. [CrossRef]
51. Nečemer, M.; Potočnik, D.; Ogrinc, N. Discrimination between Slovenian cow, goat and sheep milk and cheese according to geographical origin using a combination of elemental content and stable isotope data. *J. Food Compos. Anal.* **2016**, *52*, 16–23. [CrossRef]
52. Bontempo, L.; Larcher, R.; Camin, F.; Hölzl, S.; Rossmann, A.; Horn, P.; Nicolini, G. Elemental and Isotopic Characterisation of Typical Italian Alpine Cheeses. *Int. Dairy J.* **2011**, *21*, 441–446. [CrossRef]
53. Danezis, G.P.; Pappas, A.C.; Tsiplakou, E.; Pappa, E.C.; Zacharioudaki, M.; Tsagkaris, A.S.; Papachristidis, C.A.; Sotirakoglou, K.; Zervas, G.; Georgiou, C.A. Authentication of Greek Protected Designation of Origin Cheeses through Elemental Metabolomics. *Int. Dairy J.* **2020**, *104*, 104599. [CrossRef]
54. Camin, F.; Wehrens, R.; Bertoldi, D.; Bontempo, L.; Ziller, L.; Perini, M.; Nicolini, G.; Nocetti, M.; Larcher, R. H, C, N and S Stable Isotopes and Mineral Profiles to Objectively Guarantee the Authenticity of Grated Hard Cheeses. *Anal. Chim. Acta* **2012**, *711*, 54–59. [CrossRef]
55. Yang, J.; Xu, J.; Zhang, X.; Wu, C.; Lin, T.; Ying, Y. Deep Learning for Vibrational Spectral Analysis: Recent Progress and a Practical Guide. *Anal. Chim. Acta* **2019**, *1081*, 6–17. [CrossRef] [PubMed]
56. Ozaki, Y. Infrared Spectroscopy—Mid-Infrared, Near-Infrared, and Far-Infrared/Terahertz Spectroscopy. *Anal. Sci.* **2021**, *37*, 1193–1212. [CrossRef]
57. De Marchi, M.; Penasa, M.; Zidi, A.; Manuelian, C.L. Invited Review: Use of Infrared Technologies for the Assessment of Dairy Products—Applications and Perspectives. *J. Dairy Sci.* **2018**, *101*, 10589–10604. [CrossRef]
58. Holroyd, S.E. The Use of near Infrared Spectroscopy on Milk and Milk Products. *J. Near Infrared Spectrosc.* **2013**, *21*, 311–322. [CrossRef]
59. Lerma-García, M.J.; Gori, A.; Cerretani, L.; Simó-Alfonso, E.F.; Caboni, M.F. Classification of Pecorino Cheeses Produced in Italy According to Their Ripening Time and Manufacturing Technique Using Fourier Transform Infrared Spectroscopy. *J. Dairy Sci.* **2010**, *93*, 4490–4496. [CrossRef]
60. Ottavian, M.; Facco, P.; Barolo, M.; Berzaghi, P.; Segato, S.; Novelli, E.; Balzan, S. Near-Infrared Spectroscopy to Assist Authentication and Labeling of Asiago d'allevo Cheese. *J. Food Eng.* **2012**, *113*, 289–298. [CrossRef]
61. Stocco, G.; Cipolat-Gotet, C.; Ferragina, A.; Berzaghi, P.; Bittante, G. Accuracy and Biases in Predicting the Chemical and Physical Traits of Many Types of Cheeses Using Different Visible and Near-Infrared Spectroscopic Techniques and Spectrum Intervals. *J. Dairy Sci.* **2019**, *102*, 9622–9638. [CrossRef]
62. Bergamaschi, M.; Cipolat-Gotet, C.; Cecchinato, A.; Schiavon, S.; Bittante, G. Chemometric Authentication of Farming Systems of Origin of Food (Milk and Ripened Cheese) Using Infrared Spectra, Fatty Acid Profiles, Flavor Fingerprints, and Sensory Descriptions. *Food Chem.* **2020**, *305*, 125480. [CrossRef] [PubMed]
63. Coppa, M.; Martin, B.; Agabriel, C.; Chassaing, C.; Sibra, C.; Constant, I.; Graulet, B.; Andueza, D. Authentication of Cow Feeding and Geographic Origin on Milk Using Visible and Near-Infrared Spectroscopy. *J. Dairy Sci.* **2012**, *95*, 5544–5551. [CrossRef] [PubMed]
64. Niemöller, A.; Holroyd, S. Applications of near infrared spectrometry for the analysis of milk and milk products. *Bull. Int. Dairy Fed.* **2019**, *497*, 19.

65. Bittante, G.; Patel, N.; Cecchinato, A.; Berzaghi, P. Invited Review: A Comprehensive Review of Visible and near-Infrared Spectroscopy for Predicting the Chemical Composition of Cheese. *J. Dairy Sci.* **2022**, *105*, 1817–1836. [CrossRef]
66. Andueza, D.; Agabriel, C.; Constant, I.; Lucas, A.; Martin, B. Using Visible or near Infrared Spectroscopy (NIRS) on Cheese to Authenticate Cow Feeding Regimes. *Food Chem.* **2013**, *141*, 209–214. [CrossRef]
67. Karoui, R.; Dufour, É.; Pillonel, L.; Schaller, E.; Picque, D.; Cattenoz, T.; Bosset, J.-O. The Potential of Combined Infrared and Fluorescence Spectroscopies as a Method of Determination of the Geographic Origin of Emmental Cheeses. *Int. Dairy J.* **2005**, *15*, 287–298. [CrossRef]
68. Karoui, R.; Mazerolles, G.; Bosset, J.; Debaerdemaeker, J.; Dufour, E. Utilisation of Mid-Infrared Spectroscopy for Determination of the Geographic Origin of Gruyère PDO and L'Etivaz PDO Swiss Cheeses. *Food Chem.* **2007**, *105*, 847–854. [CrossRef]
69. Mazzei, P.; Piccolo, A. HRMAS NMR Spectroscopy Applications in Agriculture. *Chem. Biol. Technol. Agric.* **2017**, *4*, 11. [CrossRef]
70. Mazzei, P.; Piccolo, A. 1H HRMAS-NMR Metabolomic to Assess Quality and Traceability of Mozzarella Cheese from Campania Buffalo Milk. *Food Chem.* **2012**, *132*, 1620–1627. [CrossRef]
71. Shintu, L.; Caldarelli, S. Toward the Determination of the Geographical Origin of Emmental(Er) Cheese via High Resolution MAS NMR: A Preliminary Investigation. *J. Agric. Food Chem.* **2006**, *54*, 4148–4154. [CrossRef]
72. Consonni, R.; Cagliani, L.R. Ripening and geographical characterization of Parmigiano Reggiano cheese by 1H NMR spectroscopy. *Talanta* **2008**, *76*, 200–205. [CrossRef] [PubMed]
73. Balthazar, C.F.; Guimarães, J.T.; Rocha, R.S.; Pimentel, T.C.; Neto, R.P.C.; Tavares, M.I.B.; Graça, J.S.; Alves Filho, E.G.; Freitas, M.Q.; Esmerino, E.A.; et al. Nuclear Magnetic Resonance as an Analytical Tool for Monitoring the Quality and Authenticity of Dairy Foods. *Trends Food Sci. Technol.* **2021**, *108*, 84–91. [CrossRef]
74. Emwas, A.-H.; Roy, R.; McKay, R.T.; Tenori, L.; Saccenti, E.; Gowda, G.A.N.; Raftery, D.; Alahmari, F.; Jaremko, L.; Jaremko, M.; et al. NMR Spectroscopy for Metabolomics Research. *Metabolites* **2019**, *9*, 123. [CrossRef] [PubMed]
75. Hewavitharana, G.G.; Perera, D.N.; Navaratne, S.B.; Wickramasinghe, I. Extraction Methods of Fat from Food Samples and Preparation of Fatty Acid Methyl Esters for Gas Chromatography: A Review. *Arab. J. Chem.* **2020**, *13*, 6865–6875. [CrossRef]
76. Reis Lima, M.J.; Bahri, H.; Sá Morais, J.; Veloso, A.C.A.; Fontes, L.; Lemos, E.T.; Peres, A.M. Assessing Serra Da Estrela PDO Cheeses' Origin-Production Date Using Fatty Acids Profiles. *Food Meas.* **2019**, *13*, 2988–2997. [CrossRef]
77. Ahad, T.; Nissar, J. Fingerprinting in Determining the Adultration of Food. *J. Pharmacogn. Phytochem.* **2017**, *6*, 1543–1553.
78. Holm, T. Aspects of the Mechanism of the Flame Ionization Detector. *J. Chromatogr. A* **1999**, *842*, 221–227. [CrossRef]
79. Gaspardo, B.; Lavrenčič, A.; Levart, A.; Del Zotto, S.; Stefanon, B. Use of milk fatty acids composition to discriminate area of origin of bulk milk. *J. Dairy Sci.* **2010**, *93*, 3417–3426. [CrossRef]
80. Bonanno, A.; Tornambè, G.; Bellina, V.; De Pasquale, C.; Mazza, F.; Maniaci, G.; Di Grigoli, A. Effect of farming system and cheesemaking technology on the physicochemical characteristics, fatty acid profile, and sensory properties of Caciocavallo Palermitano cheese. *J. Dairy Sci.* **2013**, *96*, 710–724. [CrossRef]
81. Fox, P.F.; Guinee, T.P.; Cogan, T.M.; McSweeney, P.L.H. Microbiology of Cheese Ripening. In *Fundamentals of Cheese Science*; Springer: Boston, MA, USA, 2017; pp. 333–390. [CrossRef]
82. Margalho, L.P.; Kamimura, B.A.; Pimentel, T.C.; Balthazar, C.F.; Araujo, J.V.A.; Silva, R.; Conte-Junior, C.A.; Raices, R.S.L.; Cruz, A.G.; Sant'Ana, A.S. A Large Survey of the Fatty Acid Profile and Gross Composition of Brazilian Artisanal Cheeses. *J. Food Compos. Anal.* **2021**, *101*, 103955. [CrossRef]
83. Danezis, G.P.; Tsiplakou, E.; Pappa, E.C.; Pappas, A.C.; Mavrommatis, A.; Sotirakoglou, K.; Georgiou, C.A.; Zervas, G. Fatty Acid Profile and Physicochemical Properties of Greek Protected Designation of Origin Cheeses, Implications for Authentication. *Eur. Food Res. Technol.* **2020**, *246*, 1741–1753. [CrossRef]
84. Gatzias, I.S.; Karabagias, I.K.; Kontominas, M.G.; Badeka, A.V. Geographical Differentiation of Feta Cheese from Northern Greece Based on Physicochemical Parameters, Volatile Compounds and Fatty Acids. *LWT* **2020**, *131*, 109615. [CrossRef]
85. Vatavali, K.; Kosma, I.; Louppis, A.; Gatzias, I.; Badeka, A.V.; Kontominas, M.G. Characterisation and Differentiation of Geographical Origin of Graviera Cheeses Produced in Greece Based on Physico-Chemical, Chromatographic and Spectroscopic Analyses, in Combination with Chemometrics. *Int. Dairy J.* **2020**, *110*, 104799. [CrossRef]
86. Bhandari, M.P.; Carmona, E.N.; Galstyan, V.; Sberveglieri, V. Quality Evaluation of Parmigiano Reggiano Cheese by a Novel Nanowire Device S3 and Evaluation of the VOCs Profile. *Procedia Eng.* **2016**, *168*, 460–464. [CrossRef]
87. Afshari, R.; Pillidge, C.J.; Dias, D.A.; Osborn, A.M.; Gill, H. Cheesomics: The Future Pathway to Understanding Cheese Flavour and Quality. *Crit. Rev. Food Sci. Nutr.* **2020**, *60*, 33–47. [CrossRef] [PubMed]
88. Turri, F.; Cremonesi, P.; Battelli, G.; Severgnini, M.; Brasca, M.; Gandini, G.; Pizzi, F. High Biodiversity in a Limited Mountain Area Revealed in the Traditional Production of Historic Rebel Cheese by an Integrated Microbiota–Lipidomic Approach. *Sci. Rep.* **2021**, *11*, 10374. [CrossRef] [PubMed]
89. Medina, S.; Perestrelo, R.; Silva, P.; Pereira, J.A.M.; Câmara, J.S. Current Trends and Recent Advances on Food Authenticity Technologies and Chemometric Approaches. *Trends Food Sci. Technol.* **2019**, *85*, 163–176. [CrossRef]
90. Pillonel, L.; Ampuero, S.; Tabacchi, R.; Bosset, J. Analytical Methods for the Determination of the Geographic Origin of Emmental Cheese: Volatile Compounds by GC/MS-FID and Electronic Nose. *Eur. Food Res. Technol.* **2003**, *216*, 179–183. [CrossRef]
91. Pluta-Kubica, A.; Domagała, J.; Gąsior, R.; Wojtycza, K.; Witczak, M. Characterisation of the Profile of Volatiles of Polish Emmental Cheese. *Int. Dairy J.* **2021**, *116*, 104954. [CrossRef]

92. Di Donato, F.; Biancolillo, A.; Mazzulli, D.; Rossi, L.; D'Archivio, A.A. HS-SPME/GC–MS Volatile Fraction Determination and Chemometrics for the Discrimination of Typical Italian Pecorino Cheeses. *Microchem. J.* **2021**, *165*, 106133. [CrossRef]
93. Caligiani, A.; Nocetti, M.; Lolli, V.; Marseglia, A.; Palla, G. Development of a Quantitative GC–MS Method for the Detection of Cyclopropane Fatty Acids in Cheese as New Molecular Markers for Parmigiano Reggiano Authentication. *J. Agric. Food Chem.* **2016**, *64*, 4158–4164. [CrossRef] [PubMed]
94. Tompa, G.; Susi, R.; Rogelj, I.; Pompe, M. Cryotrap/SPME/GC/MS Method for Profiling of Monoterpenes in Cheese and Their Clustering According to Geographic Origin. *Acta Chim. Slov.* **2013**, *60*, 595–603. [PubMed]
95. Song, S.; Yuan, L.; Zhang, X.; Hayat, K.; Chen, H.; Liu, F.; Xiao, Z.; Niu, Y. Rapid Measuring and Modelling Flavour Quality Changes of Oxidised Chicken Fat by Electronic Nose Profiles through the Partial Least Squares Regression Analysis. *Food Chem.* **2013**, *141*, 4278–4288. [CrossRef] [PubMed]
96. Fujioka, K. Comparison of Cheese Aroma Intensity Measured Using an Electronic Nose (E-Nose) Non-Destructively with the Aroma Intensity Scores of a Sensory Evaluation: A Pilot Study. *Sensors* **2021**, *21*, 8368. [CrossRef] [PubMed]
97. Gardner, J.W.; Bartlett, P.N. A Brief History of Electronic Noses. *Sens. Actuators B Chem.* **1994**, *18*, 210–211. [CrossRef]
98. Gliszczyńska-Świgło, A.; Chmielewski, J. Electronic nose as a tool for monitoring the authenticity of food. *A Rev. Food Anal. Methods* **2017**, *10*, 1800–1816. [CrossRef]
99. Ampuero, S.; Bosset, J.O. The Electronic Nose Applied to Dairy Products: A Review. *Sens. Actuators B Chem.* **2003**, *94*, 1–12. [CrossRef]
100. Cevoli, C.; Cerretani, L.; Gori, A.; Caboni, M.F.; Gallina Toschi, T.; Fabbri, A. Classification of Pecorino Cheeses Using Electronic Nose Combined with Artificial Neural Network and Comparison with GC–MS Analysis of Volatile Compounds. *Food Chem.* **2011**, *129*, 1315–1319. [CrossRef]
101. Lytou, A.E.; Panagou, E.Z.; Nychas, G.-J.E. Volatilomics for Food Quality and Authentication. *Curr. Opin. Food Sci.* **2019**, *28*, 88–95. [CrossRef]
102. Roy, M.; Yadav, B.K. Electronic Nose for Detection of Food Adulteration: A Review. *J. Food Sci. Technol.* **2022**, *59*, 846–858. [CrossRef]
103. Gómez-Ruiz, J.Á.; Ballesteros, C.; Viñas, M.Á.G.; Cabezas, L.; Martínez-Castro, I. Relationships between Volatile Compounds and Odour in Manchego Cheese: Comparison between Artisanal and Industrial Cheeses at Different Ripening Times. *Lait* **2002**, *82*, 613–628. [CrossRef]
104. Ardö, Y.; McSweeney, P.L.; Magboul, A.A.; Upadhyay, V.K.; Fox, P.F. Biochemistry of cheese ripening: Proteolysis. In *Cheese*; Academic Press: Cambridge, MA, USA, 2017; pp. 445–482. [CrossRef]
105. Galle, S.A.; Koot, A.; Soukoulis, C.; Cappellin, L.; Biasioli, F.; Alewijn, M.; van Ruth, S.M. Typicality and Geographical Origin Markers of Protected Origin Cheese from The Netherlands Revealed by PTR-MS. *J. Agric. Food Chem.* **2011**, *59*, 2554–2563. [CrossRef] [PubMed]
106. D'Incecco, P.; Limbo, S.; Hogenboom, J.; Rosi, V.; Gobbi, S.; Pellegrino, L. Impact of extending hard-cheese ripening: A multiparameter characterization of Parmigiano reggiano cheese ripened up to 50 months. *Foods* **2020**, *9*, 268. [CrossRef] [PubMed]
107. Quigley, L.; O'Sullivan, O.; Beresford, T.P.; Ross, R.P.; Fitzgerald, G.F.; Cotter, P.D. High-Throughput Sequencing for Detection of Subpopulations of Bacteria Not Previously Associated with Artisanal Cheeses. *Appl. Environ. Microbiol.* **2012**, *78*, 5717–5723. [CrossRef]
108. Kleerebezem, R.; Stouten, G.; Koehorst, J.; Langenhoff, A.; Schaap, P.; Smidt, H. Experimental Infrastructure Requirements for Quantitative Research on Microbial Communities. *Curr. Opin. Biotechnol.* **2021**, *67*, 158–165. [CrossRef] [PubMed]
109. Wolfe, B.E.; Button, J.E.; Santarelli, M.; Dutton, R.J. Cheese Rind Communities Provide Tractable Systems for In Situ and In Vitro Studies of Microbial Diversity. *Cell* **2014**, *158*, 422–433. [CrossRef]
110. Ceugniez, A.; Taminiau, B.; Coucheney, F.; Jacques, P.; Delcenserie, V.; Daube, G.; Drider, D. Fungal Diversity of "Tomme d'Orchies" Cheese during the Ripening Process as Revealed by a Metagenomic Study. *Int. J. Food Microbiol.* **2017**, *258*, 89–93. [CrossRef]
111. Duru, I.C.; Laine, P.; Andreevskaya, M.; Paulin, L.; Kananen, S.; Tynkkynen, S.; Auvinen, P.; Smolander, O.-P. Metagenomic and Metatranscriptomic Analysis of the Microbial Community in Swiss-Type Maasdam Cheese during Ripening. *Int. J. Food Microbiol.* **2018**, *281*, 10–22. [CrossRef]
112. Hugerth, L.W.; Andersson, A.F. Analysing Microbial Community Composition through Amplicon Sequencing: From Sampling to Hypothesis Testing. *Front. Microbiol.* **2017**, *8*, 1561. [CrossRef]
113. Tilocca, B.; Costanzo, N.; Morittu, V.M.; Spina, A.A.; Soggiu, A.; Britti, D.; Roncada, P.; Piras, C. Milk Microbiota: Characterization Methods and Role in Cheese Production. *J. Proteom.* **2020**, *210*, 103534. [CrossRef]
114. Reitmeier, S.; Hitch, T.C.A.; Treichel, N.; Fikas, N.; Hausmann, B.; Ramer-Tait, A.E.; Neuhaus, K.; Berry, D.; Haller, D.; Lagkouvardos, I.; et al. Handling of Spurious Sequences Affects the Outcome of High-Throughput 16S RRNA Gene Amplicon Profiling. *ISME Commun.* **2021**, *1*, 31. [CrossRef]
115. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME Allows Analysis of High-Throughput Community Sequencing Data. *Nat. Methods* **2010**, *7*, 335–336. [CrossRef] [PubMed]

116. 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]
117. 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]
118. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857. [CrossRef]
119. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools. *Nucleic Acids Res.* **2012**, *41*, D590–D596. [CrossRef]
120. Nilsson, R.H.; Larsson, K.-H.; Taylor, A.F.S.; Bengtsson-Palme, J.; Jeppesen, T.S.; Schigel, D.; Kennedy, P.; Picard, K.; Glöckner, F.O.; Tedersoo, L.; et al. The UNITE Database for Molecular Identification of Fungi: Handling Dark Taxa and Parallel Taxonomic Classifications. *Nucleic Acids Res.* **2019**, *47*, D259–D264. [CrossRef]
121. Mizrahi-Man, O.; Davenport, E.R.; Gilad, Y. Taxonomic Classification of Bacterial 16S RRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs. *PLoS ONE* **2013**, *8*, e53608. [CrossRef]
122. Dalmasso, A.; Soto del Rio, M.d.l.D.; Civera, T.; Pattono, D.; Cardazzo, B.; Bottero, M.T. Characterization of Microbiota in Plaisentif Cheese by High-Throughput Sequencing. *LWT-Food Sci. Technol.* **2016**, *69*, 490–496. [CrossRef]
123. Gobbetti, M.; Di Cagno, R.; Calasso, M.; Neviani, E.; Fox, P.F.; De Angelis, M. Drivers That Establish and Assembly the Lactic Acid Bacteria Biota in Cheeses. *Trends Food Sci. Technol.* **2018**, *78*, 244–254. [CrossRef]
124. Piqueras, J.; Chassard, C.; Callon, C.; Rifa, E.; Theil, S.; Lebecque, A.; Delbès, C. Lactic Starter Dose Shapes S. Aureus and STEC O26:H11 Growth, and Bacterial Community Patterns in Raw Milk Uncooked Pressed Cheeses. *Microorganisms* **2021**, *9*, 1081. [CrossRef] [PubMed]
125. Porcellato, D.; Smistad, M.; Bombelli, A.; Abdelghani, A.; Jørgensen, H.J.; Skeie, S.B. Longitudinal Study of the Bulk Tank Milk Microbiota Reveals Major Temporal Shifts in Composition. *Front. Microbiol.* **2021**, *12*, 616429. [CrossRef] [PubMed]
126. Skeie, S.B.; Håland, M.; Thorsen, I.M.; Narvhus, J.; Porcellato, D. Bulk Tank Raw Milk Microbiota Differs within and between Farms: A Moving Goalpost Challenging Quality Control. *J. Dairy Sci.* **2019**, *102*, 1959–1971. [CrossRef]
127. Guo, X.; Yu, Z.; Zhao, F.; Sun, Z.; Kwok, L.-Y.; Li, S. Both Sampling Seasonality and Geographic Origin Contribute Significantly to Variations in Raw Milk Microbiota, but Sampling Seasonality Is the More Determining Factor. *J. Dairy Sci.* **2021**, *104*, 10609–10627. [CrossRef] [PubMed]
128. Licitra, G.; Caccamo, M.; Lortal, S. Artisanal Products Made With Raw Milk. In *Raw Milk*; Elsevier: San Diego, CA, USA, 2019; pp. 175–221. [CrossRef]
129. Mounier, J.; Gelsomino, R.; Goerges, S.; Vancanneyt, M.; Vandemeulebroecke, K.; Hoste, B.; Scherer, S.; Swings, J.; Fitzgerald, G.F.; Cogan, T.M. Surface Microflora of Four Smear-Ripened Cheeses. *Appl. Environ. Microbiol.* **2005**, *71*, 6489–6500. [CrossRef]
130. Bokulich, N.A.; Mills, D.A. Facility-Specific "House" Microbiome Drives Microbial Landscapes of Artisan Cheesemaking Plants. *Appl. Environ. Microbiol.* **2013**, *79*, 5214–5223. [CrossRef]
131. Calasso, M.; Ercolini, D.; Mancini, L.; Stellato, G.; Minervini, F.; Di Cagno, R.; De Angelis, M.; Gobbetti, M. Relationships among House, Rind and Core Microbiotas during Manufacture of Traditional Italian Cheeses at the Same Dairy Plant. *Food Microbiol.* **2016**, *54*, 115–126. [CrossRef]
132. Sun, L.; D'Amico, D.J. Composition, Succession, and Source Tracking of Microbial Communities throughout the Traditional Production of a Farmstead Cheese. *mSystems* **2021**, *6*, e00830-21. [CrossRef]
133. Knights, D.; Kuczynski, J.; Charlson, E.S.; Zaneveld, J.; Mozer, M.C.; Collman, R.G.; Bushman, F.D.; Knight, R.; Kelley, S.T. Bayesian Community-Wide Culture-Independent Microbial Source Tracking. *Nat. Methods* **2011**, *8*, 761–763. [CrossRef]
134. Montel, M.-C.; Buchin, S.; Mallet, A.; Delbes-Paus, C.; Vuitton, D.A.; Desmasures, N.; Berthier, F. Traditional Cheeses: Rich and Diverse Microbiota with Associated Benefits. *Int. J. Food Microbiol.* **2014**, *177*, 136–154. [CrossRef]
135. Busetta, G.; Garofalo, G.; Mangione, G.; Botta, L.; Franciosi, E.; Di Gerlando, R.; Todaro, M.; Licitra, G.; Scatassa, M.L.; Gaglio, R.; et al. Polyphasic Characterization of Microbiota of "Mastredda", a Traditional Wooden Tool Used during the Production of PDO Provola Dei Nebrodi Cheese. *Appl. Sci.* **2021**, *11*, 8647. [CrossRef]
136. Zago, M.; Bardelli, T.; Rossetti, L.; Nazzicari, N.; Carminati, D.; Galli, A.; Giraffa, G. Evaluation of Bacterial Communities of Grana Padano Cheese by DNA Metabarcoding and DNA Fingerprinting Analysis. *Food Microbiol.* **2021**, *93*, 103613. [CrossRef] [PubMed]
137. Kamimura, B.A.; De Filippis, F.; Sant'Ana, A.S.; Ercolini, D. Large-Scale Mapping of Microbial Diversity in Artisanal Brazilian Cheeses. *Food Microbiol.* **2019**, *80*, 40–49. [CrossRef]
138. Sant'Anna, F.M.; Wetzels, S.U.; Cicco, S.H.S.; Figueiredo, R.C.; Sales, G.A.; Figueiredo, N.C.; Nunes, C.A.; Schmitz-Esser, S.; Mann, E.; Wagner, M.; et al. Microbial Shifts in Minas Artisanal Cheeses from the Serra Do Salitre Region of Minas Gerais, Brazil throughout Ripening Time. *Food Microbiol.* **2019**, *82*, 349–362. [CrossRef]
139. Li, J.; Zheng, Y.; Xu, H.; Xi, X.; Hou, Q.; Feng, S.; Wuri, L.; Bian, Y.; Yu, Z.; Kwok, L.-Y.; et al. Bacterial Microbiota of Kazakhstan Cheese Revealed by Single Molecule Real Time (SMRT) Sequencing and Its Comparison with Belgian, Kalmykian and Italian Artisanal Cheeses. *BMC Microbiol.* **2017**, *17*, 13. [CrossRef] [PubMed]

140. Lucchini, R.; Cardazzo, B.; Carraro, L.; Negrinotti, M.; Balzan, S.; Novelli, E.; Fasolato, L.; Fasoli, F.; Farina, G. Contribution of Natural Milk Culture to Microbiota, Safety and Hygiene of Raw Milk Cheese Produced in Alpine Malga. *Ital. J. Food Saf.* **2018**, *7*, 6967. [CrossRef]
141. Xue, J.; Yang, Y.; Wang, Z.; Guo, Y.; Shao, Y. Bacterial Diversity in Chinese Rushan Cheese From Different Geographical Origins. *Front. Microbiol.* **2018**, *9*, 1920. [CrossRef]
142. Johnson, J.; Curtin, C.; Waite-Cusic, J. The Cheese Production Facility Microbiome Exhibits Temporal and Spatial Variability. *Front. Microbiol.* **2021**, *12*, 644828. [CrossRef]
143. Kamilari, E.; Anagnostopoulos, D.A.; Papademas, P.; Kamilaris, A.; Tsaltas, D. Characterizing Halloumi Cheese's Bacterial Communities through Metagenomic Analysis. *LWT* **2020**, *126*, 109298. [CrossRef]
144. Aldrete-Tapia, A.; Escobar-Ramírez, M.C.; Tamplin, M.L.; Hernández-Iturriaga, M. High-Throughput Sequencing of Microbial Communities in Poro Cheese, an Artisanal Mexican Cheese. *Food Microbiol.* **2014**, *44*, 136–141. [CrossRef]
145. Porcellato, D.; Skeie, S.B. Bacterial Dynamics and Functional Analysis of Microbial Metagenomes during Ripening of Dutch-Type Cheese. *Int. Dairy J.* **2016**, *61*, 182–188. [CrossRef]
146. Cardinali, F.; Ferrocino, I.; Milanović, V.; Belleggia, L.; Corvaglia, M.R.; Garofalo, C.; Foligni, R.; Mannozzi, C.; Mozzon, M.; Cocolin, L.; et al. Microbial Communities and Volatile Profile of Queijo de Azeitão PDO Cheese, a Traditional Mediterranean Thistle-Curdled Cheese from Portugal. *Food Res. Int.* **2021**, *147*, 110537. [CrossRef] [PubMed]
147. Biolcati, F.; Bottero, M.T.; Dalmasso, A. Microbiological Analysis of the Robiola Di Roccaverano Cheese by Means of Traditional Culturing Methods. *Ital. J. Food Saf.* **2019**, 8. [CrossRef] [PubMed]
148. Martín, J.F.; Coton, M. Blue cheese: Microbiota and fungal metabolites. In *Fermented Foods in Health and Disease Prevention*; Academic Press: San Diego, CA, USA, 2017; pp. 275–303. [CrossRef]
149. Coton, E.; Coton, M.; Hymery, N.; Mounier, J.; Jany, J.-L. Penicillium Roqueforti: An Overview of Its Genetics, Physiology, Metabolism and Biotechnological Applications. *Fungal Biol. Rev.* **2020**, *34*, 59–73. [CrossRef]
150. Gillot, G.; Jany, J.-L.; Poirier, E.; Maillard, M.-B.; Debaets, S.; Thierry, A.; Coton, E.; Coton, M. Functional Diversity within the Penicillium Roqueforti Species. *Int. J. Food Microbiol.* **2017**, *241*, 141–150. [CrossRef]
151. Garnier, L.; Valence, F.; Mounier, J. Diversity and Control of Spoilage Fungi in Dairy Products: An Update. *Microorganisms* **2017**, *5*, 42. [CrossRef]
152. Mounier, J.; Coton, M. *Kluyveromyces* spp. In *Encyclopedia of Dairy Sciences*; McSweeney, P.L.H., McNamara, J.P., Eds.; Elsevier: San Diego, CA, USA, 2022; Volume 4, pp. 569–574.
153. Lundin, D.; Severin, I.; Logue, J.B.; Östman, Ö.; Andersson, A.F.; Lindström, E.S. Which Sequencing Depth Is Sufficient to Describe Patterns in Bacterial α- and β-Diversity? Sequencing Depth in Diversity Research. *Environ. Microbiol. Rep.* **2012**, *4*, 367–372. [CrossRef]
154. Rubin, B.E.R.; Sanders, J.G.; Hampton-Marcell, J.; Owens, S.M.; Gilbert, J.A.; Moreau, C.S. DNA Extraction Protocols Cause Differences in 16S RRNA Amplicon Sequencing Efficiency but Not in Community Profile Composition or Structure. *MicrobiologyOpen* **2014**, *3*, 910–921. [CrossRef]
155. Nearing, J.T.; Douglas, G.M.; Comeau, A.M.; Langille, M.G.I. Denoising the Denoisers: An Independent Evaluation of Microbiome Sequence Error-Correction Approaches. *PeerJ* **2018**, *6*, e5364. [CrossRef]
156. Unno, R.; Suzuki, T.; Matsutani, M.; Ishikawa, M. Evaluation of the Relationships Between Microbiota and Metabolites in Soft-Type Ripened Cheese Using an Integrated Omics Approach. *Front. Microbiol.* **2021**, *12*, 681185. [CrossRef]
157. Piwosz, K.; Shabarova, T.; Pernthaler, J.; Posch, T.; Šimek, K.; Porcal, P.; Salcher, M.M. Bacterial and Eukaryotic Small-Subunit Amplicon Data Do Not Provide a Quantitative Picture of Microbial Communities, but They Are Reliable in the Context of Ecological Interpretations. *mSphere* **2020**, *5*, e00052-20. [CrossRef]
158. Yeh, Y.-C.; Needham, D.M.; Sieradzki, E.T.; Fuhrman, J.A. Taxon Disappearance from Microbiome Analysis Reinforces the Value of Mock Communities as a Standard in Every Sequencing Run. *mSystems* **2018**, *3*, e00023-18. [CrossRef]
159. Brumfield, K.D.; Huq, A.; Colwell, R.R.; Olds, J.L.; Leddy, M.B. Microbial Resolution of Whole Genome Shotgun and 16S Amplicon Metagenomic Sequencing Using Publicly Available NEON Data. *PLoS ONE* **2020**, *15*, e0228899. [CrossRef] [PubMed]
160. Johnson, J.S.; Spakowicz, D.J.; Hong, B.-Y.; Petersen, L.M.; Demkowicz, P.; Chen, L.; Leopold, S.R.; Hanson, B.M.; Agresta, H.O.; Gerstein, M.; et al. Evaluation of 16S RRNA Gene Sequencing for Species and Strain-Level Microbiome Analysis. *Nat. Commun.* **2019**, *10*, 5029. [CrossRef]
161. Cassago, A.L.L.; Artêncio, M.M.; de Moura Engracia Giraldi, J.; Da Costa, F.B. Metabolomics as a Marketing Tool for Geographical Indication Products: A Literature Review. *Eur. Food Res. Technol.* **2021**, *247*, 2143–2159. [CrossRef]
162. Ranjan, R.; Rani, A.; Metwally, A.; McGee, H.S.; Perkins, D.L. Analysis of the Microbiome: Advantages of Whole Genome Shotgun versus 16S Amplicon Sequencing. *Biochem. Biophys. Res. Commun.* **2016**, *469*, 967–977. [CrossRef]
163. Quince, C.; Walker, A.W.; Simpson, J.T.; Loman, N.J.; Segata, N. Shotgun Metagenomics, from Sampling to Analysis. *Nat. Biotechnol.* **2017**, *35*, 833–844. [CrossRef]
164. Beghini, F.; McIver, L.J.; Blanco-Míguez, A.; Dubois, L.; Asnicar, F.; Maharjan, S.; Mailyan, A.; Manghi, P.; Scholz, M.; Thomas, A.M.; et al. Integrating Taxonomic, Functional, and Strain-Level Profiling of Diverse Microbial Communities with BioBakery 3. *eLife* **2021**, *10*, e65088. [CrossRef] [PubMed]
165. Mahony, J.; van Sinderen, D. Virome Studies of Food Production Systems: Time for 'Farm to Fork' Analyses. *Curr. Opin. Biotechnol.* **2022**, *73*, 22–27. [CrossRef]

166. Afshari, R.; Pillidge, C.J.; Read, E.; Rochfort, S.; Dias, D.A.; Osborn, A.M.; Gill, H. New Insights into Cheddar Cheese Microbiota-Metabolome Relationships Revealed by Integrative Analysis of Multi-Omics Data. *Sci. Rep.* **2020**, *10*, 3164. [CrossRef] [PubMed]
167. Goyal, K.; Kumar, P.; Verma, K. Food Adulteration Detection Using Artificial Intelligence: A Systematic Review. *Arch. Comput. Methods Eng.* **2022**, *29*, 397–426. [CrossRef]

Chapter 2: Characterization of typical mountain Caciotta cheese

The aim of this chapter is to investigate the ecological drivers of the microbiome of typical Caciotta cheese produced in different regions with similar cheesemaking know-how. As traditional cheeses are recognized for their diverse and distinctive sensory properties (Montel et al., 2014), Caciotta cheese from mountain areas is renowned and appreciated by consumers because of its typical flavour. The connection between the cheese microbiota and its sensory attributes is an ongoing subject of investigation, especially when employing a multi-omics approach (Afshari et., 2020). The composition and activity of the cheese microbiota, influenced by factors such as the type of starter cultures used, the environment, and the ageing conditions, can directly affect the volatilome of the cheese (Gobbeti et al., 2018). Changes in the microbial populations and in their functional potential can lead to different production of volatile compounds, resulting in changes in the aroma and flavour profile of the cheese. Shotgun metagenomics and headspace gas chromatography mass spectrometry techniques were combined to better understand the complex interactions between the cheese microbiota, its functional potential and the observed volatilome. In this context, this study examined how biotic and abiotic factors shape traditional mountain Caciotta cheese microbiome and contribute to the development of its typical flavour. Our results could help in the development of improved cheese production techniques, selection of starter cultures, and understanding the factors influencing the flavor characteristics of different cheese varieties.

Ecological diversity and associated volatilome of typical mountain Caciotta cheese from Italy

Marco Cardin[1,2], Barbara Cardazzo[1*], Monika Coton[2], Lisa Carraro[1], Rosaria Lucchini[3], Enrico Novelli[1], Emmanuel Coton[2] and Jérôme Mounier[2]

[1] *Department of Comparative Biomedicine and Food Science, University of Padova, Viale Università 16, 35020 Legnaro, PD, Italy*

[2] *Univ Brest, INRAE, Laboratoire Universitaire de Biodiversité et Écologie Microbienne, F-29280 Plouzané, France*

[3] *Italian Health Authority and Research Organization for Animal Health and Food Safety (Istituto zooprofilattico sperimentale delle Venezie), Viale Università 10, 35020 Legnaro, Pd, Italy*

**corresponding author*

Keywords: cheese, shotgun metagenomics, volatile organic compounds, cheese ecology, functional profiling, strain diversity

Abstract

Traditional products are particularly appreciated by consumers and among these products, cheese is a major contributor to the Italian mountainous area economics. In this study, shotgun metagenomics and volatilomics were used to understand the biotic and abiotic factors contributing to mountain Caciotta cheese typicity and diversity. Results showed that the geographical origin of cheese played a significant role; however, curd cooking temperature, pH, salt concentration and water activity also had an impact. Viral communities exhibited higher biodiversity and discriminated cheese origins. Among the most dominant bacteria, *Streptococcus thermophilus* showed higher intraspecific diversity and closer relationship to cheese origin when compared to *Lactobacillus delbrueckii.* However, despite a few cases in which the starter culture was phylogenetically separated from the most dominant strains sequenced in the cheese, starter cultures and dominant cheese strains clustered together suggesting substantial starter colonization in mountain Caciotta cheese. The Caciotta cheese volatilome contained prominent levels of alcohols and ketones, accompanied by lower proportions of terpenes. Volatile profile not only demonstrated a noticeable association with cheese origin but also significant differences in the relative abundances of enzymes connected to flavor development. Moreover, correlations of different non-homologous isofunctional enzymes highlighted specific contributions to the typical flavor of mountain Caciotta cheese. Overall, this study provides a deeper understanding of the factors shaping typical mountain Caciotta cheese, and the potential of metagenomics for characterizing and potentially authenticating food products.

1. Introduction

Dairy production is a major contributor to the Italian mountain area economy, cheese production being a principal part of the regional gastronomical culture. Moreover, the place of origin of food products is also a powerful attraction for consumers (Paxson 2010). For example, cheeses produced in a traditional way in specific areas (*e.g.* "product of island farming" or "mountain product") are purchased by consumers who associate them with unique typical sensorial characteristics and traditional know-how. Consumers are also ready to pay a premium price for these products.

Cheese typicity results from a combination of factors, including the milk type used for production, manufacturing practices such as the addition of natural whey or starter cultures, the use of traditional equipment, ripening conditions and the geographical origin (Kamilari et al., 2019; Cardin et al., 2022). Much emphasis has been given to cheese origin and, in Europe, this can be linked to the Protected Designation of Origin (PDO), Protected Geographical Indication (PGI) and Traditional Specialities Guaranteed (TGS) labels. However, knowledge regarding whether microbial patterns associated with regional production practices exist in ripened cheeses remains limited (Kamilari et al., 2022) and comparison with other factors shaping cheese microbiota are scarce.

If the environmental microbiome can be a source for cheese spoilage and pathogenic microorganisms (Possas et al., 2021), environmental microbiota can also harbor beneficial microorganisms such as non-starter lactic acid bacteria and fungi important for cheese ripening (Tilocca et al., 2020). Microbial populations can be introduced with raw milk, starters and traditional equipment. They are then further selected during cheesemaking processing and hygiene practices and might become resident when appropriate niches are found (Alegría et al., 2009; Calasso et al., 2016). Many researchers have studied cheese microbiota diversity and hypothesized a microbial signature exists from different processing environments, named "house microbiota" (Calasso et al., 2016; Gobbetti et al., 2018). However, despite detailed microbial community characterization, most researchers mainly focused on 16S rRNA markers using metagenetics, and only few of them have

combined DNA and metabolomic approaches. Moreover, the taxonomic resolution of metagenetics is uneven and can, for a large number of taxa, discriminate microorganisms up to the genus level, while cheese is constituted of a complex microbiota in which many bacterial, fungal and viral strains might play a key role in microbiota dynamics (Cardin et al., 2022). In this context, a different and more in-depth method to study microbial communities is to use shotgun metagenomics, in which total DNA in a sample is sequenced (Usyk et al., 2023). Many issues of amplicon sequencing, such as introduction of biases due to amplification and inability to study microorganisms such as viruses, can be overcome with shotgun metagenomics since all DNA present in the sample is fragmented and sequenced without previous amplification (Abellan-Schneyder et al., 2021; Maske et al., 2021). Shotgun metagenomics can provide the opportunity to study this microbiota at an unprecedented depth (De Filippis et al., 2021).

Caciotta is a type of Italian cheese made from cow's milk. It is a traditional, short-medium (from 15–20 days to 6 month) ripened cheese characterized by a cylindrical shape (4 to 8 cm high and 8 to 16 cm in diameter), weighs usually 1 kg, has a dark ivory colored rind and pale-yellow core (Bancalari et al., 2020). The production of Caciotta in the northern Italian mountain area is often based on small-scale dairy farms that use traditional methods. Mountain areas such as Alti Pascoli della Lessina and Trento province are well known and recognized for their cheesemaking history, especially Caciotta, Monteveronese and Spressa delle Giudicarie cheeses (Apolito 2018). Cheeses from such mountain areas harbor highly diverse microbiota (Carafa et al., 2019). However, scant information is available using high throughput sequencing approaches in traditional mountain cheese production (Turri et al., 2021).

In this study, we combined shotgun metagenomics and volatilomics to investigate the factors shaping the microbiota and aroma of typical mountain Caciotta cheese.

2. Materials and Methods

2.1 Cheesemaking and sample collection

Overall, 42 Caciotta cheeses were sampled in triplicate from 5 closely located producers (51 ± 26 km range) from Trentino Alto-Adige and Veneto mountain regions belonging to Alti Pascoli della Lessinia, Giudicarie esteriori and Trento province areas (Supplementary Figure 1). Raw milk collected from four milking sessions, which included a full day's milk, as well as the previous evening's milk and the following day's morning milk, was stored at 8-10°C before use (2 days total). Milk was heated to 37°C then lyophilized starter or natural milk culture (*i.e.* through backslopping) was added. After 40 minutes, the milk mass was heated to 39°C and animal rennet was added. The curd was manually broken into rice/corn-sized pieces, cooked for 15 minutes at 43, 44 or 45 °C and placed in perforated molds. Cheeses were pressed 24 hours to drain whey. Molds were removed and cheeses were salted by dry-salt washing for two days or immersed in brine (20% NaCl w/v) for four days. Ripening was performed at 12-14°C and 80% relative humidity for a duration of 60±14 days. Table 1 summarizes the cheesemaking practices employed in mountain Caciotta cheese production.

Table 1. Summary of the cheesemaking practices employed in mountain Caciotta cheese production

Origin	Cooking temperature (°C)	Starter type	Salting method	Ripening days
Producer 1	44	Natural milk culture and lyophyized starter (Lyofast, Sacco, Italy)	Salt washing	64
Producer 2	44	Natural milk culture	Brine	54
Producer 3	43	Lyophilized starter (T1/CD and TB1/B -D, Bioagro, Italy)	Brine	70
Producer 4	45	Lyophilized starter (CO-02, Chr. Hansen, Denmark)	Brine	60
Producer 5	45	Lyophilized starter (TCC-20 and MC-18, Chr. Hansen, Denmark	Brine	52

Cheeses, obtained from milk during the cold (October to May) and warm periods (June to September) of 2020 and 2021, were sampled at the end of the ripening period. Core cheese samples of 2.5 g and 150 mg were employed for volatilomic and metagenomic analyses as described by Penland et al. (2021) and Carraro et al. (2011), respectively. Three biological replicates were obtained for each of the above-mentioned analyses yielding a total of 126 samples.

2.2 Chemical characterization

For chemical analyses, namely pH, salt content, water activity and humidity, a slice of cheese of around 50 g (from crust to core) was homogenized using a knife mill (Retsch Grindomix GM200, Hann, Germany). For pH determination, 2 g of sample were diluted with distilled water (1/10, w/v) and homogenized for 30 sec at 8,000 rpm with a T10 Ultra-Turrax (IKA-Werke, Staufen, Germany). The homogenate was centrifuged for 5 min at 2,000 rpm (Eppendorf 5804, Hamburg, Germany) then paper filtered (Whatman grade 1, Fisher Scientific Italia, Rodano, Italy) and pH was measured on the filtrate using a Hanna HI5221

pH meter (Ronchi di Villafranca Padovana, Italy). Salt content was determined using the Volhard method (AOAC, 1990), while residual moisture was determined gravimetrically by oven-drying a 10 g sample at 103°C until constant weight. Water activity was determined using a dew point hygrometer (AquaLab 4 TEV; Decagon Devices, Pullman, WA, USA). Each measure was done in duplicate.

2.3 Shotgun metagenomics

Total DNA was extracted using the DNeasy PowerSoil kit (Qiagen, Hilden, Germany) following manufacturer's instructions. After extraction, DNA quantity was determined using a Qubit dsDNA HS Assay (Invitrogen, Life Technologies, Italy). Libraries were prepared using the Nextera XT DNA Sample Preparation Kit (Illumina, Inc., San Diego, USA) and IDT for Illumina Nextera DNA UD Indexes. Final libraries were assembled in equimolar pools. Quality was checked with an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) and quantified using a Qubit Assay Kit HS. Subsequently, libraries were sequenced by UC Davis Genome Center (California, US) using a NOVASEQ Sp500 platform (250 bp forward and reverse for ~8 million reads per sample). A total of 1.1 billion reads were generated from the sequencing service. Raw reads were checked for their quality using the FASTQC software (v.0.11.9, Brown et al., 2017). Subsequently, the bioBakery3 platform was used for quality control, contaminant depletion (KneadData), taxonomic (MetaPhlAn 3) assignment and functional (HUMAnN 3) profiling (Beghini et al., 2021). Low quality, repetitive sequence and adapters were removed with KneadData. A quality score cut-off of 35 was used. High quality microbial reads were taxonomically profiled using MetaPhlAn3, an assembly free taxonomic profiler (Segata et al., 2012; Beghini et al., 2021). This computational tool mapped the quality-controlled shotgun reads to a database of unique clade-specific marker genes (read-based profiling) with high discriminatory power, estimating the relative abundance of each microbial clade in the samples with species-level resolution (Segata et al., 2012; Beghini et al., 2021). Bowtie2, a fast DNA aligner, was used by MetaPhlAn3 to map the metagenomic reads against the unique

clade-specific marker genes. MetaPhlAn3 viral profiles were obtained using the *–add_viruses* command to profile potential DNA viruses in the samples.

HUMAnN 3 was applied to assess the functional aspect of genes and pathways in sample metagenomes by utilizing the native UniRef90 annotations from the pangenomes of species analyzed by ChocoPhlAn. Gene abundance was reported in reads per kilobase. This was achieved by computing the sum of the alignments scores (*i.e.* number of matches to the reference gene) for all alignments for a gene family and normalizing it taking into consideration the alignments for a single sequence to multiple reference genes. Alignments that did not meet the e-value, identity and coverage thresholds were not included in the analysis (Beghini et al., 2021).

Strain level diversity of the most abundant bacterial species from cheese and starter (*i.e. Streptococcus thermophilus* and *Lactobacillus delbrueckii*) was investigated with StrainPhlAn 3 and PanPhlAn 3 (Truong et al., 2017; Beghini et al., 2021) which uses the core gene families to generate precise markers for the genetic characterization at the strain level (Beghini et al., 2021). Inferred branch supports from the phylogenetic tree were constructed with 1,000 bootstrap replicates based on parsimony splits as implemented in SplitsTree 4.0 (Huson and Bryant, 2006). Computational performances were optimized with UFBoot2 (Hoang et al., 2018), while graphical representation was elaborate using iTol v6 (Letunic, and Bork, 2021).

All raw sequence data in read-pairs format were deposited in the National Centre for Biotechnology Information (NCBI) in the Sequence Read Archive (SRA) under the project PRJNA922379 and PRJNA922380, for cheese samples and starter cultures, respectively.

2.4 Volatilome analysis

Headspace thermal desorption coupled with gas chromatography (GC) mass spectrometry (MS) was performed using HiSorb probes desorbed with UNITY-xr (both from Markes International, UK) combined with a 5977B GC-MS (Agilent Technologies, US). Cheese was placed in 2.5 ml vials and headspace sampled using an HiSorb Agitator at 40°C and 200 rpm for one hour. After sampling, probes were thermodesorbed using UNITY-xr at 280°C for 12

min and purge flow of 50 ml/min for 1 min. Flow path was set at 200°C with trap low at 25°C and trap high at 290°C. Injection in the GC was performed using a low split 5 ml/min flow. A DB-5ms capillary column 60 m × 250 μm × 0,25 μm (Agilent Technologies, US) was used. The oven temperature program was as follows: initial 40°C held for 2 min, then ramped 3°C/min up to 180°C, and again ramped 20°C/min up to 260°C for 5 min, and finally held for 6 min. The constant flow rate of helium carrier gas was 1 mL/min. The MS analyses were done in a full scan mode (TIC mode), with a scan range of 33 to 350 amu. Molecules were identified using the in-built National Institute of Standards and Technology library (identification criteria >85% ion profile match). Forty-four standard molecules were injected to validate detected peaks (Supplementary Table S1). Semi-quantitative analysis was performed with the MassHunter quantitative analysis workstation (v.11.1, Agilent Technologies, US).

2.5 Statistical analyses

Statistical analyses and visual representations of microbiome taxonomic abundance and functional activity were assessed with MicrobiomeAnalyst (v. 2.0, Dhariwal et al., 2017; Chong et al., 2020), an online tool designed to facilitate the statistical analysis, interactive visualization and meta-analysis of microbiome data. Low abundance taxa were filtered according to mean abundance values and interquartile range of 10%. Alpha-diversity was calculated using Shannon and Chao1 indices. Kruskal–Wallis test was used to compare indices and taxa relative abundance. Beta-diversity was assessed using permutational multivariate analysis of variance (PERMANOVA) to test factor effects on microbial communities, while visual representation was obtained with non-metric multidimensional scaling (NMDS). Significant factors ($p<0.05$) yielding high pseudo F-ratio were evaluated using pairwise comparison. Jaccard and Bray-Curtis distances were evaluated to test compositional dissimilarities between presence/absence or presence and relative abundance of detected species.

Statistical analysis of volatile organic compounds (VOCs) was performed in R version 4.0.5 (R Foundation for Statistical Computing, Vienna, Austria). Peaks from validated molecules

were integrated and normalized using total area normalization. Relative abundance of each compound was compared using Kruskal–Wallis test. For multivariate analysis, integrated peaks were first log transformed and then normalized using total area normalization. Principal component analysis was used for visual representation. Euclidean distances were used in one-way PERMANOVA, while pair-wise comparison was performed only on the significant factor yielding the highest pseudo F-ratio. All *p-values* from multiple comparisons were adjusted with the Benjamini & Hochberg (1995) correction method.

Barplots of functional activity were obtained using the *humann_barplot* function of HUMAnN 3 (Beghini et al., 2021). Correlations between gene abundances and VOCs were calculated with Spearman's rank correlation coefficient (Best and Roberts, 1975). The analysis was limited to strong ($\rho>0.5$ and $\rho<-0.5$) and significant (FDR-corrected *p-values* <0.05) correlations for the three most abundant VOCs for each class of compounds (Benjamini and Hochberg, 1995). From these, the top 2 and bottom 2 correlations were reported. When strong correlations were not available, ρ limits were reduced to 0.15 or -0.3 corresponding to weak and modest correlations.

3. Results

3.1 Physico-chemical composition of Caciotta cheese

The physico-chemical properties of Caciotta cheese - pH, water activity, salt content and humidity - are shown in Table 2.

Table 2. Chemical characterization of Caciotta cheese from mountain areas.

Parameter	Mean±SD	Median	1st quartile	3rd quartile
pH	5.55±0.17	5.52	5.32	5.99
Water activity	0.95±0.01	0.955	0.928	0.969
Salt (g/100 g)	1.99±0.40	2.02	1.44	2.94
Humidity (%)	37.57±6.82	35.6	30.7	50.58

These parameters showed some variability between samples. For example, in Caciotta cheese, optimal acidification is considered to be achieved when the curd pH is between 5.20 and 5.40. In this study, at the end of the ripening period, observed pH values ranged from 5.32 to 5.99 reflecting different acidification kinetics and thus microbial metabolism. Similar technological importance is given to salt content which directly affects water activity and microbial growth. Salt concentration varied between 1.44 and 2.94 g/100 g of Caciotta. Similarly, water activity varied from 0.928 to 0.965 according to the considered cheese. On the other hand, the variation in cheese humidity was in agreement with the "semi-hard cheese" classification and salt concentration. The distribution of pH, water activity values and salt content were further considered to study their effects on bacterial and viral community structure.

3.2. Taxonomical diversity using metagenomics

The sequencing of the cheese sample DNA yielded a total of 1.1 billion reads. After quality filtering 7,087,825 ± 3,722,973 high quality reads per sample were obtained. Two replicates from two distinct samples presented very different compositions in comparison with the

other corresponding replicates and were discarded. Bacterial reads accounted for 53.73 ± 41.84% of total reads while viral reads accounted for 46.27 ± 41.89%. A total of 75 bacterial and 84 viral species were identified using MetaPhlAn. From these, only 57 bacterial and 58 viral species were kept after data filtering in MicrobiomeAnalyst. Bacterial community consisted of 5 phyla, 5 classes, 11 order, 16 families and 26 genera. Among them, the phyla *Bacillota*, class *Bacilli*, order *Lactobacillales*, family *Lactobacillaceae* and genus *Streptococcus, Lactococcus, Lentilactobacillus* were the most represented and diverse (*i.e.* number of associated species). The most abundant bacterial species were *Streptococcus thermophilus* (73.58%) and *Lactobacillus delbrueckii* (21.37%) followed by *Lactococcus lactis* (0.92%), *Leuconostoc pseudomesenteroides* (0.83%), *Lactobacillus helveticus* (0.70%) and *Leuconostoc mesenteroides* (0.59%).

The viral community was composed of 2 phyla, 2 classes, 11 families and 25 genera. Among them, the phyla *Uroviricota*, class *Caudoviricetes*, family "unclassified *Caudoviricetes"* and *Aliceevansviridae* and genus "unclassified *Caudoviricetes", Moineauvirus, Brussowvirus* were the most represented and diverse. The most abundant viral species were Streptococcus virus phiAbc2m (18.26%), Streptococcus phage TP778L (16.95%), Streptococcus virus DT1 (16.82%), Streptococcus virus 7201 (8.15%), Lactobacillus phage A2 (6.22%), Streptococcus virus Sfi21 (5.67%), Lactococcus phage bIL310 (3.71%) and Lactobacillus phage Lrm1 (3.10%).

Considering starter cultures, DNA sequencing yielded a total of 340 M reads from which 8.08 ± 0.08 M high quality reads per sample were obtained. Four bacterial species were identified using MetaPhlAn. After data filtering in MicrobiomeAnalyst, only *S. thermophilus* (85.60%) and *L. delbrueckii* (14.40%) were found to compose cheese starter cultures. Natural milk cultures were characterized by the sole presence of *S. thermophilus.*

3.3 *Factors affecting bacterial and viral beta-diversity*

Multiple categorical variables, based on the distribution of the chemical parameters (Table 2) and origin, sampling season and year, use of lyophilized starter or natural milk culture

as well as curd cooking temperature, were investigated (Supplementary Table S2). PERMANOVA based on Bray-Curtis and Jaccard distances was performed on bacterial and viral cheese communities (Table 3).

Table 3. PERMANOVA of Caciotta cheese bacterial and viral species based on Bray-Curtis and Jaccard distances.

m.o.	Distance	Variable	Pseudo F ratio	*p*-value	Distance	Pseudo F ratio	*p*-value
Bacteria	Bray-Curtis	Origin	19.48	0.001	Jaccard	14.33	0.001
		Cooking Temperature	12.92	0.001		10.82	0.001
		Ripening days	12.32	0.001		9.75	0.001
		Salt	8.22	0.001		6.33	0.001
		Water activity	5.90	0.001		4.94	0.001
		Starter	5.39	0.001		4.74	0.001
		Season	4.61	0.001		4.01	0.001
		Salting	4.56	0.004		4.83	0.001
		pH	3.65	0.001		3.72	0.001
		Year	1.62	0.096		2.11	0.021
Virus	Bray-Curtis	Origin	22.35	0.001	Jaccard	17.38	0.001
		Cooking Temperature	17.97	0.001		13.56	0.001
		Salting	17.01	0.001		15.60	0.001
		Ripening days	15.69	0.001		12.04	0.001
		pH	5.32	0.001		4.74	0.001
		Salt	4.69	0.001		4.75	0.001
		Water activity	3.72	0.001		3.40	0.001
		Season	2.69	0.044		2.25	0.041
		Year	0.83	0.488		1.32	0.201
		Starter	0.66	0.611		3.17	0.006

Many factors shaped the microbial communities of typical Caciotta cheese. Bacterial community composition (*i.e.* presence/absence of bacterial species) was significantly affected by all the tested factors. Among these, origin, cooking temperature and ripening time presented the highest pseudo F-ratio, indicating the largest separation between the investigated groups. On the other hand, salting method, pH, and sampling year had the lowest pseudo F-ratio and the investigated groups were less separated. Similar results were

obtained after the comparison of composition and relative abundance (Bray-Curtis distance) except that the "sampling year" factor did not significantly affect the bacterial community structure.

Viral communities of Caciotta cheese were significantly different in composition and relative abundance according to origin, cooking temperature, salting, ripening days, pH, salt concentration and water activity. The "sampling year" and "starter type" were non-significant factors. Similar results were obtained for compositional data (Jaccard distances), although only sampling year was found to not affect the presence or absence of viral species in Caciotta cheese (p=0.201). As for bacterial communities, origin and cooking temperature presented the highest pseudo F-ratio in the viral communities. These two factors were further investigated through qualitative (NMDS) and quantitative (pairwise PERMANOVA) methods to understand if microbial communities were differently affected within groups.

Qualitative estimation on the effect of origin was performed using NMDS based on Bray-Curtis distances (Figure 1).

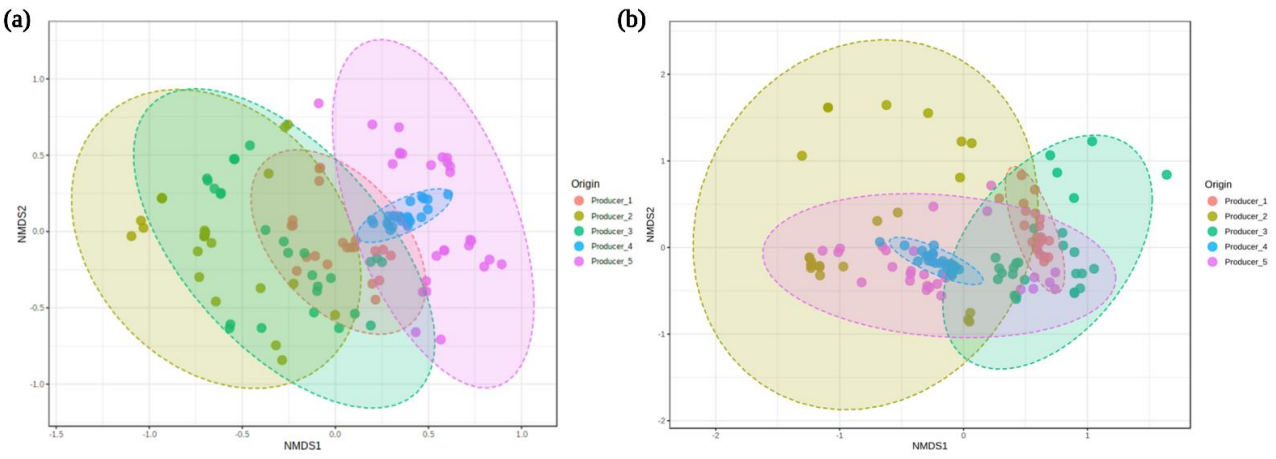


Figure 1. Non-metric multidimensional scaling plot based on Caciotta cheese origin, showing the similarities among cheeses from different producers by investigating (**a**) bacterial (stress value 0.212) and (**b**) viral (stress value 0.125) species using Bray-Curtis distances.

Clusters of cheeses organized according to producer location were observed for both bacterial and viral communities. Viral communities presented narrower clusters for each cheese origin in comparison to the bacterial ones. On the other hand, while bacterial

communities showed a separation between regions of origin (Producer 4 and Producer 5 being in Veneto while the other producers were located in Trentino Alto-Adige), a similar separation could not be observed for the viral communities.

A posteriori test, using pairwise PERMANOVA confirmed that bacterial communities (Supplementary Table S3) centroids (*i.e.* location in multivariate space that represent the average value of all species) were significantly different in terms of composition and relative abundance for producers from different regions. Moreover, no significant differences were observed in the comparison between Producer 1 and, Producer 2 or Producer 3, and between Producer 4 and Producer 5 (Figure 1a). The results of pairwise PERMANOVA using Jaccard distance were analogous to those obtained using the Bray-Curtis dissimilarity. On the other hand, a posteriori test on viral communities (Supplementary Table S3) showed that viral communities were significantly different in relative abundance and composition for each of the investigated origins.

Also, the cooking temperature was shown to affect bacterial and viral communities (Supplementary Figure S2). Significant changes in composition and relative abundance of bacterial and viral communities were observed for the "cooking temperature" parameter (Supplementary Table S4). While all the pairwise comparisons between the 43, 44, and 45°C cooking temperatures resulted in significantly different centroids for viral species communities, the pairwise comparison between 44 and 45°C did not show significance in the bacterial communities.

3.4 *Alpha Diversity*

The effect of origin was further investigated by considering species diversity within the sample (Figure 2). Chao 1 (estimator of total richness) and Shannon (estimator of richness and evenness) indices were used to compare cheese diversity (Figure 2 b & 2d).

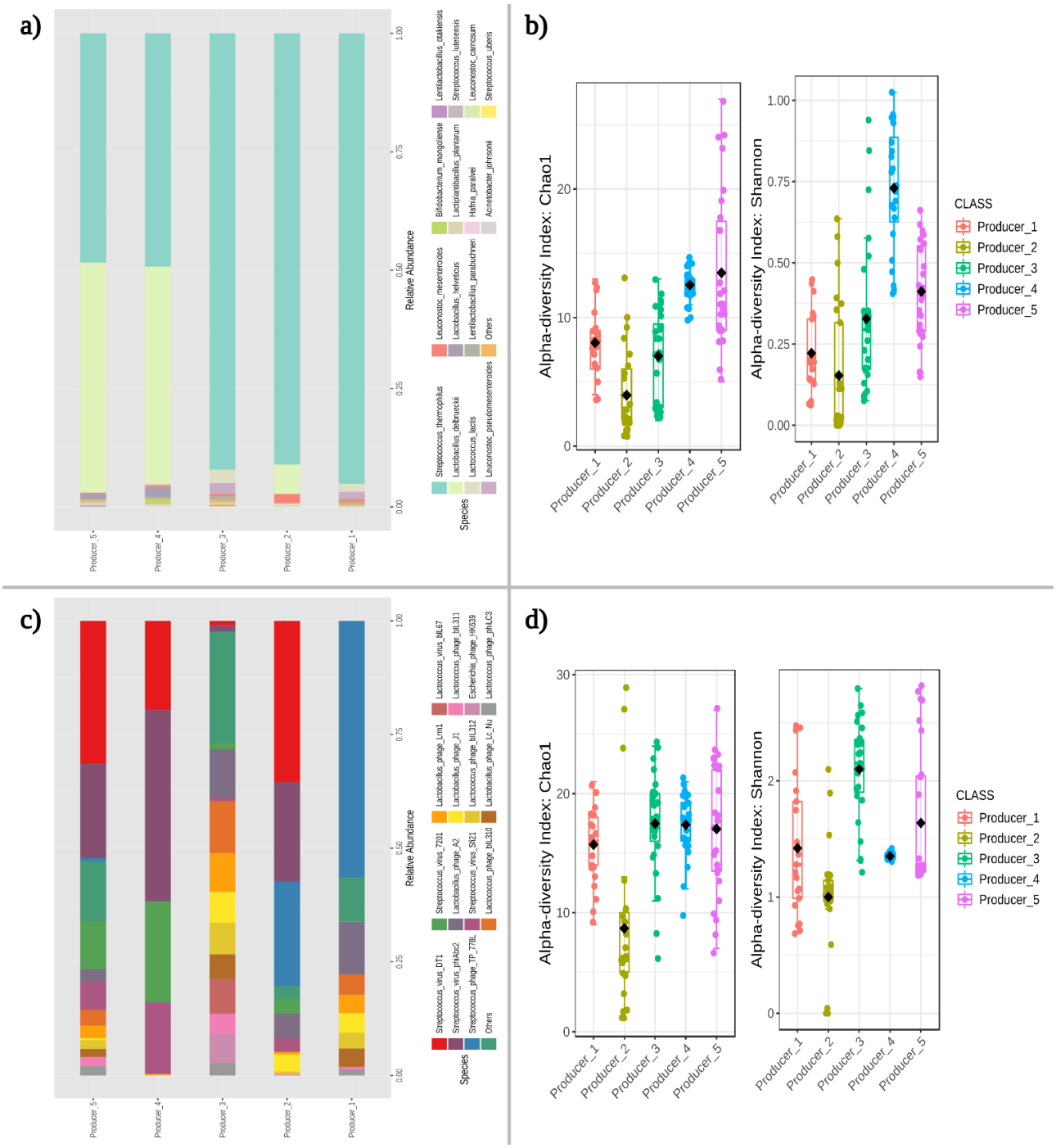


Figure 2. Summary of alpha diversity of typical mountain Caciotta cheese. (**a**) Relative abundance of bacterial communities at species level according to origin; (**b**) Chao 1 and Shannon index of bacterial species; (**c**) relative abundance of viral communities at species level according to origin; (**d**) Chao 1 and Shannon index of viral species.

For both microbial communities, Chao 1 and Shannon indices significantly differed according to cheese origin (Kruskal–Wallis test; *p<0.001)*. Both indices presented higher values for viral communities in comparison to the bacterial ones.

3.5. Taxonomic composition of microbial communities

A total of 45 bacterial and 44 viral species were found to have a significantly different mean relative abundance (Kruskal–Wallis test, $p< 0.05$) according to their origin. Among bacterial species, both starter (*i.e. S. thermophilus* and *L. delbrueckii*) and non-starter lactic acid bacteria (such as *Lactococcus raffinolactis*, *Lactilactobacillus curvatus, Lentilactobacillus parabuchneri, Leuconostoc pseudomesenteroides* and *Lactobacillus brevis*) significantly differed in relative abundance (Supplementary Table S5). Other components of the secondary microbiota (*i.e.* non-inoculated microbiota that develops during the late stages of ripening) were only found for some origins, as for example *Lactiplantibacillus paraplantarum, Lentilactobacillus buchneri* and *Lactococcus piscium.* As starter lactic acid bacteria were the most abundant bacterial species, associated bacteriophages, such as Streptococcus phage TP 778L, Streptococcus virus DT1, Streptococcus virus phiAbc2 and Lactobacillus phage A2, were the most abundant and significantly differed according to Caciotta origin (Supplementary Table S6). Other bacteriophages connected to the infection of species belonging to the *Enterobacteriaceae* family were found to significantly differ as well.

3.6 Intraspecific diversity

Using phylogenetic analysis, we further studied the most abundant species *(S. thermophilus* and *L. delbrueckii)* to characterize their strain diversity. A PhyloPhlAn and the following StrainPhlAn analysis of *S. thermophilus* yielded a total of 160 samples (124 cheese and 36 starters), while in the strain analysis of *L. delbrueckii,* 59 cheese and 24 starter culture samples were discarded due to the low number and/or poor quality of the reconstructed markers, yielding 77 samples. In the analyzed samples, the evolutionary relationship for each of these two species showed complex patterns for mountain Caciotta cheese.

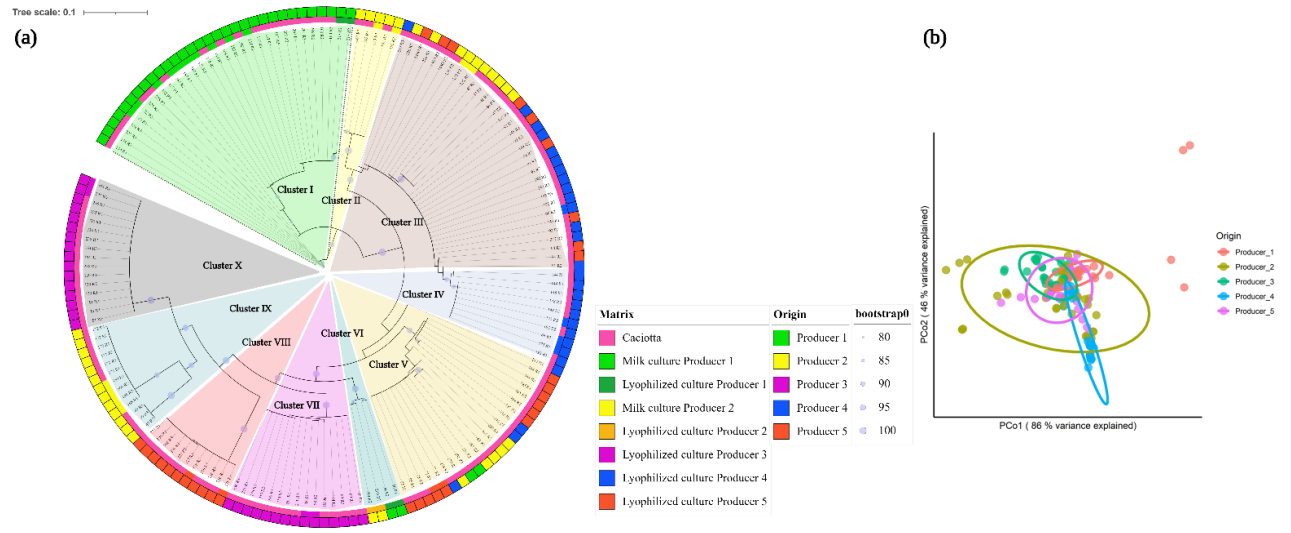


Figure 3. Intraspecific analysis of *Streptococcus thermophilus* in mountain Caciotta cheese. (**a**) Phylogenetic analysis on 160 *Streptococcus thermophilus* strains from mountain Caciotta cheese (124) and starters employed during cheesemaking (36). The inner circle represents the cheese or starter matrix, while the external circle reports the sample origins. Major clusters from I to X were identified through bootstrap values. Phylogenetic analysis was performed using PhyloPhlAn and StrainPhlAn. Bootstrap was assessed with IqTree and the graphical representation of phylogenetic trees was obtained using iTOL. (**b**) Principal coordinate analysis (PCoA) of strain dissimilarities was created in the BioBakery environment using distmat function and Kimura correction method.

In Fig 3a, bootstrap values showed 10 major clusters characterized by different levels of strain diversity. Clusters II, III, IV, VII, VIII, IX and X showed limited diversity within themselves. Notably, strains occurring in mountain Caciotta cheese were clustered together at high bootstrap support with the starter strains from either natural milk or lyophilized starter cultures indicating their establishment and implication in the cheese production process. This was the case for the II, IV, VII, VIII, IX, X clusters but not for cluster I (Fig 3a) in which the starter strains were different from those actually observed in the cheese. Moreover, for these cheeses, the relationship between the cheese and the producer was evident. On the other hand, some samples grouped independently of the employed starter or the cheese origin (Fig 3a, III, V, VI). This phenomenon can be clearly noticed in the strain diversity multivariate analysis (PCoA) (Fig 3b). Cheese and starter strains tended to cluster according to their origin (as shown by the confidence intervals of the ellipses).

The analysis of *L. delbrueckii* strains showed a different pattern. Most of the strain diversity was shared among the samples (Fig 4a, Cluster I) and only a few small groups were differentiated (Fig 4a, Clusters II, III, IV). In a similar way, the multivariate analysis (Fig 4b) showed completely overlapped confidence intervals and few separated samples.

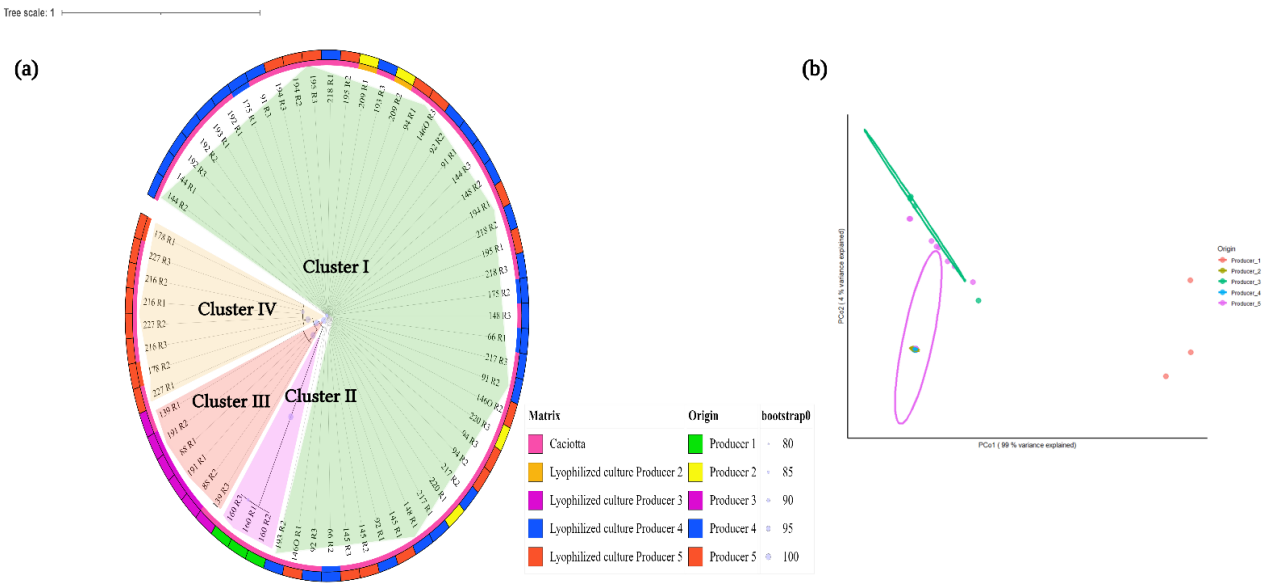


Figure 4. Intraspecific analysis of *Lactobacillus delbrueckii* in mountain Caciotta cheese. (**a**) Phylogenetic analysis on 77 *Lactobacillus delbrueckii* strains from mountain Caciotta cheese (65) and starters employed during the cheesemaking (12). The inner circle represents the cheese or starter matrix, while the external circle reports the sample origins. Major clusters from I to IV were identified through bootstrap values. Phylogenetic analysis was performed using PhyloPhlAn and StrainPhlAn. Bootstrap was assessed with IqTree and the graphical representation of phylogenetic trees was obtained using iTOL. (**b**) Principal coordinate analysis (PCoA) of strain dissimilarities was created in the BioBakery environment using distmat function and Kimura correction method.

3.7 Volatile organic compounds of Caciotta cheese

Mountain Caciotta cheese presented a complex volatilome from which frequent molecules were selected for the following validation. Forty-four VOCs, including 14 alcohols, 4 aldehydes, 8 ketones, 3 carboxylic acid, 2 fatty acids, 6 esters, 4 terpenes and 3 hydrocarbons were validated by injecting pure standards (Supplementary Table S1). The most abundant VOCs in typical mountain Caciotta cheese corresponded to acetic acid (14.45%), butan-2-ol (8.99%), 3-methyl-butan-1-ol (8.49%), butan-2-one (7.69%), ethanol (7.61%) and 3-hydroxybutan-2-one (6.92%). Among the investigated molecules, only the mean values of

3-hydroxybutan-2-one and 1-acetophenyletanone were not significantly different according to Caciotta origin (Kruskal–Wallis test; Supplementary Table S7). For all the other compounds, a significant effect of Caciotta origin was found (Supplementary Table S7). Further PERMANOVA analysis (supplementary Table S8) confirmed that origin significantly affected the VOCs of typical Caciotta cheese. However, all tested factors actually significantly affected the cheese volatile profile. The origin, pH and season had the highest pseudo F-ratio, while the year, curd cooking temperature and starter type showed the lowest values. Pairwise PERMANOVA analysis showed that each producer of typical mountain Caciotta cheese presented a more or less unique VOC profile.

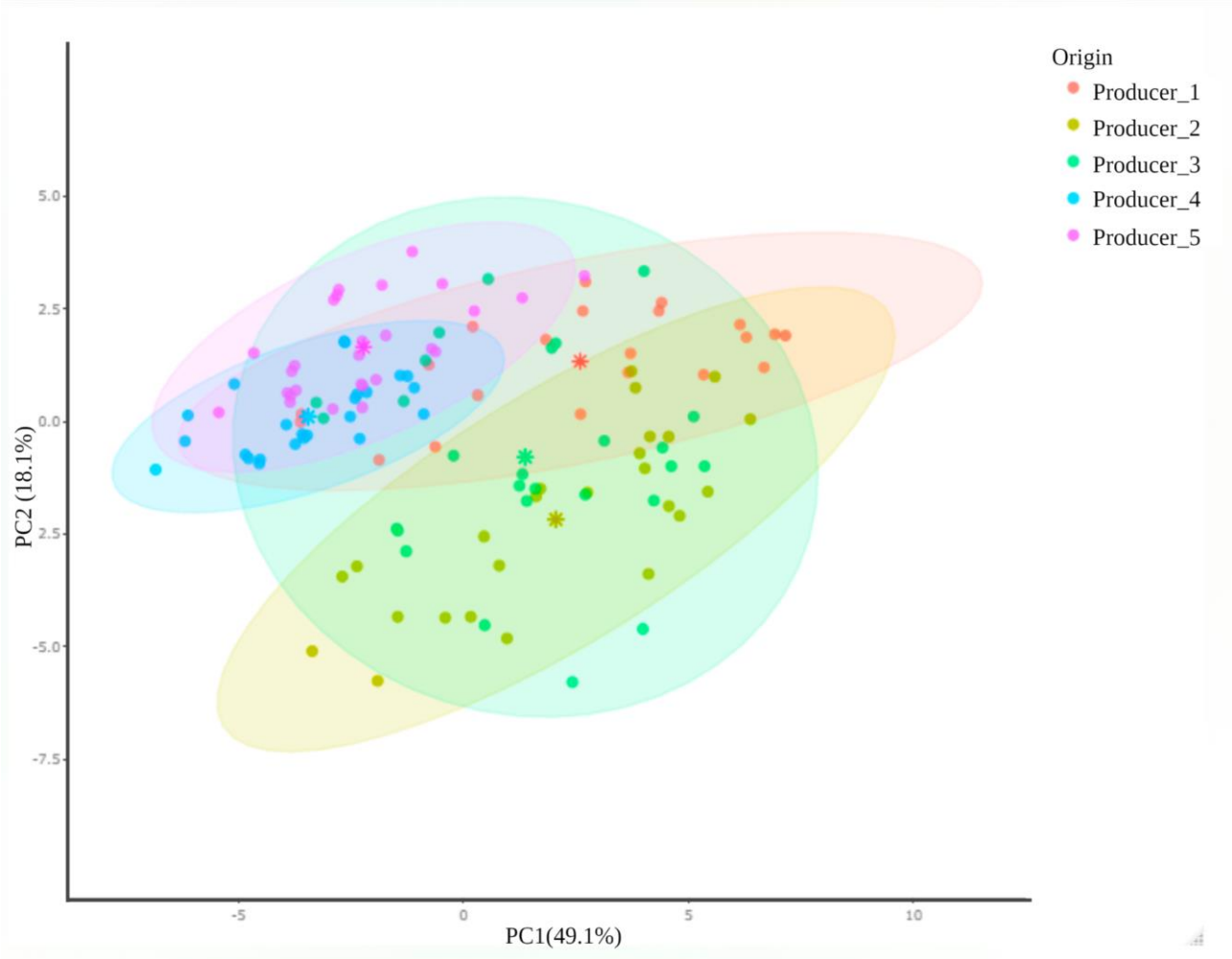


Figure 5. Principal component analysis plot based on Caciotta volatile organic compounds showing the effect of cheese origin. Asterisks (*) represent group centroids.

Finally, as shown in Figure 5 in which a PCA plot of the volatile profiles of Caciotta cheeses according to their origin is represented, cheese volatile profiles from each producer grouped well together, indicating relatively stable volatile profiles over sampling times and between batches, with the exception of those from producer 3. The confidence ellipses of producers 4, 5 and 1 overlapped together indicating quite similar overall volatile profiles while cheese samples from producer 2 showed distinct volatile profiles from those of producers 4 and 5.

3.8 Functional aspects linked to volatilome

From a functional point of view, analysis of the gathered metagenomics data yielded a total of 26,360 genes that presented different associations with Caciotta microbiota. Among them, 1,750 genes were unclassified, 2,462 belonged to *S. thermophilus* and 412 genes were associated with *L. delbrueckii.* The remaining functional annotations were associated with other components of the cheese microbiota. For example, *L. lactis*, *L. pseudomesenteroides* and *L. helveticus* were found to be associated with 551, 415 and 334 genes, respectively. In order to link the obtained metagenomics data to an actual impact of the microbial metabolism, we focused our analysis on enzymes associated with cheese flavor during cheesemaking and ripening and thus linked this to the determined volatilomes (Supplementary Table S10). For this analysis, we created various metabolic classes in which associated EC were pooled. The "proteases and aminopeptidases" class (e.g. serine proteinases, endopeptidases, proline peptidases) was represented by 778 EC (enzyme commission number), while 2987 EC were associated with the "generic amino acid degradation" (*e.g.* keto acid dehydrogenase, acyl kinase) class, 277 EC were associated with "cysteine and methionine degradation" (*e.g.* homocysteine S-methyltransferase, cystathionine beta lyase) class, 126 EC were associated with "citrate fermentation" (*e.g.* citrate lyase, acetolactate synthase) class and 115 EC were associated with "lipid metabolism" (*e.g.* triglycerides esterase, glycerol ester hydrolases). Among them, 530 EC were found to have significantly different relative abundances according to cheese origin (Supplementary Table S11). The most represented classes were those corresponding to "generic amino acid degradation" (344 EC) and "proteases and

aminopeptidases "(111 EC) while "cysteine and methionine metabolism" (47 EC), "citrate metabolism" (18 EC) and "lipid metabolism" (10 EC) were less frequent.

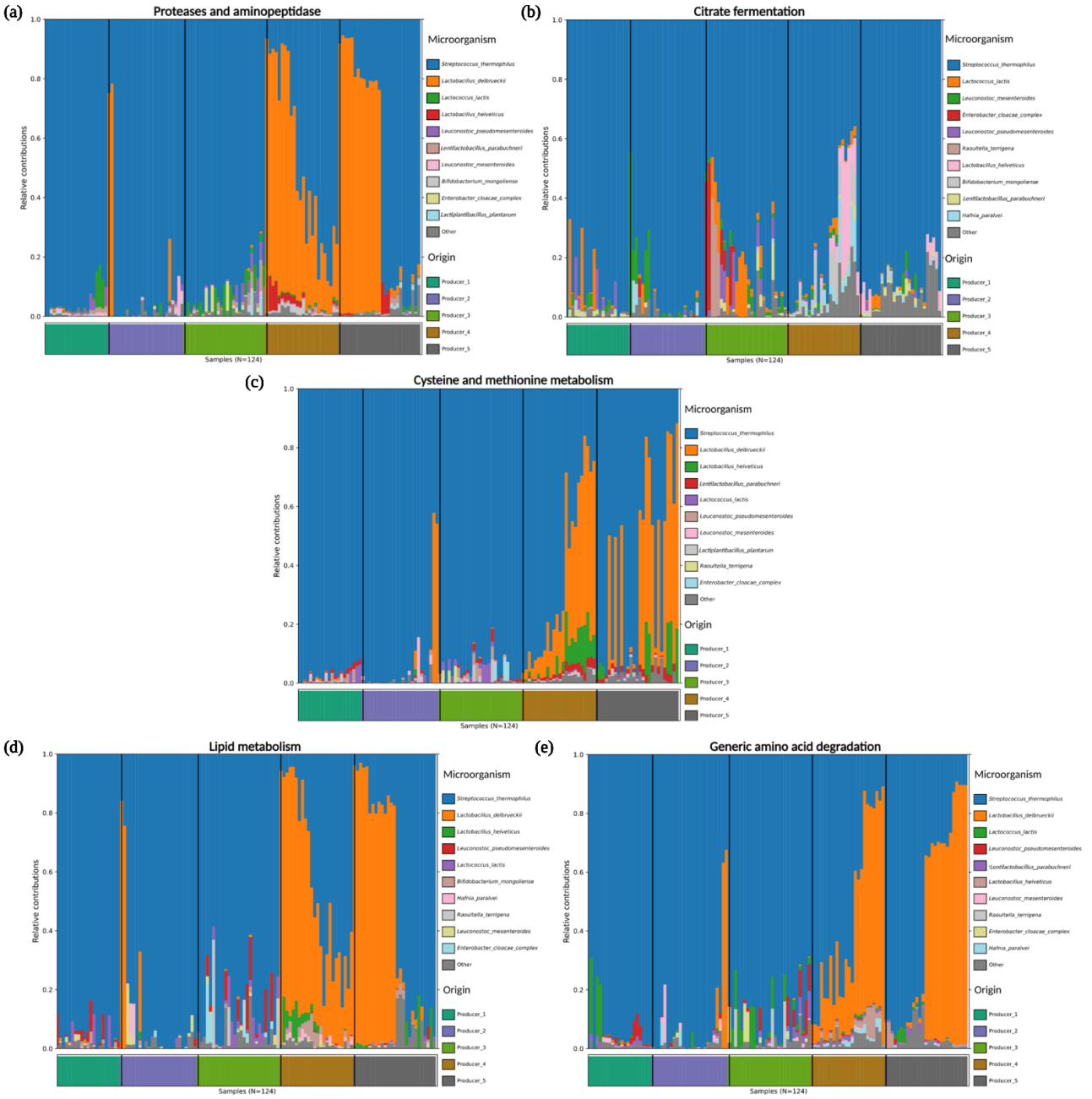


Figure 6. Barplot of gene abundances and microorganism contribution describing functional diversity for the five investigated metabolic classes grouped according to Caciotta origin, namely **a)** "Amino acid metabolism", **b)** "Citrate fermentation", **c)** "Cysteine and methionine metabolism", **d)** "Lipid metabolism" and **e)** "Generic amino acid metabolism". Relative contribution is expressed as the percentage of reads assigned per kilobase (RPK).

We further investigated the functional diversity of typical Caciotta cheese through barplots of genes and associated microorganisms (Fig. 6). Among the investigated classes, it was

possible to notice microbiota functional differences connected to Caciotta origin. Generally, it was observed that *S. thermophilus* exhibited the highest potential for various metabolic activities in typical Caciotta. The "proteases and aminopeptidases" class (Fig. 6a) of producers 4 and 5 were primarily characterized by the strong contribution of *L. delbrueckii* and *L. helveticus*. On the other hand, producers 1, 2, and 3 showed a larger contribution from those belonging to *L. lactis* and *L. pseudomesenteroides*. Similarly, *L. helveticus* strongly contributed to the "citrate fermentation" class (Fig. 6 b) of producer 4 and 5. The highest contribution from *L. lactis* was linked to Producers 1 and 3, while producer 1 and 5 shared a considerable contribution from *L. parabuchneri*. The greatest diversity in the "citrate fermentation" class was observed for producers 3, 4, and 5.

Regarding the "cysteine and methionine metabolism" class (Fig. 6c), *L. delbrueckii* and *L. helveticus* strongly contributed to the functional potential of Caciotta in producers 4 and 5. *L. parabuchneri* showed noticeable relative contributions across most origins. Higher contributions were seen in producers 4 and 5 compared to producers 1 and 3. Producer 2 had a negligible contribution to this class from this microorganism. For the "lipid metabolism" class (Fig. 6d), both *S. thermophilus* and *L. delbrueckii* made similar contributions in producers 4 and 5. Producers 1 and 3 showed high contributions from *L. pseudomesenteroides* and *L. lactis*, respectively, while producer 2 lipid metabolism contribution was predominantly connected to *S. thermophilus*. Similar trends in microorganism functional contributions were observed for the "generic amino acid degradation" class (Fig. 6e). In this class, the highest contributions were observed from *S. thermophilus* and *L. delbrueckii*, with minor contributions from *L. parabuchneri*, *L. lactis*, and *L. helveticus* in producers 4 and 5. Producers 1, 3 and 2 were characterized by the contributions of *L. lactis*, *L. pseudomesenteroides*, and *Leuconostoc mesenteroides*, respectively.

Lastly, we used Spearman's correlation to link bacteria, gene abundances and VOCs (Fig. 7 and Supplementary Table S12).

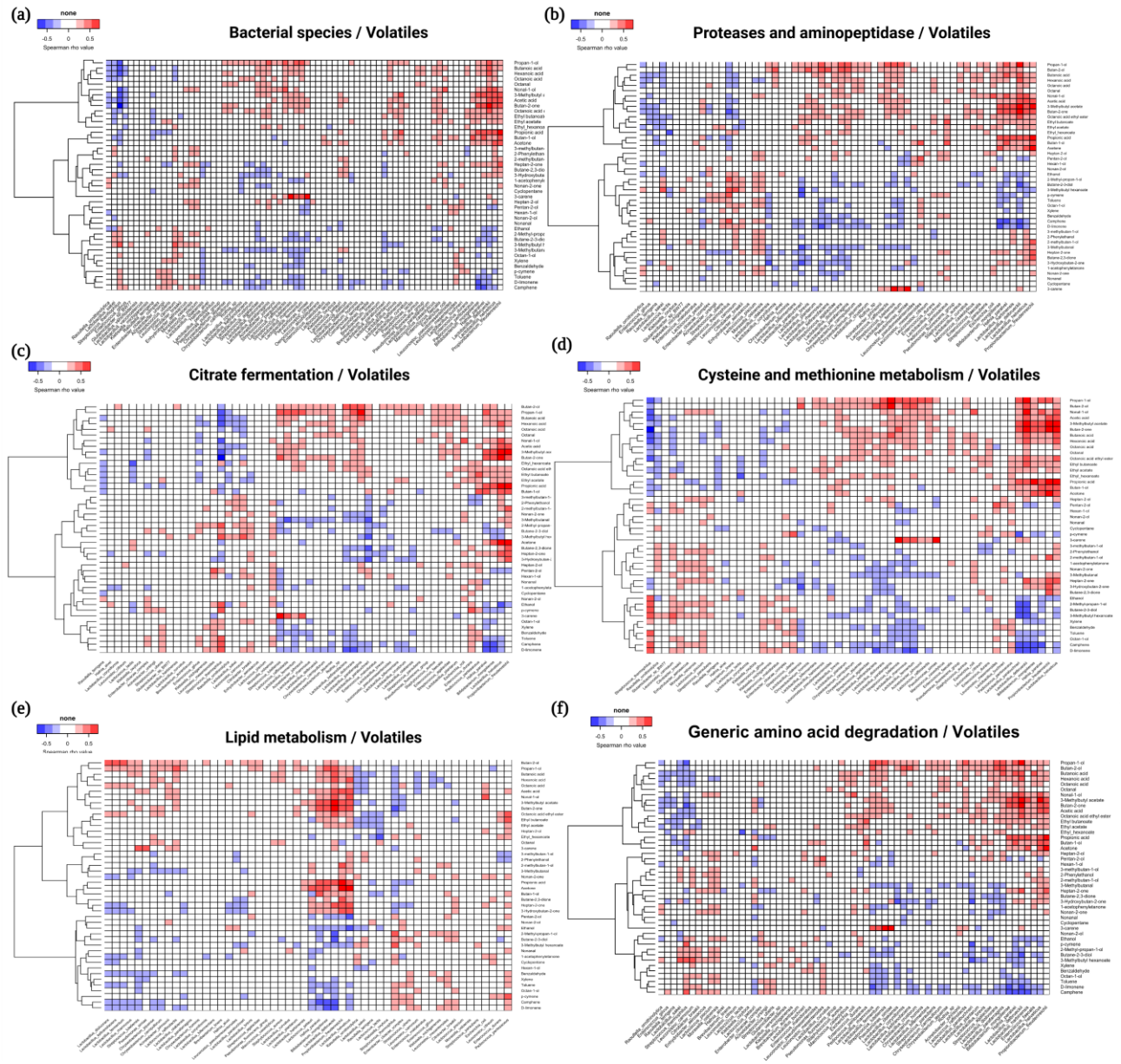


Figure 7. Heatmap of all significant Spearman correlations (FRD <0.05) describing the relationship between volatiles and **a)** bacterial species, **b)** "Amino acid metabolism", **c)** "Citrate fermentation", **d)** "Cysteine and methionine metabolism", **e)** "Lipid metabolism" and **f)** "Generic amino acid metabolism".

The correlation between taxa relative abundance and VOCs (Fig. 7a) highlighted that NSLAB such as *L. paraplantarum, Propionibacterium freudenreichii, L. helveticus* and Enterobacterales, like *Hafnia paralvei,* presented strong correlations with the most abundant VOCs of Caciotta cheese, namely alcohols and ketones. Similarly, the sum of each EC belonging to the defined class was used to further evaluate the metabolic correlation

between microorganisms and VOCs (Fig. 7 b-f). Analogous results were observed for all the investigated classes with the exception of the "lipid metabolism" one.

We obtained a considerable number of significant correlations and decided to focus on the three most abundant VOCs for each of the investigated classes (excluding terpenoids). Overall, acetic acid was significantly correlated with 28 EC, propionic acid with 41 EC, butanoic acid with 34 EC, butan-2-ol with 61 EC, ethanol with 116 EC, 3-methylbutan-1-ol with 69 EC, butan-2-one with 78 EC, 3-hydroxybutan-2-one with 91 EC, butan-2,3-dione with 11 EC, ethyl acetate with 44 EC, ethyl butanoate with 35 EC, ethyl hexanoate with 55 EC, cyclopentane with 56 EC, toluene with 40 EC, xylene with 74 EC, hexanoic acid with 17 EC, octanoic acid with 46 EC, benzaldehyde with 82 EC, nonanal with 68 EC and 3-methylbutanal with 99 EC. We found a total of 76 strong correlations from which the most and least represented classes were "generic amino acid degradation" and "citrate fermentation", with 58 and 1 correlations, respectively. We further focused our attention on the five most abundant alcohols and ketones which are usually associated with typical Caciotta flavor.

Butan-2-ol exhibited negative correlations with EC 2.3.1.189 (keto acid dehydrogenase) and EC 1.1.1.313 (D-hydroxy acid dehydrogenase), while positively correlating with EC 2.3.1.242 and EC 2.3.1.118 (both keto acid dehydrogenase). *P. freudenreichii* and *Pseudomonas fluorescens* were associated with the genes that showed negative correlations, whereas *Enterobacter cloacae, Raoultella ornithinolytica, Hafnia alvei, Lactococcus lactis, Escherichia coli,* and *Klebsiella michiganensis* demonstrated a positive association with butan-2-ol.

In the case of 3-methyl-butan-1-ol, it displayed negative correlations with EC 1.1.1.291 (D-hydroxyacid dehydrogenase) and EC 4.4.1.11 (Cystathionine-Gamma-lyase), while showing positive correlations with EC 4.1.1.59 (acetolactate decarboxylase) and EC 1.1.1.405 (D-hydroxyacid dehydrogenase). Among identified microorganisms, *Acinetobacter johnsonii, L. delbrueckii,* and *Escherichia coli* were associated with the genes showing negative correlations, whereas *L. plantarum* demonstrated a positive association with this compound.

Concerning butan-2-one, it exhibited negative correlations with EC 1.1.1.105 (Alcohol dehydrogenase) and EC 1.1.1.202 (D-hydroxyacid dehydrogenase), while positive correlations were observed with EC 1.1.1.91 and EC 2.3.1.86 (both D-hydroxyacid dehydrogenase). *L. lactis, L. plantarum, L. parabuchneri*, and *Pseudomonas simiae* were associated with the negatively correlated genes, while *Bifidobacterium mongoliense* was positively associated with this VOC.

3-Hydroxybutan-2-one was negatively correlated with EC 4.4.1.11 (cystathionine-gamma-lyase) and positively with EC 1.1.1.301 (D-hydroxyacid dehydrogenase), EC 2.7.2.7 (acylkinase), and EC 1.1.1.62 (D-hydroxyacid dehydrogenase). Genes showing negative correlations belonged to *L. delbrueckii, Escherichia coli and Raoultella terrigena* , while those with positive correlations belonged to *Enterococcus gilvus, Enterococcus faecalis,* and *L. lactis*.

Finally, ethanol was negatively correlated with EC 1.1.1.291 (D-hydroxy acid dehydrogenase) and EC 3.4.24.11 (Metalloproteinases), while showing positive correlations with EC 3.4.22.49 (cysteine proteinase) and EC 2.3.1.109 (keto acid dehydrogenase). Among the associated microorganisms, *Acinetobacter johnsonii and B. mongoliense* harbored genes showing negative correlations, while *Limosilactobacillus fermentum, Pseudomonas simiae and Acinetobacter johnsonii* harbored those with positive correlations.

4. Discussion

In this study, we used "omics" approaches, namely metagenomics and volatilomics, to study bacterial and viral communities, and volatilomes of typical mountain Caciotta cheese. This pinpointed the factors significantly shaping the microbiota and volatile profile of this traditional Italian cheese. As for the metagenomic aspects, the Caciotta bacterial microbiota was similar to previous studies (Calasso et al. 2016). It was mainly dominated by the *Lactobacillaceae* family, *Streptococcus, Lactococcus, Lentilactobacillus* genera and *S. thermophilus*, *L. delbrueckii, L. lactis* and *L. pseudomesenteroides* species. *S. thermophilus* and *L. delbrueckii,* often present in lyophilized cultures used in cheesemaking, or selected during the thermization and incubation process of natural milk cultures, were found to represent nearly 90% of bacterial relative abundance. However, the remaining fraction of bacterial communities showed a more diverse population mainly characterized by non-starter lactic acid bacteria.

Concerning viral communities, to our knowledge, the Caciotta virome has never been explored. To do so, we limited our study to DNA viruses of typical mountain Caciotta cheese without using previous enrichment procedures as described by Walsh et al. (2020) and Yang et al. (2021). We noticed high variability of viral relative abundances possibly connected with phage propagation during cheesemaking. Bacteriophages play a significant role in cheese microbiota, especially regarding their ability to potentially harm the bacteria involved in the fermentation process (Mayo et al., 2021). Indeed, bacteriophages are obligate intracellular parasites that require specific bacterial hosts for their replication. Our findings support previous research that identified the presence and high relative abundance of *Streptococcus, Lactobacillus* and *Lactococcus* phages when these genera were detected in starters and cheese (Walsh et al., 2020; Queiroz et al., 2023).

To investigate how different biotic and abiotic factors shape the microbial communities of raw milk Caciotta cheese, we used PERMANOVA. We found cheese origin had the highest impact on the microbiota and volatilomes of raw milk Caciotta cheese from mountain areas among other tested factors. In particular, while bacterial microbiota significantly differed

among producers from different areas, viromes formed characteristic assemblages for each considered producer. Moreover, viral microbiota appeared more stable during the sampled years compared to the bacterial one. The potential connection between the stability of bacteriophages in cheese and their origin might be attributed to their ubiquitous presence, their selection through cheesemaking practices and their ability to withstand the typical cleaning methods used in cheese production environments (Paillet et al., 2022; Queiroz et al., 2023).

Considering the other factors affecting Caciotta microbiota such as curd cooking temperature, pH, salt concentration and water activity, the observed significant differences could be explained by the variation of the cheesemaking know-how among producers. Cheesemaking induces heat-related, acidic, osmotic, and oxidative stresses on microorganisms and is responsible for modification in heat load, pH, water activity (a_w), and redox potential gradients in the matrix (Beresford et al., 2001). These factors are known to shape bacterial and viral communities in different ways contributing to how specific origin connected communities are shaped (Cardin et al., 2022). This could explain the highest pseudo-F value obtained for the origin factor.

We also characterized strain diversity of typical Caciotta cheese focusing on the most abundant bacterial species, namely *S. thermophilus* and *L. delbrueckii*. Strains occurring in mountain Caciotta cheese were often phylogenetically similar to those from the starter (*i.e.* either natural milk culture or lyophilized starter) indicating their establishment and implication in the cheese production process. Our findings regarding strain diversity showed comparable trends to those observed by Sommerville et al. (2022) for *S. thermophilus* and *L. delbrueckii* in the starter culture and Swiss hard cheese. This indicates a significant presence of strains originating from the used starter, resulting in limited biodiversity as noted in the study by Piquerras et al. (2021). However, we observed a diverse volatile profile of Caciotta cheese suggesting a relevant contribution of the NSLAB population of Caciotta cheese.

Bacterial communities have been reported to establish specific patterns in the cheesemaking environment contributing to typical aromas (Calasso et al., 2016); therefore, we also explored the Cacciotta volatilome as a direct result of microbiota metabolism. Concerning volatile profiles, some of the investigated Caciotta cheese VOCs were previously reported by Bancalari et al. (2020). Similarly, the main alcohols and ketones, such as butan-2-ol, 3-methyl-butan-1-ol, butan-2-one, ethanol and 3-hydroxybutan-2-one, represented the most abundant volatiles in typical Caciotta cheese. These compounds are often associated with odor descriptors such as sweet apricot, fermented, fruity, ethereal, alcoholic and sour milk, which combined with other components of the volatilome create the typical aromatic profile of this mountain cheese. Unfortunately, we could not compare the concentrations of volatiles since the method we used was semi-quantitative. Since all Caciotta producers grazed their herds, we investigated the terpenoids VOC fraction, namely *p*-cymene, D-limonene, camphene and 3-carene (Supplementary Table S7). We found that terpenoid relative abundances differed according to cheese origin. Moreover, some terpenoids such as *p*-cymene and 3-carene were not systematically detected for all origins. These results are in agreement with the study of Turri et al. (2021). They showed that the association of pasture-producers had significantly different terpenoids fractions linked to the cheese origin.

Interestingly, using metagenomics, we had a clear view of both microbial diversity data and metabolic functions. Our study aimed to assess the functional capabilities of Caciotta cheese microbiota, with a specific focus on enzymes influencing flavor. We discovered significant variations in the relative abundances of these enzymes, along with distinct profiles of associated microorganisms linked to the cheese origins. In this context, we specifically analysed genes associated with aromatic compounds synthesis, which were directly related to the most abundant VOCs. We reported the most significant and strong correlations identified between non-homologous isofunctional enzymes and the most abundant VOCs. Among them we found that D-hydroxy acid dehydrogenase, identified as EC 1.1.1.405, EC 1.1.1.291 and EC 1.1.1.62, was positively and negatively correlated with many volatiles. In particular, the enzymes were positively correlated to hexanoic acid, benzaldehyde and

ethanol while negatively correlated to 3-hydroxybutan-2-one, butane-2,3-dione, and ethyl hexanoate. Generally, the class of EC 1.1.1- enzymes are zinc proteins which act on primary or secondary alcohols or hemi-acetals with very broad specificity thus could explain the obtained results. These enzymes were previously described for *L. fermentum* (Hossain 2022). This species was among the most frequently associated microorganisms with these enzymes.

An interesting case is that of EC 3.4.24.11 and ethanol. The enzyme was associated with *B. mongoliense* (metalloproteinases assigned to "amino acid metabolism" class, ρ=-0.5114 *p*-value <0.001). It hydrolyses protein and peptide substrates preferentially on the carboxyl side. It is known that the antiporter/decarboxylase systems constitute indirect proton pumps and play an important role as energy sources (Fernandez and Zuniga, 2006) and this could explain the observed negative correlation.

5. Conclusions

Overall, using a combined microbiota and volatilome omics approach, we performed an in-depth characterization of the typical mountain Caciotta cheese, a raw milk cheese produced in the Italian alpine area.

Many factors significantly affected the microbiota and volatilome structure among the studied biotic and abiotic factors. Cheese origin was a significant driver for bacterial and viral communities as well as for volatilome differences. Overall, viral communities showed higher biodiversity and narrowed sample clusters that could be used for future in depth analysis on cheese origin authenticity. As origin was the major driver for the observed differences, further studies should evaluate the performance of origin authentication by comparing bacterial and viral communities, and volatilome.

Author contributions

Marco Cardin Resources, Investigation, Data Curation, Formal analysis, Visualization, Writing - Original Draft, **Barbara Cardazzo,** Conceptualization, Supervision, Writing - review & editing, **Monika Coton** Methodology, Investigation, Writing - review & editing, **Lisa Carraro** Methodology, Investigation, **Rosaria Lucchini** Resources, Methodology, **Enrico Novelli** Conceptualization, Supervision, Founding acquisition, Writing - review & editing, **Emmanuel Coton** Conceptualization, Supervision, Founding acquisition, Writing - review & editing, **Jérôme Mounier** Methodology, Investigation, Visualization, Methodology, Formal analysis, Writing - review & editing.

Acknowledgements

This work was completed in the framework of Marco Cardin's Ph.D. funded by the Università degli Studi di Padova (Italy, Funding number 20192413451/ISCR_D1) and the Université de Bretagne Occidentale (France). The authors would like to thank Elisabeth Poirier from LUBEM for her support in the volatilome analysis.

References

Abellan-Schneyder, I., Matchado, M. S., Reitmeier, S., Sommer, A., Sewald, Z., Baumbach, J., ... & Neuhaus, K. (2021). Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. Msphere, 6(1), e01202-20.

Alegría, Á., Álvarez-Martín, P., Sacristán, N., Fernández, E., Delgado, S., & Mayo, B. (2009). Diversity and evolution of the microbial populations during manufacture and ripening of Casín, a traditional Spanish, starter-free cheese made from cow's milk. International journal of food microbiology, 136(1), 44-51.

AOAC, C. (1990). Official methods of analysis of the Association of Analytical Chemists International. Official Methods: Washington, DC, USA.

Apolito, A. (2018). Programma Rete Rurale-Nazionale 2014-2020-Rapporto di valutazione delle attività. Available at https://www.reterurale.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/17423 (accessed July 18, 2023)

Bancalari, E., Montanari, C., Levante, A., Alinovi, M., Neviani, E., Gardini, F., & Gatti, M. (2020). *Lactobacillus paracasei* 4341 as adjunct culture to enhance flavor in short ripened Caciotta-type cheese. Food Research International, 135, 109284.

Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., ... & Segata, N. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife, 10, e65088.

Benavides, A. B., Ulcuango, M., Yépez, L., & Tenea, G. N. (2016). Assessment of the in vitro bioactive properties of lactic acid bacteria isolated from native ecological niches of Ecuador. Revista Argentina de microbiologia, 48(3), 236-244.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.

Beresford, T. P., Fitzsimons, N. A., Brennan, N. L., & Cogan, T. M. (2001). Recent advances in cheese microbiology. International dairy journal, 11(4-7), 259-274.

Best, D. J., & Roberts, D. E. (1975). Algorithm AS 89: the upper tail probabilities of Spearman's rho. Journal of the Royal Statistical Society. Series C (Applied Statistics), 24(3), 377-379.

Broadbent, J. R., Cai, H., Larsen, R. L., Hughes, J. E., Welker, D. L., De Carvalho, V. G., ... & Steele, J. L. (2011). Genetic diversity in proteolytic enzymes and amino acid metabolism among *Lactobacillus helveticus* strains. Journal of dairy science, 94(9), 4313-4328.

Brown, J., Pirrung, M., & McCue, L. A. (2017). FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics, 33(19), 3137-3139.

Carafa, I., Stocco, G., Franceschi, P., Summer, A., Tuohy, K. M., Bittante, G., & Franciosi, E. (2019). Evaluation of autochthonous lactic acid bacteria as starter and non-starter cultures for the production of Traditional Mountain cheese. Food Research International, 115, 209-218.

Cardin, M., Cardazzo, B., Mounier, J., Novelli, E., Coton, M., & Coton, E. (2022). Authenticity and Typicity of Traditional Cheeses: A Review on Geographical Origin Authentication Methods. Foods, 11(21), 3379.

Calasso, M., Ercolini, D., Mancini, L., Stellato, G., Minervini, F., Di Cagno, R., ... & Gobbetti, M. (2016). Relationships among house, rind and core microbiotas during manufacture of traditional Italian cheeses at the same dairy plant. *Food Microbiology*, *54*, 115-126.

Carraro, L., Maifreni, M., Bartolomeoli, I., Martino, M. E., Novelli, E., Frigo, F., ... & Cardazzo, B. (2011). Comparison of culture-dependent and-independent methods for bacterial community monitoring during Montasio cheese manufacturing. Research in microbiology, 162(3), 231-239.

Chong, J., Liu, P., Zhou, G., & Xia, J. (2020). Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nature protocols*, *15*(3), 799-821.

Coll-Marqués, J. M., Bäuerl, C., Zúñiga, M., & Pérez-Martínez, G. (2020). Differences in the expression of cell envelope proteinases (CEP) in two *Lactobacillus paracasei* probiotic strains. FEMS Microbiology Letters, 367(13), fnaa102.

Christiansen, J. K., Hughes, J. E., Welker, D. L., Rodríguez, B. T., Steele, J. L., & Broadbent, J. R. (2008). Phenotypic and genotypic analysis of amino acid auxotrophy in *Lactobacillus helveticus* CNRZ 32. Applied and Environmental Microbiology, 74(2), 416-423.

De Filippis, F., Valentino, V., Alvarez-Ordóñez, A., Cotter, P. D., & Ercolini, D. (2021). Environmental microbiome mapping as a strategy to improve quality and safety in the food industry. Current Opinion in Food Science, 38, 168-176.

den Besten, H. M., van der Mark, E. J., Hensen, L., Abee, T., & Zwietering, M. H. (2010). Quantification of the effect of culturing temperature on salt-induced heat resistance of Bacillus species. Applied and environmental microbiology, 76(13), 4286-4292.

Dhariwal, A., Chong, J., Habib, S., King, I. L., Agellon, L. B., & Xia, J. (2017). MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic acids research, 45(W1), W180-W188.

Dobric, N., Bruinenberg, P. G., Limsowtin, G. K. Y., Hillier, A. J., & Davidson, B. E. (1998). Cystathionine (gamma-lyase) from *Lactococcus lactis* subsp. *cremoris* SK11. Australian Journal of Dairy Technology, 53(2), 131.

Fernandez, M., & Zuniga, M. (2006). Amino acid catabolic pathways of lactic acid bacteria. Critical reviews in microbiology, 32(3), 155-183.

Guillot, A., Gitton, C., Anglade, P., & Mistou, M. Y. (2003). Proteomic analysis of *Lactococcus lactis*, a lactic acid bacterium. Proteomics, 3(3), 337-354.

Gobbetti, M., Di Cagno, R., Calasso, M., Neviani, E., Fox, P. F., & De Angelis, M. (2018). Drivers that establish and assembly the lactic acid bacteria biota in cheeses. Trends in Food Science & Technology, 78, 244-254.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. Molecular biology and evolution, 35(2), 518-522.

Holland, R., Liu, S. Q., Crow, V. L., Delabre, M. L., Lubbers, M., Bennett, M., & Norris, G. (2005). Esterases of lactic acid bacteria and cheese flavour: Milk fat hydrolysis, alcoholysis and esterification. International Dairy Journal, 15(6-9), 711-718.

Hossain, T. J. (2022). Functional genomics of the lactic acid bacterium *Limosilactobacillus fermentum* LAB-1: metabolic, probiotic and biotechnological perspectives. Heliyon, 8(11), e11412.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. Molecular biology and evolution, 23(2), 254-267.

Jang, S. Y., Heo, J., Park, M. R., Song, M. H., Kim, J. N., Jo, S. H., ... & Oh, S. (2017). Genome characteristics of *Lactobacillus fermentum* strain JDFM216 for application as probiotic bacteria. Journal of microbiology and biotechnology, 27(7), 1266-1271.

Kamilari, E., Tomazou, M., Antoniades, A., & Tsaltas, D. (2019). High throughput sequencing technologies as a new toolbox for deep analysis, characterization and potentially authentication of protection designation of origin cheeses?. International journal of food science, 2019.

Kamilari, E., Tsaltas, D., Stanton, C., & Ross, R. P. (2022). Metataxonomic mapping of the microbial diversity of Irish and eastern mediterranean cheeses. *Foods*, *11*(16), 2483.

Kieliszek, M., Pobiega, K., Piwowarek, K., & Kot, A. M. (2021). Characteristics of the proteolytic enzymes produced by lactic acid bacteria. Molecules, 26(7), 1858.

Kim, K. H., Chun, B. H., Baek, J. H., Roh, S. W., Lee, S. H., & Jeon, C. O. (2020). Genomic and metabolic features of *Lactobacillus sakei* as revealed by its pan-genome and the metatranscriptome of kimchi fermentation. Food microbiology, 86, 103341.

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic acids research, 49(W1), W293-W296.

Liu, M., Nauta, A., Francke, C., & Siezen, R. J. (2008). Comparative genomics of enzymes in flavor-forming pathways from amino acids in lactic acid bacteria. Applied and environmental microbiology, 74(15), 4590-4600.

Liu, M., Bayjanov, J. R., Renckens, B., Nauta, A., & Siezen, R. J. (2010). The proteolytic system of lactic acid bacteria revisited: a genomic comparison. BMC genomics, 11, 1-15.

Liu, M., Prakash, C., Nauta, A., Siezen, R. J., & Francke, C. (2012). Computational analysis of cysteine and methionine metabolism and its regulation in dairy starter and related bacteria. Journal of bacteriology, 194(13), 3522-3533.

Maske, B. L., de Melo Pereira, G. V., da Silva Vale, A., Souza, D. S. M., Lindner, J. D. D., & Soccol, C. R. (2021). Viruses in fermented foods: are they good or bad? Two sides of the same coin. Food Microbiology, 98, 103794.

Matoba, Y., Noda, M., Yoshida, T., Oda, K., Ezumi, Y., Yasutake, C., ... & Sugiyama, M. (2020). Catalytic specificity of the *Lactobacillus plantarum* cystathionine γ-lyase presumed by the crystallographic analysis. Scientific Reports, 10(1), 1-16.

Mayo, B., Rodríguez, J., Vázquez, L., & Flórez, A. B. (2021). Microbial interactions within the cheese ecosystem and their application to improve quality and safety. Foods, 10(3), 602.

Medina, R. B., Katz, M. B., González, S., & Oliver, G. (2004). Determination of esterolytic and lipolytic activities of lactic acid bacteria. Public Health Microbiology: Methods and Protocols, 465-470.

Paillet, T., Lossouarn, J., Figueroa, C., Midoux, C., Rué, O., Petit, M. A., & Dugat-Bony, E. (2022). Virulent Phages Isolated from a Smear-Ripened Cheese Are Also Detected in Reservoirs of the Cheese Factory. Viruses, 14(8), 1620.

Paxson, H. (2010). Locating value in artisan cheese: reverse engineering terroir for new-world landscapes. American Anthropologist, 112(3), 444-457.

Penland, M., Falentin, H., Parayre, S., Pawtowski, A., Maillard, M. B., Thierry, A., ... & Deutsch, S. M. (2021). Linking Pélardon artisanal goat cheese microbial communities to aroma compounds during cheese-making and ripening. International Journal of Food Microbiology, 345, 109130.

Piqueras, J., Chassard, C., Callon, C., Rifa, E., Theil, S., Lebecque, A., & Delbès, C. (2021). Lactic Starter Dose Shapes S. aureus and STEC O26: H11 Growth, and Bacterial Community Patterns in Raw Milk Uncooked Pressed Cheeses. Microorganisms, 9(5), 1081.

Possas, A., Bonilla-Luque, O. M., & Valero, A. (2021). From cheese-making to consumption: Exploring the microbial safety of cheeses through predictive microbiology models. Foods, 10(2), 355.

Queiroz, L. L., Lacorte, G. A., Isidorio, W. R., Landgraf, M., de Melo Franco, B. D. G., Pinto, U. M., & Hoffmann, C. (2023). High Level of Interaction between phages and bacteria in an artisanal raw milk cheese microbial community. Msystems, 8(1), e00564-22.

Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., & Huttenhower, C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. Nature methods, 9(8), 811-814.

Somerville, V., Berthoud, H., Schmidt, R. S., Bachmann, H. P., Meng, Y. H., Fuchsmann, P., ... & Engel, P. (2022). Functional strain redundancy and persistent phage infection in Swiss hard cheese starter cultures. The ISME Journal, 16(2), 388-399.

Stressler, T., Eisele, T., Schlayer, M., Lutz-Wahl, S., & Fischer, L. (2013). Characterization of the recombinant exopeptidases PepX and PepN from *Lactobacillus helveticus* ATCC 12046 important for food protein hydrolysis. PLoS One, 8(7), e70055.

Tilocca, B., Costanzo, N., Morittu, V. M., Spina, A. A., Soggiu, A., Britti, D., ... & Piras, C. (2020). Milk microbiota: Characterization methods and role in cheese production. Journal of Proteomics, 210, 103534.

Truong, D. T., Tett, A., Pasolli, E., Huttenhower, C., & Segata, N. (2017). Microbial strain-level population structure and genetic diversity from metagenomes. Genome research, 27(4), 626-638.

Turri, F., Cremonesi, P., Battelli, G., Severgnini, M., Brasca, M., Gandini, G., & Pizzi, F. (2021). High biodiversity in a limited mountain area revealed in the traditional production of Historic Rebel cheese by an integrated microbiota–lipidomic approach. Scientific Reports, 11(1), 1-14.

Usyk, M., Peters, B. A., Karthikeyan, S., McDonald, D., Sollecito, C. C., Vazquez-Baeza, Y., ... & Burk, R. D. (2023). Comprehensive evaluation of shotgun metagenomics, amplicon sequencing, and harmonization of these platforms for epidemiological studies. Cell Reports Methods, 3(1).

Walsh, A. M., Macori, G., Kilcawley, K. N., & Cotter, P. D. (2020). Meta-analysis of cheese microbiomes highlights contributions to multiple aspects of quality. Nature Food, 1(8), 500-510.

Yang, C., You, L., Kwok, L. Y., Jin, H., Peng, J., Zhao, Z., & Sun, Z. (2021). Strain-level multiomics analysis reveals significant variation in cheeses from different regions. LWT, 151, 112043.

Supplementary materials - Ecological diversity and associated volatilome of typical mountain Caciotta cheese from Italy

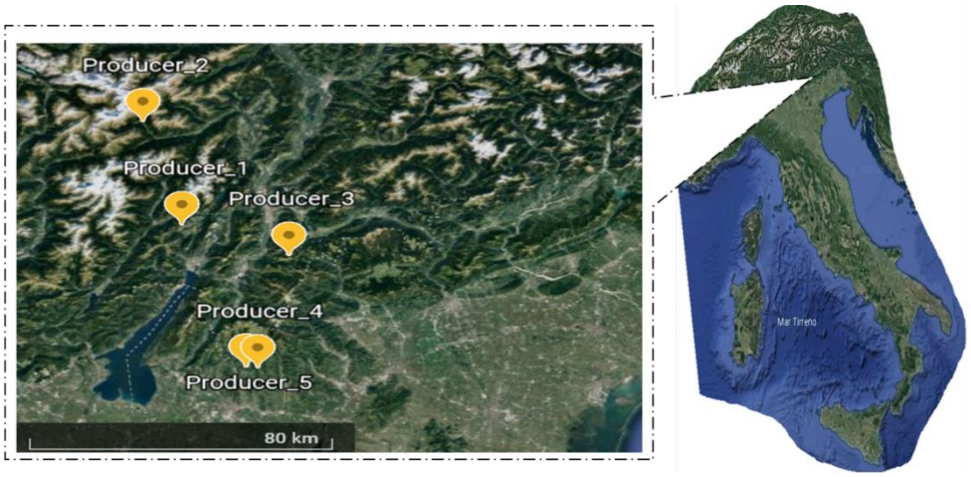
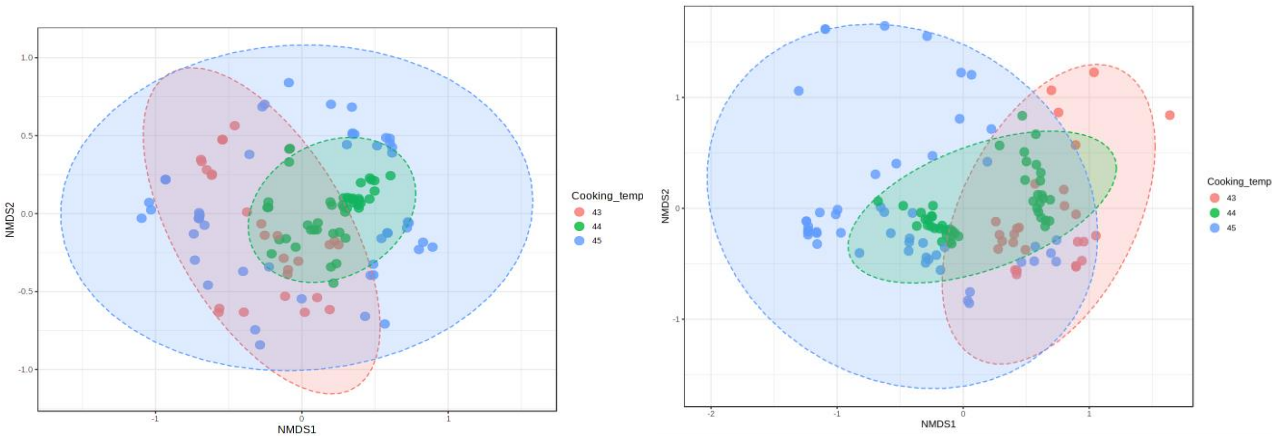


Figure S1. Map of the origin of mountain Caciotta producers. Producer 1 is located in Giudicarie Esteriori area while Producer 2 and Producer 3 are located in Trento province. Producer 4 and Producer 5 are located in Alti Pascoli della Lessinia area.



(**a**) (**b**)

Figure S2. Non-metric multidimensional scaling plot based on curd cooking temperature, showing the similarities among cheeses cooked at different temperatures by investigating (**a**) bacterial and (**b**) viral species using Bray-Curtis distances.

Table S1. Investigated volatile organic compounds in Caciotta cheese

Family	Compound	Formula	MW (g/mol)	Bibliography RT	EIC (m/z)	CAS n°	Odor descriptor *
Alcohols	Ethanol	C_2H_6O	46.07	421-496	31	64-17-5	Alcoholic
	Propan-1-ol	C_3H_8O	60.10	524-572	45	71-23-8	Alcoholic, fermented, fusel
	Butan-1-ol	$C_4H_{10}O$	74.12	660-684	56	71-36-3	Fusel oil, sweet balsam, whiskey
	Butan-2-ol	$C_4H_{10}O$	74.12	570-603	134	78-92-2	Sweet apricot
	2-methyl-propan-1-ol	$C_4H_{10}O$	74.12	594-654	43	78-83-1	Ethereal, winey, cortex
	3-Methylbutan-1-ol	$C_5H_{12}O$	88.15	706-738	55	123-51-3	Fermented, fruity, pungent
	2-Methylbutan-1-ol	$C_5H_{12}O$	88.15	718-740	57	137-32-6	Ethereal, fusel, alcoholic
	Butane-2,3-diol	$C_4H_{10}O2$	90.12	747.5-824	45	513-85-9	Fruity, creamy, buttery
	Pentan-2-ol	$C_5H_{12}O$	88.15	664-704	59	584-02-1	Sweet, herbal, nutty
	Hexan-1-ol	$C_6H_{14}O$	102.17	841-871	69	111-27-3	Green, fruity, pear
	Heptan-2-ol	$C_7H_{16}O$	116.20	877-924	45	543-49-7	Fresh lemon, grass, weet floral
	Octan-1-ol	$C_8H_{18}O$	130.23	1052-1093	56	111-87-5	Waxy, green, orange
	Nonal-1-ol	$C_9H_{20}O$	144.25	1149-1168	56	143-08-8	Fresh clean, fatty, rose
	Nonal-2-ol	$C_9H_{20}O$	144.25	1084-1100	56	628-99-9	Waxy, green, creamy
	2-Phenylethanol	$C_8H_{10}O$	122.16	1080-1150	91	60-12-8	Floral, sweet, rosey

Aldehydes	3-Methylbutanal	$C_5H_{10}O$	86.13	615-665	58	590-86-3	Malty
	Benzaldehyde	C_7H_6O	106.12	921-1005	105	100-52-7	Almond
	Octanal	$C_8H_{16}O$	128.21	977-1036	43	124-13-0	Green, herbal, fresh, fatty
	Nonanal	$C_9H_{18}O$	142.24	1081-1128	82	124-19-6	Fatty, waxy
Carboxylic acids	Acetic acid	$C_2H_4O_2$	60.05	594-648	60	64-19-7	Pungent vinegar
	Propionic acid	$C_3H_6O_2$	74.08	684-743	74	79-09-4	Pungent, acidic, cheesy
	Butanoic acid	$C_4H_8O_2$	88.11	775-830	60	107-92-6	Sharp, acetic, cheese
Esters	Ethyl acetate	$C_4H_8O_2$	88.105	577- 616	70	141-78-6	Ethereal, fruity, sweet
	Ethyl butanoate	$C_6H_{12}O_2$	116.16	778-793	71	105-54-4	Fruity, juicy, pineapple
	3-Methylbutyl acetate	$C_7H_{14}O_2$	130.18	843-884	70	123-92-2	Fruity, green, ripe
	Ethyl hexanoate	$C_8H_{16}O_2$	144.21	976-1011	88	123-66-0	Sweet, fruity, pineapple
	Ethyl octanoate	$C_{10}H_{20}O_2$	172.26	1175-1194	88	106-32-1	Fruity, winey
	3-methylbutyl hexanoate	$C_{11}H_{22}O_2$	186.29	1253-1238	70	2198-61-0	Fruity, banana, apple,
Fatty acids	Hexanoic acid	$C_6H_{12}O_2$	116.16	951-1013	87	142-62-1	Fruity, fatty, sour
	Octanoic acid	$C_8H_{16}O_2$	144.21	1154-1209	73	124-07-2	Oily, rancid, capric
Ketones	Propan-2-one	C_3H_6O	58.08	469-475	43	67-64-1	Solvent, ethereal, apple
	Butan-2-one	C_4H_8O	72.11	543-587	43	78-93-3	Acetone-like, ethereal, fruity
	Butane-2,3-dione	$C_4H_6O_2$	86.09	550-619	86	431-03-8	Buttery
	3-Hydroxybutan-2-one	$C_4H_8O_2$	88.11	662-714	45	513-86-0	Sour milk

	Pentane-2,3-dione	$C_5H_8O_2$	100.12	650-681	43	600-14-6	Pungent, sweet, butter
	Heptan-2-one	$C_7H_{14}O$	114.19	846.9-880	114	110-43-0	Fruity, green banana
	1-phenylethanone	C_8H_8O	120.15	1039-1068	105	98-86-2	Sweet, pungent, hawthorn
	Nonan-2-one	$C_9H_{18}O$	142.24	1066-1093	58	821-55-6	Fresh, sweet, green
Terpens	2,2-dimethyl-3-methylidenebicyclo[2.2.1]heptane	$C_{10}H_{16}$	136.23	941-980	93	79-92-5	Woody, herbal, fir, needle
	3,7,7-trimethylbicyclo[4.1.0]hept-3-ene	$C_{10}H_{16}$	136.23	N.D.	93.1	13466-78-9	Citrus, terpenic, herbal
	1-methyl-4-propan-2-ylbenzene	$C_{10}H_{14}$	134.22	1011-1037.54	119.2	99-87-6	Fresh, citrus, terpene
	(4R)-1-methyl-4-prop-1-en-2-ylcyclohexene	$C_{10}H_{16}$	136.23	N.D.	68	5989-27-5	Citrus, orange, fresh, sweet
Others	Cyclopentane	C_5H_{10}	70.13	553.7-687	42.1	287-92-3	Petroleum
	Toluene	C_7H_8	92.14	755-770	91	108-88-3	Sweet
	1,4-xylene	C_8H_{10}	106.16	847-882	91	106-42-3	NA

MW: molecular weight, RT: retention time. * odor descriptor associated with the compounds according to Thegoodscentscompany database (http://www.thegoodscentscompany.com).

Table S2. Tested factors reporting groups levels and sample population for the tested volatile organic compounds, bacterial and viral communities.

Variable	Levels	Sample population	Variable	Levels	Sample population
Origin	Producer 1	21	Salt (g/100 g)	1.30 to 1.53 g	21
	Producer 2	25		1.54 to 1.90 g	26
	Producer 3	27		1.91 to 2.10 g	33
	Producer 4	24		2.11 to 2.45 g	32
	Producer 5	27		2.46 to 3.06 g	12
pH	5.30 to 5.45	32	Water activity	0.914 to 0.930	18
	5.46 to 5.60	57		0.931 to 0.948	30
	5.61 to 5.81	23		0.949 to 0.960	39
	5.81 to 6.00	12		0.961 to 0.969	37
Season	Cold period	65	Salting metod	Brine	103
	Hot period	59		Salt washing	21
Sampling year	2020	65	Starter	Lyofized	97
	2021	59		Natural milk culture	27
Cooking temperature	43°C	27	Ripening days	44 to 54 days	53
	44°C	45		55 to 72 days	56
	45°C	52		73 to 91 days	15

Table S3. Pairwise PERMANOVA test on microbial beta diversity of typical mountain Caciotta cheese based on cheese origin.

m.o.	Distance method	Group 1	Group 2	Sample size	Permutations	Pseudo F ratio	*p*-value
Bacteria	Bray-Curtis	Producer_1	Producer_2	46	999	1.86	0.238
		Producer_1	Producer_3	48	999	1.77	0.216
		Producer_1	Producer_4	45	999	53.01	0.001
		Producer_1	Producer_5	48	999	25.96	0.001
		Producer_2	Producer_3	49	999	33.23	0.001
		Producer_2	Producer_4	49	999	33.23	0.001
		Producer_2	Producer_5	52	999	19.47	0.001
		Producer_3	Producer_4	51	999	64.58	0.001
		Producer_3	Producer_5	51	999	64.58	0.001
		Producer_4	Producer_5	51	999	0.28	0.603
	Jaccard	Producer_1	Producer_2	46	999	2.11	0.107
		Producer_1	Producer_3	48	999	1.88	0.118
		Producer_1	Producer_4	45	999	39.06	0.001
		Producer_1	Producer_5	48	999	24.30	0.001
		Producer_2	Producer_3	49	999	26.27	0.001
		Producer_2	Producer_4	49	999	26.27	0.001
		Producer_2	Producer_5	52	999	17.66	0.001
		Producer_3	Producer_4	51	999	44.79	0.001
		Producer_3	Producer_5	51	999	44.79	0.001
		Producer_4	Producer_5	51	999	2.44	0.107

Virus	Bray-Curtis	Producer_1	Producer_2	46	999	25.55	0.001
		Producer_1	Producer_3	48	999	20.74	0.001
		Producer_1	Producer_4	45	999	21.48	0.001
		Producer_1	Producer_5	48	999	37.98	0.001
		Producer_2	Producer_3	49	999	19.29	0.001
		Producer_2	Producer_4	49	999	19.29	0.001
		Producer_2	Producer_5	52	999	3.97	0.016
		Producer_3	Producer_4	51	999	66.86	0.001
		Producer_3	Producer_5	51	999	66.86	0.001
		Producer_4	Producer_5	51	999	13.52	0.001
	Jaccard	Producer_1	Producer_2	46	999	19.34	0.001
		Producer_1	Producer_3	48	999	14.66	0.001
		Producer_1	Producer_4	45	999	11.84	0.001
		Producer_1	Producer_5	48	999	28.64	0.001
		Producer_2	Producer_3	49	999	25.52	0.001
		Producer_2	Producer_4	49	999	25.52	0.001
		Producer_2	Producer_5	52	999	4.15	0.010
		Producer_3	Producer_4	51	999	48.74	0.001
		Producer_3	Producer_5	51	999	48.74	0.001
		Producer_4	Producer_5	51	999	18.16	0.001

Table S4. Pairwise PERMANOVA test on microbial beta diversity of typical mountain Caciotta cheese based on curd cooking temperature.

m.o.	Distance method	Group 1	Group 2	Sample size	Permutations	Pseudo F ratio	*p*-value
Bacteria	Bray-Curtis	43°C	44°C	72	999	15.96	0.003
		43°C	45°C	79	999	12.64	0.003
		44°C	45°C	97	999	0.40	0.536
	Jaccard	43°C	44°C	72	999	12.82	0.002
		43°C	45°C	79	999	11.73	0.002
		44°C	45°C	97	999	1.64	0.189
Virus	Bray-Curtis	43°C	44°C	72	999	18.67	0.002
		43°C	45°C	79	999	25.53	0.002
		44°C	45°C	97	999	6.09	0.009
	Jaccard	43°C	44°C	72	999	15.93	0.001
		43°C	45°C	79	999	19.85	0.001
		44°C	45°C	97	999	9.04	0.001

Table S5. Kruskal–Wallis test of bacterial species relative abundance (%) based on typical mountain Caciotta cheese origin.

Species	Producer 1	Producer 2	Producer 3	Producer 4	Producer 5	*p*-value
Streptococcus thermophilus	95.04	90.71	91.76	48.81	47.68	<0.001
Lactobacillus delbrueckii	0.15	6.14	0.08	45.46	47.78	<0.001
Leuconostoc mesenteroides	0.53	1.85	0.43	0.35	0.07	0.158
Hafnia paralvei	0.04	0.33	0.002	0.45	0.04	0.102
Loigolactobacillus coryniformis	0	0.25	0.03	0	0.02	<0.001
Acinetobacter johnsonii	0.01	0.12	0.10	0	0.31	<0.001
Lactococcus raffinolactis	0.01	0.08	0	0	1.37	<0.001
Lactiplantibacillus plantarum	0.16	0.05	0.39	0.09	0.18	0.078
Lactococcus lactis	1.47	0.05	2.77	0.22	0.21	0.029
Latilactobacillus curvatus	0	0.05	0.06	0.01	0.01	0.050
Lentilactobacillus parabuchneri	0.50	0.04	0.91	0.42	0.33	<0.001
Leuconostoc pseudomesenteroides	1.56	0.04	2.26	0.25	0.04	0.010
Enterococcus gilvus	0.01	0.03	0.08	0	0	0.002
Streptococcus parauberis	0	0.03	0	0	0.01	<0.001
Lactobacillus sakei	0	0.01	0.02	0	0	<0.001
Lactococcus piscium	0	0.01	0	0	0	0.013
Citrobacter braakii	0	0.01	0.01	0.00	0.00	0.112
Escherichia coli	0	0.01	0	0	0.02	<0.001
Moraxella osloensis	0	0.004	0	0	0	0.004
Lactobacillus brevis	0.002	0.005	0.02	0.11	0.01	0.048
Raoultella terrigena	0	0.002	0.36	0	0	<0.001
Enterobacter cloacae complex	0	0.002	0.14	0	0.00002	<0.001
Enhydrobacter aerosaccus	0	0.001	0	0	0	0.124
Pediococcus pentosaceus	0.07	0.00	0.00	0.07	0.00	<0.001
Brevibacterium aurantiacum	0	0.0005	0	0	0.0001	0.591
Brevibacterium linens	0.0008	0.0004	0	0	0.00001	0.309
Enterococcus italicus	0	0.00005	0	0.0004	0	<0.001
Bifidobacterium mongoliense	0.18	0	0.05	0.96	0.17	<0.001

Propionibacterium freudenreichii	0.14	0	0	0.85	0	0.017
Leuconostoc citreum	0.11	0	0.01	0	0.02	<0.001
Lentilactobacillus parafarraginis	0.01	0	0	0	0.02	<0.001
Enterococcus durans	0.01	0	0	0	0.002	0.092
Lentilactobacillus buchneri	0.01	0	0	0	0	<0.001
Schleiferilactobacillus harbinensis	0.005	0	0.01	0.04	0.03	0.124
Lacticaseibacillus rhamnosus	0.0003	0	0	0	0.02	0.027
Acinetobacter ursingii	0	0	0.01	0	0	0.027
Chryseobacterium carnipullorum	0	0	0	0	0.01	0.124
Chryseobacterium jejuense	0	0	0	0	0.00	0.027
Enterococcus faecalis	0	0	0	0.003	0.01	0.011
Enterococcus malodoratus	0	0	0.01	0.01	0	0.223
Hafnia alvei	0	0	0.06	0.004	0	0.003
Klebsiella michiganensis	0	0	0.01	0	0	<0.001
Lactiplantibacillus paraplantarum	0	0	0	0.001	0	<0.001
Lactobacillus helveticus	0	0	0	1.89	1.15	<0.001
Lactobacillus kefiranofaciens	0	0	0	0	0.05	<0.001
Lactococcus petauri	0	0	0	0.01	0	<0.001
Lentilactobacillus diolivorans	0	0	0	0	0.04	<0.001
Loigolactobacillus bifermentans	0	0	0	0	0.01	<0.001
Macrococcus caseolyticus	0	0	0	0.0002	0.01	0.027
Pseudomonas fluorescens group	0	0	0	0	0.002	<0.001
Pseudomonas simiae	0	0	0	0	0.02	<0.001
Raoultella ornithinolytica	0	0	0.06	0	0	0.027
Rothia sp	0	0	0	0	0.001	<0.001
Staphylococcus aureus	0	0	0	0	0.01	<0.001
Streptococcus agalactiae	0	0	0	0	0.02	0.027
Streptococcus lutetiensis	0	0	0	0	0.18	<0.001

Table S6. Kruskal–Wallis test of viral species relative abundance (%) based on typical mountain Caciotta cheese origin.

Species	Producer 1	Producer 2	Producer 3	Producer 4	Producer 5	*p*-value
Streptococcus virus DT1	0.01	35.46	0.97	19.54	31.22	<0.001
Streptococcus phage TP 778L	56.37	23.05	0.09	0.04	0.80	<0.001
Streptococcus virus phiAbc2	0	21.79	1.37	42.10	20.63	<0.001
Lactobacillus phage A2	11.59	5.34	10.57	0.23	2.86	<0.001
Lactobacillus phage J1	4.37	3.77	6.79	0.01	0.41	<0.001
Streptococcus virus 7201	0.02	3.22	1.12	22.07	10.27	<0.001
Streptococcus virus Sfi21	0	3.03	0.80	15.56	6.13	<0.001
Lactococcus phage bIL285	0.70	1.11	2.07	0.01	0.57	<0.001
Lactococcus phage bIL312	3.48	0.61	6.89	0.16	1.92	<0.001
Lactobacillus phage Sha1	0.34	0.57	0.25	0.02	2.95	<0.001
Lactobacillus phage Lrm1	3.99	0.54	8.54	0.13	2.75	<0.001
Salmonella phage Fels1	0	0.50	0	0	0.0002	0.591
Lactococcus phage bIL286	2.87	0.13	2.16	0.02	0.75	<0.001
Escherichia phage HK639	0	0.11	6.46	0.0002	0.03	0.002
Lactococcus phage bIL311	0.56	0.11	4.44	0.002	1.92	<0.001
Lactococcus phage bIL309	1.52	0.09	3.61	0.01	0.83	<0.001
Enterobacteria phage HK225	0	0.08	0.20	0	0	0.004
Lactococcus phage BM13	0.02	0.06	0.38	0.01	0.76	<0.001
Streptococcus phage TP J34	0	0.05	0	0	0	<0.001
Lactococcus phage bIL310	4.40	0.05	11.43	0.04	3.46	<0.001
Enterobacteria phage P4	0	0.05	0.15	0	0.00004	0.010
Enterococcus phage phiEf11	0	0.05	0	0.0008	0	0.180
Enterobacteria phage mEp237	0	0.04	0.50	0	0.0004	0.005
Lactococcus phage phiLC3	1.33	0.04	2.68	0.002	2.13	<0.001

Lactococcus phage ul36	0.59	0.03	1.74	0.02	0.42	0.001
Salmonella phage SSU5	0	0.025	0.16	0	0.17	0.003
Lactococcus phage P335 sensu lato	0.45	0.02	3.67	0.01	0.39	<0.001
Lactococcus phage r1t	0.25	0.01	1.34	0.002	0.24	<0.001
Lactococcus phage TP901 1	1.12	0.007	1.04	0.0004	0.98	<0.001
Lactococcus phage BK5 T	1.11	0.01	2.06	0.01	0.97	<0.001
Enterobacterial phage mEp390	0	0.005	0	0	0	0.017
Lactococcus phage Tuc2009	0.74	0.002	0.98	0.0003	0.14	<0.001
Escherichia virus P2	0	0.0009	0	0	0	0.092
Enterobacteria phage mEp460	0	0.00	0.03	0	0.01	0.517
Lactococcus phage 1706	0	0.0004	0	0.001	0.19	0.239
Stx2 converting phage 1717	0	0.0003	0.15	0	0.01	0.783
Salmonella virus Epsilon15	0	0.0003	0.03	0	0.001	0.046
Enterobacteria phage mEp235	0	0.0003	2.36	0	0.11	0.004
Escherichia virus P1	0	0.0002	0	0	0.03	0.114
Escherichia phage D108	0	0	0.02	0	0.04	0.316
Acinetobacter virus 133	0	0	0	0	0.20	0.124
Enterobacteria phage phiP27	0	0	0	0	0.02	<0.001
Lactobacillus prophage Lj771	0	0	0	0	0.15	0.027
Salmonella virus PsP3	0	0	0.12	0	0	0.005
Lactococcus virus bIL67	0.0009	0	7.71	0	0.003	<0.001
Streptococcus virus SPQS1	0	0	0	0	0.01	<0.001
Mycobacterium virus Papyrus	0	0	0	0.001	0	<0.001
Streptococcus virus O1205	0	0	0.003	0	0	0.124
Enterobacteria phage ES18	0	0	0.01	0	0.01	0.256
Lactobacillus phage LF1	0	0	0	0	0.32	<0.001

Lactobacillus phage LcNu	4.03	0	5.29	0.0002	1.75	<0.001
Lactobacillus phage phiGb1	0	0	0.06	0	0	0.124
Lactobacillus phage phig1e	0.09	0	0.01	0	3.16	<0.001
Lactobacillus prophage Lj928	0	0	0	0	0.15	0.005
Microbacterium phage Min1	0.01	0	0	0.0004	0	0.056
Streptococcus phage SMP	0	0	0	0.003	0	<0.001
Lactococcus phage 340	0.03	0	0	0	0.02	<0.001
Musca hytrovirus	0	0	0	0	0.05	0.124

Table S7. Kruskal–Wallis test of investigated volatile organic compound relative abundance (%) based on typical mountain Caciotta cheese origin.

Volatile organic compound	Producer_1	Producer_2	Producer_3	Producer_4	Producer_5	*p*-value
Acetic acid	17.82	8.85	12.74	18.46	14.37	<0.001
Ethanol	11.51	9.62	7.63	5.32	3.96	<0.001
Cyclopentane	8.23	4.70	6.50	4.22	3.53	0.015
Ethyl acetate	7.76	1.52	6.04	3.25	6.76	<0.001
Butanoic acid	7.71	2.15	3.52	4.06	5.66	<0.001
Ethyl butanoate	7.14	2.22	1.95	4.07	3.11	<0.001
Toluene	4.08	3.04	2.60	1.59	0.94	<0.001
Butane-2,3-dione	3.96	3.70	4.47	3.85	0.86	0.028
Pentan-2-ol	3.88	3.02	0.98	2.48	0.77	0.004
Ethyl hexanoate	3.55	8.18	1.46	2.18	2.35	<0.001
3-Hydroxybutan-2-one	3.31	6.16	12.83	9.18	3.14	0.196
Xylene	3.19	2.86	2.18	1.79	0.97	<0.001
Butan-2-one	2.96	0.57	1.88	16.13	16.92	<0.001
3-methyl-butan-1-ol	2.95	20.15	11.49	3.56	4.31	<0.001
1-acetophenyletanone	1.58	0.85	1.28	1.06	0.54	0.083
Hexanoic acid	1.15	0.44	0.94	0.89	1.42	<0.001
Butan-2-ol	1.14	1.70	10.35	7.99	23.78	<0.001
Butan-1-ol	0.91	0.14	0.16	0.87	0.23	<0.001
2-Methyl-propan1-ol	0.78	5.51	3.06	0.36	0.31	<0.001
Propan-1-ol	0.72	0.42	0.64	1.42	3.44	<0.001
Propionic acid	0.71	0.20	0.31	2.32	0.25	<0.001
Benzaldehyde	0.58	0.48	0.40	0.37	0.17	<0.001
Octanoic acid ethyl ester	0.53	0.09	0.20	0.37	0.36	<0.001
Heptan-2-one	0.49	2.09	0.27	0.49	0.27	0.019
2-methylbutan-1-ol	0.45	1.70	1.49	0.72	0.33	<0.001
Heptan-2-ol	0.42	1.24	0.21	0.27	0.22	0.001
Hexan-1-ol	0.40	0.23	0.12	0.11	0.09	<0.001
p-cymene	0.37	0.05	0.01	0	0	<0.001
Nonanal	0.31	0.19	0.19	0.13	0.09	<0.001
Butane-2,3-diol	0.30	1.56	1.03	0.12	0.04	<0.001
Acetone	0.26	0.06	0.96	1.21	0	<0.001
D-limonene	0.20	0.20	0.14	0.08	0.04	<0.001
Octanoic acid	0.15	0.11	0.27	0.09	0.14	0.001
3-Methylbutanal	0.15	0.37	0.22	0.11	0.06	<0.001
Nonan-2-one	0.12	1.34	0.11	0.04	0.04	<0.001
3-Methylbutyl acetate	0.08	0.10	0.05	0.48	0.21	<0.001
Camphene	0.07	0.09	0.05	0.02	0.01	<0.001
Octan-1-ol	0.05	0.05	0.03	0.03	0.02	<0.001
Nonal-1-ol	0.03	0.04	0.01	0.03	0.04	<0.001
Octanal	0.01	0	0.01	0.01	0.01	<0.001
Nonan-2-ol	0.01	0.01	0.01	0.01	0.01	<0.001
3-carene	0.01	0.01	0	0	0.02	0.021
2-Phenylethanol	0	3.79	1.21	0.26	0.21	<0.001

3-Methylbutyl hexanoate	0	0.20	0.002	0	0	<0.001

Table S8. PERMANOVA of Caciotta cheese volatile organic compounds.

Variable	**Pseudo F-ratio**	**p-value**
Origin	13.43	0.001
pH	7.08	0.001
Season	7.06	0.001
Salt	6.87	0.001
Water activity	6.23	0.001
Salting	5.69	0.002
Ripening days	3.97	0.001
Year	3.27	0.011
Cooking temperature	2.59	0.025
Starter	2.57	0.017

Table S9. Pairwise PERMANOVA test on volatile organic compounds of typical mountain Caciotta cheese based on origin factor.

Distance method	Group 1	Group 2	Sample size	Permutations	Pseudo F-ratio	p-value
Euclidean	Producer_1	Producer_2	46	999	5.98	0.002
	Producer_1	Producer_3	48	999	5.41	0.001
	Producer_1	Producer_4	45	999	23.06	0.001
	Producer_1	Producer_5	48	999	15.13	0.001
	Producer_2	Producer_3	49	999	24.24	0.001
	Producer_2	Producer_4	49	999	24.24	0.001
	Producer_2	Producer_5	52	999	21.08	0.001
	Producer_3	Producer_4	51	999	19.15	0.001
	Producer_3	Producer_5	51	999	19.15	0.001
	Producer_4	Producer_5	51	999	19.15	0.001

Table S10. Selection of enzyme commision number for focused functional analysis.

Metabolism	Enzyme type	Enzyme commision number (EC)	Reference
Proteases and amminopeptidase	Cell-wall bound proteinase	3.4.21.96	Liu et al., 2010; Coll-Marqués et al., 2020
	Serine Proteinases	3.4.21.-	Kieliszek et al., 2021
	Cysteine Proteinases	3.4.22.-	Kieliszek et al., 2021
	Aspartyl proteinases	3.4.23.-	Kieliszek et al., 2021
	Metalloproteinases	3.4.24.-	Kieliszek et al., 2021
	Threonine peptidase	3.4.25.-	Kieliszek et al., 2021
	Aminopeptidase	3.4.11.2	Liu et al., 2010; Broadbent et al., 2011; Stressler et al., 2013
	Unique aminopeptidases	3.4.11.18	Liu et al., 2010; Broadbent et al., 2011
	Unique aminopeptidases	3.4.11.7	Liu et al., 2010; Broadbent et al., 2011
	Unique aminopeptidases	3.4.19.3	Liu et al., 2010; Broadbent et al., 2011
	endopeptidase	3.4.24.-	Liu et al., 2010; Broadbent et al., 2011
	dipeptidase	3.4.13.–	Liu et al., 2010; Broadbent et al., 2011
	tripeptidase	3.4.11.4	Liu et al., 2010; Broadbent et al., 2011
	Proline peptidase	3.4.14.11	Liu et al., 2010; Broadbent et al., 2011; Stressler et al., 2013
	Proline peptidase	3.4.11.5	Liu et al., 2010; Broadbent et al., 2011
	Proline peptidase	3.4.11.5	Liu et al., 2010; Broadbent et al., 2011
	Proline peptidase	3.4.11.5	Liu et al., 2010; Broadbent et al., 2011
	Proline peptidase	3.4.11.9	Liu et al., 2010; Broadbent et al., 2011
	Proline peptidase	3.4.11.9	Liu et al., 2010; Broadbent et al., 2011
Generic aminoacid degradation	Branched-chain aminotransferase	2.6.1.42	Liu et al., 2008
	Aromatic aminotransferase	2.6.1.57	Liu et al., 2010
	Aspartate aminotransferase	2.6.1.1	Guillot et al., 2003
	Glutamate dehydrogenase	1.4.1.2	Liu et al., 2010
	Keto acid decarboxylase	2.3.1.16	Liu et al., 2010
	Alcohol dehydrogenase	1.1.1.1	Guillot et al., 2003
	Aldehyde dehydrogenase	1.1.1.169	Liu et al., 2010
	Keto acid dehydrogenase complex	1.8.1.4	Guillot et al., 2003
	Keto acid dehydrogenase complex	1.2.4.4	Fernandez and Zuniga, 2006

	Keto acid dehydrogenase complex	2.3.1.-	Guillot et al., 2003
	Branched Chain phosphotransacylase	2.3.1.19	Christiansen et al., 2008
	Acylkinase	2.7.2.-	Guillot et al., 2003
	D-hydroxyacid dehydrogenase	1.1.1.-	Hossain et al., 2022
	L-hydroxyacid dehydrogenase	1.1.1.-	Hossain et al., 2022
	Esterase A	3.1.1.1	Liu et al., 2010
	Serine acteyltranferase	2.3.1.30	Liu et al., 2008
	O-acetylserine sulfhydrolase	2.5.1.47	Christiansen et al., 2008
	Cystathionine beta lyase	4.4.1.8	den Besten et al., 2010
	Cystathionine Gamma lyase	4.4.1.1	Dobric et al., 1988
	Cystathionine beta synthase	4.2.1.22	Matoba et al., 2020
	Cystathionine gamma synthase	2.5.1.48	Benavides et al., 2016
Cysteine and methionine degradation	Homocysteine S-methyltransferase	2.1.1.10	Jang et al., 2017
	Homocysteine methyltransferase	2.1.1.14	Christiansen et al., 2008
	Homoserine O-succinyltransferase	2.3.1.46	Christiansen et al., 2008
	O-acetylhomoserine sulfhydrolase	2.5.1.49	Liu et al., 2012
	Citrate lyase	4.1.3.6	Guillot et al., 2003
Citrate fermentation	Acetolactate decarboxylase	4.1.1.5	Hossain et al., 2022
	Acetolactate synthase	2.2.1.6	Hossain et al., 2022
	Oxaloacetate decarboxylase	4.1.1.112	Kim et al., 2020
Lipids metabolism	Triglycerides esterase	3.1.1.2	Holland et al., 2005
	Glycerol ester hydrolases	3.1.1.3	Medina et al., 2004

Table S11. Kruskal–Wallis test of investigated genes abundance (expressed as logarithm of reads assigned per kilobase) of typical mountain Caciotta cheese origin. Only significantly different genes are reported.

Gene	Function	Metabolism	Associated m.o.	Producer 1	Producer 2	Producer 3	Producer 4	Producer 5	p-values
1.1.1.1	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,57	-1,6	-1,62	-1,75	-1,81	<0.001
1.1.1.1	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,57	-1,60	-1,62	-1,75	-1,81	<0.001
1.1.1.1	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,60	-1,52	-1,66	-1,94	-2,02	<0.001
1.1.1.1	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,60	-1,52	-1,66	-1,94	-2,02	<0.001
1.1.1.1	Alcohol dehydrogenase	Generic amino acid degradation	*Hafnia paralvei*	-4,59	-3,86	-5,91	-3,59	-4,74	0,001
1.1.1.100	Alcohol dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,06	-2,46	-4,62	-1,61	-1,54	<0.001
1.1.1.100	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,81	-1,72	-1,80	-1,92	-2,07	<0.001

1.1.1.100	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,81	-1,72	-1,80	-1,92	-2,07	<0.001
1.1.1.100	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,24	-2,70	-2,42	-2,77	-2,45	<0.001
1.1.1.100	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,24	-2,70	-2,42	-2,77	-2,45	<0.001
1.1.1.100	Alcohol dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,15	0,00	-4,63	-3,42	-4,04	<0.001
1.1.1.100	Alcohol dehydrogenase	Generic amino acid degradation	*Lactobacillus helveticus*	0,00	0,00	0,00	-2,81	-3,02	<0.001
1.1.1.100	Alcohol dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,09	-4,50	-3,05	-4,10	-3,94	0,003
1.1.1.100	Alcohol dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,18	-4,99	-3,14	-4,21	-4,46	<0.001
1.1.1.130	Alcohol dehydrogenase	Generic amino acid degradation	*Hafnia paralvei*	-5,00	-4,09	-5,53	-3,80	-4,96	0,002204
1.1.1.133	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,33	-2,38	-2,18	-2,23	-2,48	<0.001
1.1.1.133	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,33	-2,38	-2,18	-2,23	-2,48	<0.001
1.1.1.133	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,89	-1,91	-2,00	-2,35	-2,20	<0.001
1.1.1.133	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,89	-1,91	-2,00	-2,35	-2,20	<0.001
1.1.1.133	Alcohol dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,44	-5,26	-3,47	-4,52	-3,05	0,003
1.1.1.133	Alcohol dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,31	-5,44	-3,30	-4,35	-4,90	0,002704
1.1.1.169	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,72	-1,56	-1,67	-1,80	-1,78	<0.001
1.1.1.169	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,72	-1,56	-1,67	-1,80	-1,78	<0.001
1.1.1.169	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,75	-2,07	-1,88	-2,11	-2,30	<0.001
1.1.1.169	Alcohol dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,75	-2,07	-1,88	-2,11	-2,30	<0.001
1.1.1.169	Alcohol dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,47	0,00	-5,02	-3,65	-4,41	<0.001
1.1.1.169	Alcohol dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,55	-4,52	-2,97	0,00	-4,69	<0.001
1.1.1.169	Alcohol dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,29	-4,29	-3,12	-4,06	-4,76	<0.001

1.1.1.17	Alcohol dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	0,00	-3,28	-5,70	-2,41	-2,29	<0.001
1.1.1.17	Alcohol dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,38	-4,87	-3,16	-4,29	-4,14	<0.001
1.1.1.179	Alcohol dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,60	0,00	-4,92	-3,70	-4,48	<0.001
1.1.1.188	Alcohol dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,48	0,00	-4,79	-3,81	-4,46	<0.001
1.1.1.193	Alcohol dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,51	-5,41	-3,57	-4,51	-4,48	0,004
1.1.1.205	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,11	-2,12	-2,23	-2,21	-2,37	<0.001
1.1.1.205	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,11	-2,12	-2,23	-2,21	-2,37	<0.001
1.1.1.205	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,96	-1,94	-1,91	-2,20	-2,18	<0.001
1.1.1.205	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,96	-1,94	-1,91	-2,20	-2,18	<0.001
1.1.1.205	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,17	-3,20	-5,40	-2,48	-2,47	<0.001
1.1.1.205	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,17	0,00	-4,66	-3,47	-4,17	<0.001
1.1.1.205	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,30	-4,84	-3,16	-4,12	-4,21	<0.001
1.1.1.205	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,67	-5,45	-3,44	-4,47	-4,87	<0.001
1.1.1.218	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,88	-1,86	-2,05	-2,35	-2,27	<0.001
1.1.1.218	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,88	-1,86	-2,05	-2,35	-2,27	<0.001
1.1.1.218	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,30	-2,47	-2,14	-2,27	-2,37	<0.001
1.1.1.218	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,30	-2,47	-2,14	-2,27	-2,37	<0.001
1.1.1.218	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,64	-2,84	-4,99	-2,04	-1,89	<0.001
1.1.1.22	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,07	-2,07	-2,62	-2,03	-2,53	<0.001
1.1.1.22	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,42	-5,53	-3,16	-3,59	-2,71	<0.001
1.1.1.23	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,51	0,00	-4,87	-3,69	-4,43	<0.001

1.1.1.23	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,39	-4,89	-3,23	-4,56	-4,59	<0.001
1.1.1.23	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,60	0,00	-3,54	-4,76	-5,81	<0.001
1.1.1.23	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,80	-1,82	-1,86	-2,08	-2,20	<0.001
1.1.1.25	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,47	-4,96	-3,20	-4,44	-4,30	<0.001
1.1.1.25	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,58	-4,44	-3,42	-4,33	-4,89	<0.001
1.1.1.25	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,57	-2,27	-2,40	-2,10	-2,60	<0.001
1.1.1.25	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,57	-2,27	-2,40	-2,10	-2,60	<0.001
1.1.1.25	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,85	-1,93	-1,91	-2,81	-2,15	<0.001
1.1.1.25	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,85	-1,93	-1,91	-2,81	-2,15	<0.001
1.1.1.26	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,49	0,00	-5,16	-3,65	-4,66	<0.001
1.1.1.27	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,60	-2,79	-4,68	-2,02	-1,88	<0.001
1.1.1.27	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-2,91	-4,45	-2,60	-3,68	-3,64	<0.001
1.1.1.27	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,51	-5,14	-3,41	-4,36	-4,96	<0.001
1.1.1.27	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,60	-2,49	-2,51	-2,68	-2,73	<0.001
1.1.1.27	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,60	-2,49	-2,51	-2,68	-2,73	<0.001
1.1.1.27	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,79	-1,84	-1,86	-2,05	-2,09	<0.001
1.1.1.27	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,79	-1,84	-1,86	-2,05	-2,09	<0.001
1.1.1.274	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,68	-2,92	-4,73	-2,14	-2,02	<0.001
1.1.1.274	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,71	-4,44	-2,69	-3,96	-4,69	<0.001
1.1.1.276	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,27	0,00	-4,93	-3,50	-4,37	<0.001
1.1.1.276	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,14	-1,84	-2,50	-5,13	-2,59	<0.001

1.1.1.28	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,08	-2,31	-4,36	-1,49	-1,41	<0.001
1.1.1.28	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,12	-4,87	-2,88	-4,04	-4,72	<0.001
1.1.1.290	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,58	-3,74	-6,20	-2,91	-2,89	<0.001
1.1.1.290	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,94	-1,95	-1,94	-2,11	-2,16	<0.001
1.1.1.290	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,94	-1,95	-1,94	-2,11	-2,16	<0.001
1.1.1.290	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,24	-2,31	-2,34	-2,60	-2,55	<0.001
1.1.1.290	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,24	-2,31	-2,34	-2,60	-2,55	<0.001
1.1.1.292	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,53	-1,58	-1,57	-1,82	-1,82	<0.001
1.1.1.3	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,35	0,00	-4,90	-3,64	-4,38	<0.001
1.1.1.3	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Hafnia paralvei*	-5,16	-4,02	-5,53	-3,70	-4,88	<0.001
1.1.1.3	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,82	-3,08	-5,10	-2,17	-2,10	<0.001
1.1.1.3	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,70	-5,47	-3,50	-4,97	-5,13	<0.001
1.1.1.3	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,15	-2,15	-2,18	-2,35	-2,41	<0.001
1.1.1.3	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,15	-2,15	-2,18	-2,35	-2,41	<0.001
1.1.1.3	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,88	-1,90	-1,91	-2,10	-2,14	<0.001
1.1.1.3	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,88	-1,90	-1,91	-2,10	-2,14	<0.001
1.1.1.31	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,66	-4,77	-3,22	-4,20	-4,36	<0.001
1.1.1.317	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,00	0,00	-4,37	-3,19	-3,94	<0.001
1.1.1.329	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,55	-4,81	-3,43	-4,54	-5,28	<0.001
1.1.1.346	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,34	0,00	-4,89	-3,64	-4,30	<0.001
1.1.1.36	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,40	-5,08	-3,35	-4,48	-4,63	<0.001

1.1.1.36	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,81	-1,72	-1,80	-1,92	-2,07	<0.001
1.1.1.36	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,81	-1,72	-1,80	-1,92	-2,07	<0.001
1.1.1.36	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,24	-2,70	-2,42	-2,77	-2,45	<0.001
1.1.1.36	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,24	-2,70	-2,42	-2,77	-2,45	<0.001
1.1.1.363	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,55	-4,93	-3,35	-4,30	-5,32	<0.001
1.1.1.37	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,05	-4,66	-3,08	-4,25	-4,20	<0.001
1.1.1.4	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,43	0,00	-5,01	-3,73	-4,42	<0.001
1.1.1.4	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,21	-4,83	-2,96	-4,01	-4,15	<0.001
1.1.1.4	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,72	-1,68	-1,66	-2,28	-1,99	<0.001
1.1.1.4	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,72	-1,68	-1,66	-2,28	-1,99	<0.001
1.1.1.4	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,97	-2,04	-2,24	-1,95	-2,28	<0.001
1.1.1.4	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,97	-2,04	-2,24	-1,95	-2,28	<0.001
1.1.1.42	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,32	-2,36	-2,37	-2,54	-2,63	<0.001
1.1.1.42	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,32	-2,36	-2,37	-2,54	-2,63	<0.001
1.1.1.42	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,23	-2,15	-2,11	-2,35	-2,98	<0.001
1.1.1.42	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,23	-2,15	-2,11	-2,35	-2,98	<0.001
1.1.1.44	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-3,92	0,00	-4,58	-3,13	-3,87	<0.001
1.1.1.44	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,63	-2,84	-4,85	-2,06	-1,93	<0.001
1.1.1.44	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,29	-4,77	-3,10	-4,02	-4,06	<0.001
1.1.1.44	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,55	0,00	-3,35	-4,41	-5,57	<0.001
1.1.1.47	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,02	-3,98	-2,91	-3,64	-4,38	<0.001

1.1.1.49	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,28	0,00	-5,00	-3,54	-4,36	<0.001
1.1.1.49	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,67	-2,90	-5,17	-2,09	-2,02	<0.001
1.1.1.49	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,10	-4,84	-3,30	-4,40	-4,47	<0.001
1.1.1.60	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,58	-4,96	-3,33	-4,39	-4,74	<0.001
1.1.1.76	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,83	-3,60	0,00	-2,42	-2,53	<0.001
1.1.1.77	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Hafnia paralvei*	-4,91	-3,99	-5,61	-3,68	-4,78	<0.001
1.1.1.85	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,05	-5,72	-3,58	-4,83	-4,71	<0.001
1.1.1.85	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,64	-5,07	-3,51	-4,54	-5,20	<0.001
1.1.1.85	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,99	-1,83	-1,87	-2,09	-2,14	<0.001
1.1.1.85	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,99	-1,83	-1,87	-2,09	-2,14	<0.001
1.1.1.85	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,15	-2,55	-2,56	-2,77	-2,72	<0.001
1.1.1.85	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,15	-2,55	-2,56	-2,77	-2,72	<0.001
1.1.1.86	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,34	0,00	-4,98	-3,60	-4,31	<0.001
1.1.1.86	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,25	-4,85	-3,06	-4,28	-4,16	<0.001
1.1.1.86	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,00	-2,03	-2,04	-2,23	-2,27	<0.001
1.1.1.86	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,00	-2,03	-2,04	-2,23	-2,27	<0.001
1.1.1.86	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,00	-2,03	-2,05	-2,23	-2,27	<0.001
1.1.1.86	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,00	-2,03	-2,05	-2,23	-2,27	<0.001
1.1.1.88	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,73	-3,11	-5,59	-2,29	-2,24	<0.001
1.1.1.88	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,50	-5,28	-3,25	-4,34	-4,33	<0.001
1.1.1.88	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,59	-5,43	-3,42	-4,26	-4,90	<0.001

1.1.1.88	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,96	-1,96	-1,99	-2,31	-2,24	<0.001
1.1.1.88	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,96	-1,96	-1,99	-2,31	-2,24	<0.001
1.1.1.88	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,20	-2,26	-2,30	-2,33	-2,49	<0.001
1.1.1.88	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,20	-2,26	-2,30	-2,33	-2,49	<0.001
1.1.1.94	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,39	0,00	-4,70	-3,65	-4,42	<0.001
1.1.1.94	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,40	-2,63	-4,71	-1,79	-1,68	<0.001
1.1.1.94	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,19	-5,24	-3,26	-4,36	-4,55	<0.001
1.1.1.94	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,64	-4,69	-3,43	-4,49	-4,78	<0.001
1.1.1.94	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,98	-1,98	-2,00	-2,17	-2,22	<0.001
1.1.1.94	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,98	-1,98	-2,00	-2,17	-2,22	<0.001
1.1.1.94	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,00	-2,04	-2,06	-2,23	-2,26	<0.001
1.1.1.94	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,00	-2,04	-2,06	-2,23	-2,26	<0.001
1.1.1.95	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,21	0,00	-4,50	-3,41	-4,17	<0.001
1.1.1.95	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,58	-2,93	-5,06	-2,07	-2,02	<0.001
1.1.1.95	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,24	-5,07	-3,17	-4,12	-4,19	<0.001
1.1.1.95	D-hydroxyacid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,13	-5,12	-2,94	-3,97	-4,92	<0.001
1.8.1.4	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,94	-1,93	-1,98	-2,24	-2,24	<0.001
1.8.1.4	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,94	-1,93	-1,98	-2,24	-2,24	<0.001
1.8.1.4	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,21	-2,21	-2,27	-2,46	-2,49	<0.001
1.8.1.4	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,21	-2,21	-2,27	-2,46	-2,49	<0.001
1.8.1.4	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,54	0,00	-4,80	-3,67	-4,43	<0.001

1.8.1.4	Keto acid dehydrogenase	Generic amino acid degradation	*Hafnia paralvei*	-4,93	-3,97	-5,46	-3,70	-4,73	0,001
1.8.1.4	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,32	-4,77	-3,10	-4,13	-4,10	<0.001
1.8.1.4	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,64	-4,97	-3,52	-4,60	-5,41	<0.001
2.1.1.10	Homocysteine S-methyltransferase	Cysteine methionine degradation	*Lactobacillus delbrueckii*	-4,67	-3,77	-5,96	-2,99	-2,79	<0.001
2.1.1.14	Homocysteine methyltransferase	Cysteine methionine degradation	*Bifidobacterium mongoliense*	-4,43	0,00	-5,01	-3,74	-4,49	<0.001
2.1.1.14	Homocysteine methyltransferase	Cysteine methionine degradation	*Lactobacillus delbrueckii*	-4,91	-2,73	-4,97	-1,98	-1,86	<0.001
2.1.1.14	Homocysteine methyltransferase	Cysteine methionine degradation	*Lactobacillus helveticus*	0,00	0,00	0,00	-2,88	-3,19	<0.001
2.1.1.14	Homocysteine methyltransferase	Cysteine methionine degradation	*Lactococcus lactis*	-3,38	-4,89	-3,01	-4,24	-4,17	<0.001
2.1.1.14	Homocysteine methyltransferase	Cysteine methionine degradation	*Leuconostoc pseudomesenteroides*	-3,67	-4,82	-3,55	-4,33	-4,95	<0.001
2.1.1.14	Homocysteine methyltransferase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,10	-0,95	-1,11	-1,18	-1,52	<0.001
2.1.1.14	Homocysteine methyltransferase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,95	-1,99	-2,00	-2,01	-2,13	<0.001
2.2.1.6	Acetolactate synthase	Citrate fermentation	*Bifidobacterium mongoliense*	-3,90	0,00	-4,34	-3,09	-3,84	<0.001
2.2.1.6	Acetolactate synthase	Citrate fermentation	*Lactococcus lactis*	-2,94	-4,50	-2,83	-3,86	-3,81	<0.001
2.2.1.6	Acetolactate synthase	Citrate fermentation	*Leuconostoc pseudomesenteroides*	-3,61	-5,57	-3,48	-4,49	-4,94	<0.001
2.2.1.6	Acetolactate synthase	Citrate fermentation	*Streptococcus thermophilus*	-1,97	-1,97	-2,00	-2,19	-2,23	<0.001
2.2.1.6	Acetolactate synthase	Citrate fermentation	*Streptococcus thermophilus*	-1,97	-1,97	-2,00	-2,19	-2,23	<0.001
2.2.1.6	Acetolactate synthase	Citrate fermentation	*Streptococcus thermophilus*	-2,12	-2,16	-2,16	-2,34	-2,37	<0.001
2.2.1.6	Acetolactate synthase	Citrate fermentation	*Streptococcus thermophilus*	-2,12	-2,16	-2,16	-2,34	-2,37	<0.001
2.3.1.1	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,89	-1,85	-1,89	-2,18	-2,16	<0.001
2.3.1.1	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,89	-1,85	-1,89	-2,18	-2,16	<0.001
2.3.1.1	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,07	-2,11	-2,20	-2,32	<0.001

2.3.1.1	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,07	-2,11	-2,20	-2,32	<0.001
2.3.1.1	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,48	-4,98	-3,26	-4,58	-4,40	<0.001
2.3.1.1	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,45	-5,02	-3,31	-4,45	-5,30	<0.001
2.3.1.117	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,54	0,00	-4,92	-3,75	-4,45	<0.001
2.3.1.117	Keto acid dehydrogenase	Generic amino acid degradation	*Hafnia paralvei*	-5,38	-4,25	-5,43	-4,02	-5,04	0,009
2.3.1.128	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,16	0,00	-4,75	-3,39	-4,18	<0.001
2.3.1.128	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,43	-4,94	-3,34	-4,26	-4,25	0,001
2.3.1.157	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,02	-2,05	-2,04	-2,23	-2,21	<0.001
2.3.1.157	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,02	-2,05	-2,04	-2,23	-2,21	<0.001
2.3.1.157	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,08	-2,07	-2,11	-2,37	-2,46	<0.001
2.3.1.157	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,08	-2,07	-2,11	-2,37	-2,46	<0.001
2.3.1.157	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,70	-2,94	-4,99	-2,17	-2,04	<0.001
2.3.1.157	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,39	-5,51	-3,48	-4,65	-4,44	0,003804
2.3.1.157	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,61	-4,46	-3,57	-4,32	-5,09	<0.001
2.3.1.16	Keto acid dehydrogenase	Generic amino acid degradation	*Hafnia paralvei*	-4,68	-3,75	0,00	-3,44	-4,68	<0.001
2.3.1.179	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,63	-3,10	-4,99	-2,26	-2,18	<0.001
2.3.1.179	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,04	-2,07	-2,24	-2,29	<0.001
2.3.1.179	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,04	-2,07	-2,24	-2,29	<0.001
2.3.1.179	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,04	-2,07	-2,24	-2,29	<0.001
2.3.1.179	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,04	-2,07	-2,24	-2,29	<0.001
2.3.1.179	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,26	-4,82	-3,13	-4,00	-4,13	<0.001

2.3.1.179	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,56	-5,02	-3,41	-4,34	-4,78	<0.001
2.3.1.180	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-5,04	-3,38	-5,86	-2,51	-2,50	<0.001
2.3.1.180	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,17	-5,05	-3,14	-4,94	-4,09	<0.001
2.3.1.180	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,74	-3,06	-5,00	-2,24	-2,15	<0.001
2.3.1.180	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,22	-2,22	-2,21	-2,45	-2,48	<0.001
2.3.1.180	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,22	-2,22	-2,21	-2,45	-2,48	<0.001
2.3.1.180	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,96	-1,97	-2,03	-2,20	-2,22	<0.001
2.3.1.180	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,96	-1,97	-2,03	-2,20	-2,22	<0.001
2.3.1.180	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,90	-5,06	-3,28	-4,12	-4,38	<0.001
2.3.1.180	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,30	-4,93	-3,12	-4,22	-4,79	<0.001
2.3.1.181	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,57	0,00	-4,91	-3,70	-4,51	<0.001
2.3.1.183	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,44	0,00	-4,77	-3,68	-4,43	<0.001
2.3.1.234	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,14	-2,33	-2,22	-2,62	-2,52	<0.001
2.3.1.234	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,14	-2,33	-2,22	-2,62	-2,52	<0.001
2.3.1.234	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,90	-1,84	-1,95	-2,00	-2,09	<0.001
2.3.1.234	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,90	-1,84	-1,95	-2,00	-2,09	<0.001
2.3.1.234	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,70	-2,97	-5,24	-2,15	-2,06	<0.001
2.3.1.234	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,54	0,00	-5,30	-3,73	-4,54	<0.001
2.3.1.234	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,41	-5,24	-3,42	-4,79	-4,40	0,002
2.3.1.234	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,58	-4,34	-3,42	-4,05	-5,10	0,001
2.3.1.266	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,05	-2,06	-2,07	-2,26	-2,27	<0.001

2.3.1.266	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,05	-2,06	-2,07	-2,26	-2,27	<0.001
2.3.1.266	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,90	-1,91	-1,93	-2,11	-2,14	<0.001
2.3.1.266	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,90	-1,91	-1,93	-2,11	-2,14	<0.001
2.3.1.266	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,72	-5,66	-3,66	-4,70	-4,81	<0.001
2.3.1.266	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,52	-4,22	-3,31	-4,38	-4,87	<0.001
2.3.1.274	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,72	-3,01	-4,92	-2,18	-2,13	<0.001
2.3.1.274	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,84	-1,90	-1,96	-2,04	-2,20	<0.001
2.3.1.274	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,84	-1,90	-1,96	-2,04	-2,20	<0.001
2.3.1.274	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,16	-2,11	-2,07	-2,40	-2,27	<0.001
2.3.1.274	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,16	-2,11	-2,07	-2,40	-2,27	<0.001
2.3.1.274	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,55	-4,77	-3,26	-4,30	-4,34	<0.001
2.3.1.275	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,92	-1,94	-1,95	-2,19	-2,18	<0.001
2.3.1.275	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,92	-1,94	-1,95	-2,19	-2,18	<0.001
2.3.1.275	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,42	-2,44	-2,40	-2,64	-2,63	<0.001
2.3.1.275	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,42	-2,44	-2,40	-2,64	-2,63	<0.001
2.3.1.275	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,41	-5,81	-3,18	-4,30	-4,19	<0.001
2.3.1.30	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,64	-1,53	-1,56	-1,86	-1,82	<0.001
2.3.1.30	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,64	-1,53	-1,56	-1,86	-1,82	<0.001
2.3.1.30	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,93	-2,00	-2,08	-2,14	-2,19	<0.001
2.3.1.30	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,93	-2,00	-2,08	-2,14	-2,19	<0.001
2.3.1.30	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-3,28	-2,57	-4,50	-1,74	-1,64	<0.001

2.3.1.30	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus helveticus*	0,00	0,00	0,00	-3,26	-3,58	<0.001
2.3.1.30	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,00	-5,11	-3,26	-4,38	-4,42	<0.001
2.3.1.31	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,12	-2,10	-2,11	-2,36	-2,36	<0.001
2.3.1.31	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,12	-2,10	-2,11	-2,36	-2,36	<0.001
2.3.1.31	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,09	-2,06	-2,07	-2,32	-2,35	<0.001
2.3.1.31	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,09	-2,06	-2,07	-2,32	-2,35	<0.001
2.3.1.31	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,79	-5,04	-3,39	-4,57	-4,51	<0.001
2.3.1.31	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,64	-6,15	-3,62	-4,44	-4,92	<0.001
2.3.1.35	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,89	-1,85	-1,89	-2,18	-2,16	<0.001
2.3.1.35	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,89	-1,85	-1,89	-2,18	-2,16	<0.001
2.3.1.35	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,07	-2,11	-2,20	-2,32	<0.001
2.3.1.35	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,07	-2,11	-2,20	-2,32	<0.001
2.3.1.35	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,48	-4,98	-3,26	-4,58	-4,40	<0.001
2.3.1.35	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,45	-5,02	-3,31	-4,45	-5,30	<0.001
2.3.1.39	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,23	-2,60	-2,60	-2,74	-2,34	<0.001
2.3.1.39	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,23	-2,60	-2,60	-2,74	-2,34	<0.001
2.3.1.39	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,88	-1,77	-1,81	-1,99	-2,20	<0.001
2.3.1.39	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,88	-1,77	-1,81	-1,99	-2,20	<0.001
2.3.1.39	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,77	-3,10	-5,15	-2,24	-2,16	<0.001
2.3.1.39	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,51	-4,27	-3,38	-4,29	-4,86	<0.001
2.3.1.39	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,38	-4,88	-3,14	-4,38	-4,13	<0.001

2.3.1.41	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,43	0,00	-5,01	-3,69	-4,48	<0.001
2.3.1.46	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	0,00	-3,34	-4,74	-2,47	-2,38	<0.001
2.3.1.47	Keto acid dehydrogenase	Generic amino acid degradation	*Hafnia paralvei*	-4,92	-4,06	0,00	-3,86	-4,84	<0.001
2.3.1.51	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,64	-3,04	-4,95	-2,21	-2,14	<0.001
2.3.1.51	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,00	-2,01	-2,00	-2,24	-2,26	<0.001
2.3.1.51	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,00	-2,01	-2,00	-2,24	-2,26	<0.001
2.3.1.51	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,18	-2,18	-2,18	-2,42	-2,44	<0.001
2.3.1.51	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,18	-2,18	-2,18	-2,42	-2,44	<0.001
2.3.1.51	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,71	-5,68	-3,54	-4,39	0,00	<0.001
2.3.1.54	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,98	-1,98	-2,01	-2,06	-2,28	<0.001
2.3.1.54	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,98	-1,98	-2,01	-2,06	-2,28	<0.001
2.3.1.54	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,01	-2,04	-2,05	-2,43	-2,24	<0.001
2.3.1.54	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,01	-2,04	-2,05	-2,43	-2,24	<0.001
2.3.1.54	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,56	0,00	-5,02	-3,75	-4,48	<0.001
2.3.1.54	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,27	-4,84	-3,10	-4,03	-4,14	<0.001
2.3.1.8	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,11	-2,30	-2,32	-2,52	-2,58	<0.001
2.3.1.8	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,11	-2,30	-2,32	-2,52	-2,58	<0.001
2.3.1.8	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,08	-1,98	-2,00	-2,19	-2,19	<0.001
2.3.1.8	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,08	-1,98	-2,00	-2,19	-2,19	<0.001
2.3.1.8	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,66	-3,00	-5,02	-2,16	-2,07	<0.001
2.3.1.8	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,55	0,00	-4,85	-3,70	-4,49	<0.001

2.3.1.8	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,48	-5,45	-3,26	-4,43	-4,31	<0.001
2.3.1.8	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,52	-5,20	-3,51	-4,64	-4,98	<0.001
2.3.1.81	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,80	-5,87	-3,03	-4,85	-3,55	<0.001
2.3.1.86	Keto acid dehydrogenase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,43	0,00	-5,01	-3,69	-4,48	<0.001
2.3.1.89	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,63	-3,01	-5,09	-2,14	-2,11	<0.001
2.3.1.89	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,96	-1,98	-1,99	-2,18	-2,19	<0.001
2.3.1.89	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,96	-1,98	-1,99	-2,18	-2,19	<0.001
2.3.1.89	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,95	-1,99	-1,99	-2,17	-2,20	<0.001
2.3.1.89	Keto acid dehydrogenase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,95	-1,99	-1,99	-2,17	-2,20	<0.001
2.3.1.89	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-4,19	-5,26	-3,24	-4,20	-4,41	<0.001
2.3.1.89	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,59	-5,84	-3,52	-4,43	-4,89	<0.001
2.3.1.9	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,69	-3,08	-5,21	-2,24	-2,17	<0.001
2.3.1.9	Keto acid dehydrogenase	Generic amino acid degradation	*Lactococcus lactis*	-3,48	-5,01	-3,11	-4,09	-4,47	<0.001
2.3.1.9	Keto acid dehydrogenase	Generic amino acid degradation	*Hafnia paralvei*	-5,05	-4,07	0,00	-3,76	-4,80	0,001
2.3.1.n2	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,67	-5,62	-3,52	-4,47	-5,04	<0.001
2.3.1.n3	Keto acid dehydrogenase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,39	-2,65	-4,73	-1,84	-1,71	<0.001
2.3.1.n3	Keto acid dehydrogenase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,78	-5,33	-3,48	-4,75	-5,01	0,000501
2.5.1.47	O-acetylserine sulfhydrolase	Cysteine methionine degradation	*Hafnia paralvei*	-4,67	-3,92	0,00	-3,68	-4,43	0,004
2.5.1.47	O-acetylserine sulfhydrolase	Cysteine methionine degradation	*Lactobacillus delbrueckii*	-4,69	-3,03	-4,37	-2,19	-2,12	<0.001
2.5.1.47	O-acetylserine sulfhydrolase	Cysteine methionine degradation	*Lactococcus lactis*	-3,31	-4,87	-2,98	-3,98	-3,91	<0.001
2.5.1.47	O-acetylserine sulfhydrolase	Cysteine methionine degradation	*Leuconostoc pseudomesenteroides*	-3,65	-3,63	-3,27	-4,71	-4,35	0,03

2.5.1.47	O-acetylserine sulfhydrolase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,87	-1,78	-1,91	-1,97	-2,07	<0.001
2.5.1.47	O-acetylserine sulfhydrolase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,87	-1,78	-1,91	-1,97	-2,07	<0.001
2.5.1.47	O-acetylserine sulfhydrolase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,75	-1,83	-1,76	-2,13	-2,05	<0.001
2.5.1.47	O-acetylserine sulfhydrolase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,75	-1,83	-1,76	-2,13	-2,05	<0.001
2.5.1.48	Cystathionine gamma synthase	Cysteine methionine degradation	*Leuconostoc pseudomesenteroides*	-3,57	-5,48	-3,47	-4,38	-5,00	<0.001
2.5.1.48	Cystathionine gamma synthase	Cysteine methionine degradation	*Streptococcus thermophilus*	-3,43	-3,38	-3,18	-3,32	-3,32	<0.001
2.5.1.48	Cystathionine gamma synthase	Cysteine methionine degradation	*Streptococcus thermophilus*	-3,43	-3,38	-3,18	-3,32	-3,32	0,005
2.5.1.48	Cystathionine gamma synthase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,76	-1,75	-1,77	-1,97	-2,00	<0.001
2.5.1.48	Cystathionine gamma synthase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,76	-1,75	-1,77	-1,97	-2,00	0,005
2.5.1.49	O-acetylhomoserine sulfhydrolase	Cysteine methionine degradation	*Bifidobacterium mongoliense*	-4,59	0,00	-4,97	-3,74	-4,39	<0.001
2.5.1.49	O-acetylhomoserine sulfhydrolase	Cysteine methionine degradation	*Lactobacillus delbrueckii*	-4,74	-3,34	-5,87	-2,55	-2,43	<0.001
2.5.1.49	O-acetylhomoserine sulfhydrolase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,22	-2,04	-2,29	-2,22	-2,40	<0.001
2.5.1.49	O-acetylhomoserine sulfhydrolase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,22	-2,04	-2,29	-2,22	-2,40	<0.001
2.5.1.49	O-acetylhomoserine sulfhydrolase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,09	-2,28	-2,10	-2,86	-2,50	<0.001
2.5.1.49	O-acetylhomoserine sulfhydrolase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,09	-2,28	-2,10	-2,86	-2,50	<0.001
2.6.1.1	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,71	-1,88	-1,83	-2,14	-1,95	<0.001
2.6.1.1	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,71	-1,88	-1,83	-2,14	-1,95	<0.001
2.6.1.1	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,84	-1,69	-1,75	-1,86	-2,09	<0.001
2.6.1.1	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,84	-1,69	-1,75	-1,86	-2,09	<0.001
2.6.1.1	Aspartate aminotransferase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,45	-2,74	-5,19	-1,93	-1,86	<0.001
2.6.1.1	Aspartate aminotransferase	Generic amino acid degradation	*Hafnia paralvei*	-4,80	-3,94	-5,75	-3,67	-4,66	0,002

2.6.1.1	Aspartate aminotransferase	Generic amino acid degradation	*Lactococcus lactis*	-2,98	-4,65	-2,94	-3,74	-3,88	0,002
2.6.1.1	Aspartate aminotransferase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,62	-6,23	-3,44	-4,46	-5,22	<0.001
2.6.1.11	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,94	-1,87	-1,89	-2,17	-2,23	<0.001
2.6.1.11	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,94	-1,87	-1,89	-2,17	-2,23	<0.001
2.6.1.11	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,12	-2,29	-2,30	-2,36	-2,38	<0.001
2.6.1.11	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,12	-2,29	-2,30	-2,36	-2,38	<0.001
2.6.1.11	Aspartate aminotransferase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,40	0,00	-5,16	-3,75	-4,47	<0.001
2.6.1.11	Aspartate aminotransferase	Generic amino acid degradation	*Hafnia paralvei*	-5,11	-4,35	-5,64	-4,04	-4,86	0,001
2.6.1.11	Aspartate aminotransferase	Generic amino acid degradation	*Lactococcus lactis*	-3,36	-5,15	-3,25	-4,33	-4,22	0,001
2.6.1.11	Aspartate aminotransferase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,58	-5,19	-3,41	-4,33	-5,13	<0.001
2.6.1.16	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,83	-1,83	-1,88	-2,09	-2,15	<0.001
2.6.1.16	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,83	-1,83	-1,88	-2,09	-2,15	<0.001
2.6.1.16	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,97	-2,76	-2,72	-3,17	-2,89	<0.001
2.6.1.16	Aspartate aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,97	-2,76	-2,72	-3,17	-2,89	<0.001
2.6.1.16	Aspartate aminotransferase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,74	-3,02	-5,19	-2,19	-2,11	<0.001
2.6.1.16	Aspartate aminotransferase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,58	0,00	-4,85	-3,71	-4,44	<0.001
2.6.1.16	Aspartate aminotransferase	Generic amino acid degradation	*Lactococcus lactis*	-4,17	-5,59	-3,30	-4,21	-4,42	<0.001
2.6.1.19	Aspartate aminotransferase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,67	-5,40	-3,50	-4,42	-5,62	<0.001
2.6.1.42	Branched-chain aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,97	-1,94	-1,95	-2,15	-2,17	<0.001
2.6.1.42	Branched-chain aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,97	-1,94	-1,95	-2,15	-2,17	<0.001
2.6.1.42	Branched-chain aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,12	-2,18	-2,19	-2,41	-2,44	<0.001

2.6.1.42	Branched-chain aminotransferase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,12	-2,18	-2,19	-2,41	-2,44	<0.001
2.6.1.42	Branched-chain aminotransferase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,60	0,00	-4,91	-3,68	-4,41	<0.001
2.6.1.42	Branched-chain aminotransferase	Generic amino acid degradation	*Lactococcus lactis*	-3,35	-4,60	-3,12	-4,15	-4,23	<0.001
2.6.1.42	Branched-chain aminotransferase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,65	-5,36	-3,50	-4,39	-5,03	<0.001
2.7.1.39	Homoserine kinase	Cysteine methionine degradation	*Lactobacillus delbrueckii*	-4,65	-2,98	-5,03	-2,15	-2,04	<0.001
2.7.1.39	Homoserine kinase	Cysteine methionine degradation	*Lactococcus lactis*	-3,41	-5,26	-3,28	-4,40	-4,30	<0.001
2.7.1.39	Homoserine kinase	Cysteine methionine degradation	*Leuconostoc pseudomesenteroides*	-3,68	-5,98	-3,63	-4,62	-5,23	0,001
2.7.1.39	Homoserine kinase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,93	-1,83	-2,06	-1,94	-1,99	<0.001
2.7.1.39	Homoserine kinase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,93	-1,83	-2,06	-1,94	-1,99	0,003
2.7.1.39	Homoserine kinase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,04	-2,27	-2,00	-2,83	-2,88	<0.001
2.7.1.39	Homoserine kinase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,04	-2,27	-2,00	-2,83	-2,88	0,003
2.7.2.1	Acylkinase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,77	-2,93	-4,99	-2,09	-2,06	<0.001
2.7.2.1	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,87	-1,87	-1,89	-2,08	-2,10	<0.001
2.7.2.1	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,87	-1,87	-1,89	-2,08	-2,10	<0.001
2.7.2.1	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,24	-2,25	-2,25	-2,42	-2,46	<0.001
2.7.2.1	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,24	-2,25	-2,25	-2,42	-2,46	<0.001
2.7.2.1	Acylkinase	Generic amino acid degradation	*Lactococcus lactis*	-3,14	-4,52	-2,89	-3,96	-3,93	<0.001
2.7.2.1	Acylkinase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,36	-4,96	-3,19	-4,24	-4,68	<0.001
2.7.2.11	Acylkinase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,64	-3,05	-5,24	-2,24	-2,14	<0.001
2.7.2.11	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,02	-2,04	-2,21	-2,06	-2,35	<0.001
2.7.2.11	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,02	-2,04	-2,21	-2,06	-2,35	<0.001

2.7.2.11	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,05	-1,95	-2,56	-2,20	<0.001
2.7.2.11	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,03	-2,05	-1,95	-2,56	-2,20	<0.001
2.7.2.11	Acylkinase	Generic amino acid degradation	*Lactococcus lactis*	-3,55	-4,87	-3,32	-4,75	-4,32	0,0005
2.7.2.11	Acylkinase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,60	-4,49	-3,39	-4,16	-4,79	0,0006
2.7.2.2	Acylkinase	Generic amino acid degradation	*Lactococcus lactis*	-3,14	-5,12	-3,50	-4,35	-4,12	0,018
2.7.2.3	Acylkinase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,72	-2,98	-4,83	-2,18	-2,14	<0.001
2.7.2.3	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-1,72	-1,73	-1,77	-1,94	-1,99	<0.001
2.7.2.3	Acylkinase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,58	0,00	-5,02	-3,83	-4,59	<0.001
2.7.2.3	Acylkinase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,57	-4,97	-3,43	-4,28	-5,11	<0.001
2.7.2.3	Acylkinase	Generic amino acid degradation	*Lactococcus lactis*	-3,30	-4,96	-3,19	-4,21	-4,11	<0.001
2.7.2.4	Acylkinase	Generic amino acid degradation	*Lactobacillus delbrueckii*	-4,65	-2,89	-4,84	-2,04	-2,02	<0.001
2.7.2.4	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,09	-2,09	-2,10	-2,28	-2,32	<0.001
2.7.2.4	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,09	-2,09	-2,10	-2,28	-2,32	<0.001
2.7.2.4	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,09	-2,08	-2,10	-2,28	-2,32	<0.001
2.7.2.4	Acylkinase	Generic amino acid degradation	*Streptococcus thermophilus*	-2,09	-2,08	-2,10	-2,28	-2,32	<0.001
2.7.2.4	Acylkinase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,11	0,00	-4,42	-3,34	-4,05	<0.001
2.7.2.4	Acylkinase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,65	0,00	-3,54	-4,64	-5,01	0,0005
2.7.2.8	Acylkinase	Generic amino acid degradation	*Bifidobacterium mongoliense*	-4,61	0,00	-5,00	-3,84	-4,47	<0.001
2.7.2.8	Acylkinase	Generic amino acid degradation	*Lactococcus lactis*	-3,50	-5,07	-3,23	-4,32	-4,34	<0.001
2.7.2.8	Acylkinase	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,55	-6,05	-3,36	-4,39	-4,75	<0.001
3.1.1.1	Esterase A	Generic amino acid degradation	*Leuconostoc pseudomesenteroides*	-3,59	-5,59	-3,53	-4,57	-5,09	0,0006

3.1.1.29	Arylesterase	Lipid metabolism	*Bifidobacterium mongoliense*	-4,45	0,00	-4,91	-3,70	-4,30	<0.001
3.1.1.29	Arylesterase	Lipid metabolism	*Lactobacillus delbrueckii*	-4,66	-2,81	-5,02	-2,03	-1,90	<0.001
3.1.1.29	Arylesterase	Lipid metabolism	*Lactococcus lactis*	-3,54	-4,95	-3,25	-4,53	-4,11	<0.001
3.1.1.29	Arylesterase	Lipid metabolism	*Leuconostoc pseudomesenteroides*	-3,57	0,00	-3,39	-4,60	-4,74	<0.001
3.1.1.29	Arylesterase	Lipid metabolism	*Streptococcus thermophilus*	-1,94	-1,91	-1,94	-2,13	-2,15	<0.001
3.1.1.29	Arylesterase	Lipid metabolism	*Streptococcus thermophilus*	-1,94	-1,91	-1,94	-2,13	-2,15	<0.001
3.1.1.29	Arylesterase	Lipid metabolism	*Streptococcus thermophilus*	-2,14	-2,13	-2,13	-2,32	-2,34	<0.001
3.1.1.29	Arylesterase	Lipid metabolism	*Streptococcus thermophilus*	-2,14	-2,13	-2,13	-2,32	-2,34	<0.001
3.1.1.31	Triacylglycerol lipase	Lipid metabolism	*Bifidobacterium mongoliense*	-4,14	0,00	-4,65	-3,35	-4,14	<0.001
3.1.1.31	Triacylglycerol lipase	Lipid metabolism	*Lactobacillus delbrueckii*	-4,96	-3,08	-5,04	-2,26	-2,17	<0.001
3.4.11.18	unique aminopeptidases	Amino acid metabolism	*Streptococcus thermophilus*	-1,80	-1,80	-1,83	-2,03	-2,09	<0.001
3.4.11.18	unique aminopeptidases	Amino acid metabolism	*Streptococcus thermophilus*	-1,80	-1,80	-1,83	-2,03	-2,09	<0.001
3.4.11.18	unique aminopeptidases	Amino acid metabolism	*Streptococcus thermophilus*	-2,50	-2,51	-2,52	-2,68	-2,61	<0.001
3.4.11.18	unique aminopeptidases	Amino acid metabolism	*Streptococcus thermophilus*	-2,50	-2,51	-2,52	-2,68	-2,61	0,003
3.4.11.18	unique aminopeptidases	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,42	-2,81	-4,73	-2,02	-1,93	<0.001
3.4.11.18	unique aminopeptidases	Amino acid metabolism	*Lactococcus lactis*	-3,35	-4,75	-3,17	-4,08	-4,19	<0.001
3.4.11.18	unique aminopeptidases	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,60	-4,91	-3,44	-4,53	-5,24	<0.001
3.4.11.2	Aminopeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,09	-2,08	-2,14	-2,37	-2,42	<0.001
3.4.11.2	Aminopeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,09	-2,08	-2,14	-2,37	-2,42	<0.001
3.4.11.2	Aminopeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,24	-2,22	-2,28	-2,50	-2,48	<0.001
3.4.11.2	Aminopeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,24	-2,22	-2,28	-2,50	-2,48	<0.001
3.4.11.2	Aminopeptidase	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,69	-2,87	-4,84	-2,07	-1,95	<0.001
3.4.11.2	Aminopeptidase	Amino acid metabolism	*Lactococcus lactis*	-3,41	-4,80	-3,17	-4,19	-4,19	<0.001
3.4.11.24	Aminopeptidase	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,57	-5,26	-3,43	-4,52	-5,09	0,0005
3.4.11.4	Tripeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,18	-2,12	-2,04	-2,42	-2,45	<0.001
3.4.11.4	Tripeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,18	-2,12	-2,04	-2,42	-2,45	<0.001
3.4.11.4	Tripeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,08	-2,13	-2,34	-2,32	-2,35	<0.001

3.4.11.4	Tripeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,08	-2,13	-2,34	-2,32	-2,35	<0.001
3.4.11.4	Tripeptidase	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,35	-2,55	-4,76	-1,75	-1,69	<0.001
3.4.11.4	Tripeptidase	Amino acid metabolism	*Lactococcus lactis*	-3,31	-4,91	-3,11	-4,11	-4,11	<0.001
3.4.11.4	Tripeptidase	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,62	0,00	-3,53	-4,63	-4,96	<0.001
3.4.11.5	Proline peptidase	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,42	-2,57	-4,42	-1,76	-1,63	<0.001
3.4.11.7	unique aminopeptidases	Amino acid metabolism	*Streptococcus thermophilus*	-1,89	-1,86	-1,91	-2,17	-2,21	<0.001
3.4.11.7	unique aminopeptidases	Amino acid metabolism	*Streptococcus thermophilus*	-1,89	-1,86	-1,91	-2,17	-2,21	0,01
3.4.11.7	unique aminopeptidases	Amino acid metabolism	*Streptococcus thermophilus*	-2,23	-2,30	-2,31	-2,40	-2,41	<0.001
3.4.11.7	unique aminopeptidases	Amino acid metabolism	*Streptococcus thermophilus*	-2,23	-2,30	-2,31	-2,40	-2,41	0,01
3.4.11.7	unique aminopeptidases	Amino acid metabolism	*Lactococcus lactis*	-3,95	-4,99	-3,22	-4,10	-4,42	<0.001
3.4.11.7	unique aminopeptidases	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,67	-5,64	-3,44	-4,71	-5,32	<0.001
3.4.11.9	Proline peptidase	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,53	-2,88	-5,06	-2,08	-1,98	<0.001
3.4.11.9	Proline peptidase	Amino acid metabolism	*Lactococcus lactis*	-3,52	-5,39	-3,65	-5,63	-4,52	0,006
3.4.11.9	Proline peptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,12	-2,02	-2,10	-2,53	-2,26	<0.001
3.4.11.9	Proline peptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,12	-2,02	-2,10	-2,53	-2,26	<0.001
3.4.11.9	Proline peptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,05	-2,17	-2,13	-2,16	-2,41	<0.001
3.4.11.9	Proline peptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,05	-2,17	-2,13	-2,16	-2,41	<0.001
3.4.13.18	Dipeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,00	-1,93	-2,01	-2,21	-2,24	<0.001
3.4.13.18	Dipeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,00	-1,93	-2,01	-2,21	-2,24	<0.001
3.4.13.18	Dipeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,26	-2,31	-2,32	-2,55	-2,58	<0.001
3.4.13.18	Dipeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,26	-2,31	-2,32	-2,55	-2,58	<0.001
3.4.13.18	Dipeptidase	Amino acid metabolism	*Bifidobacterium mongoliense*	-4,58	0,00	-4,88	-3,68	-4,51	<0.001
3.4.13.20	Dipeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,00	-1,93	-2,01	-2,21	-2,24	<0.001
3.4.13.20	Dipeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,00	-1,93	-2,01	-2,21	-2,24	<0.001
3.4.13.20	Dipeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,26	-2,31	-2,32	-2,55	-2,58	<0.001
3.4.13.20	Dipeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,26	-2,31	-2,32	-2,55	-2,58	<0.001
3.4.13.9	Dipeptidase	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,29	-2,66	-4,64	-1,84	-1,74	<0.001

3.4.13.9	Dipeptidase	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,49	-4,34	-3,44	-4,40	-5,39	<0.001
3.4.14.11	Proline peptidase	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,69	-3,03	-4,82	-2,21	-2,13	<0.001
3.4.14.11	Proline peptidase	Amino acid metabolism	*Streptococcus thermophilus*	-1,96	-1,86	-1,86	-2,12	-2,13	<0.001
3.4.14.11	Proline peptidase	Amino acid metabolism	*Streptococcus thermophilus*	-1,96	-1,86	-1,86	-2,12	-2,13	<0.001
3.4.14.11	Proline peptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,05	-2,24	-2,34	-2,34	-2,48	<0.001
3.4.14.11	Proline peptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,05	-2,24	-2,34	-2,34	-2,48	<0.001
3.4.14.11	Proline peptidase	Amino acid metabolism	*Lactococcus lactis*	-3,46	-5,66	-3,61	-4,86	-4,60	0,006
3.4.14.11	Proline peptidase	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,67	-5,41	-3,53	-4,70	-5,15	<0.001
3.4.19.3	unique aminopeptidases	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,90	-3,14	-5,30	-2,27	-2,26	<0.001
3.4.19.3	unique aminopeptidases	Amino acid metabolism	*Lactococcus lactis*	-2,95	-4,64	-3,16	-3,23	-4,00	<0.001
3.4.19.3	unique aminopeptidases	Amino acid metabolism	*Leuconostoc mesenteroides*	-3,13	-3,44	-3,47	-3,30	-4,24	0,05
3.4.21.107	Other Proteinases	Amino acid metabolism	*Hafnia paralvei*	-4,74	-4,04	-5,54	-3,74	-5,13	<0.000
3.4.21.107	Other Proteinases	Amino acid metabolism	*Lactococcus lactis*	-3,74	-4,91	-3,26	-4,21	-4,32	<0.001
3.4.21.53	Other Proteinases	Amino acid metabolism	*Hafnia paralvei*	-4,98	-3,94	-5,21	-3,75	-4,86	0,006
3.4.21.53	Other Proteinases	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,78	-3,03	-4,95	-2,21	-2,16	<0.001
3.4.21.53	Other Proteinases	Amino acid metabolism	*Lactococcus lactis*	-3,41	-5,01	-3,25	-4,27	-4,13	<0.001
3.4.21.53	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,54	-2,55	-2,54	-2,70	-2,72	<0.001
3.4.21.53	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,54	-2,55	-2,54	-2,70	-2,72	<0.001
3.4.21.53	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,83	-1,85	-1,87	-2,06	-2,09	<0.001
3.4.21.53	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,83	-1,85	-1,87	-2,06	-2,09	<0.001
3.4.21.83	Other Proteinases	Amino acid metabolism	*Bifidobacterium mongoliense*	-4,59	0,00	-5,14	-3,79	-4,55	<0.001
3.4.21.88	Other Proteinases	Amino acid metabolism	*Bifidobacterium mongoliense*	-4,25	0,00	-4,78	-3,56	-4,22	<0.001
3.4.21.88	Other Proteinases	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,65	-2,95	-4,96	-2,14	-2,06	<0.001
3.4.21.88	Other Proteinases	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,63	-5,49	-3,57	-4,60	-5,30	<0.001
3.4.21.89	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,60	-1,55	-1,61	-1,80	-1,80	<0.001
3.4.21.89	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,60	-1,55	-1,61	-1,80	-1,80	<0.001
3.4.21.89	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,11	-2,25	-2,15	-2,44	-2,47	<0.001

3.4.21.89	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,11	-2,25	-2,15	-2,44	-2,47	<0.001
3.4.21.89	Other Proteinases	Amino acid metabolism	*Bifidobacterium mongoliense*	-4,59	0,00	-4,82	-3,59	-4,33	<0.001
3.4.21.89	Other Proteinases	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,76	-3,15	-4,86	-2,40	-2,36	<0.001
3.4.21.89	Other Proteinases	Amino acid metabolism	*Lactococcus lactis*	-4,23	-5,04	-3,33	-4,33	-4,38	<0.001
3.4.21.89	Other Proteinases	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,30	-4,94	-3,08	-4,23	-5,01	<0.001
3.4.21.92	Other Proteinases	Amino acid metabolism	*Bifidobacterium mongoliense*	-4,26	0,00	-4,41	-3,36	-4,12	<0.001
3.4.21.92	Other Proteinases	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,61	-2,86	-5,07	-1,97	-1,97	<0.001
3.4.21.92	Other Proteinases	Amino acid metabolism	*Lactococcus lactis*	-3,39	-4,81	-3,13	-4,49	-4,07	<0.001
3.4.21.92	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,83	-1,79	-1,82	-2,05	-2,07	<0.001
3.4.21.92	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,83	-1,79	-1,82	-2,05	-2,07	<0.001
3.4.21.92	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,34	-2,34	-2,32	-2,53	-2,55	<0.001
3.4.21.92	Other Proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,34	-2,34	-2,32	-2,53	-2,55	<0.001
3.4.21.96	Cell-wall bound proteinase	Amino acid metabolism	*Lactococcus lactis*	-3,25	-4,72	-2,82	-3,30	-3,90	<0.001
3.4.22.40	Aminopeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,24	-2,23	-2,25	-2,52	-2,51	<0.001
3.4.22.40	Aminopeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-2,24	-2,23	-2,25	-2,52	-2,51	<0.001
3.4.22.40	Aminopeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-1,92	-1,94	-1,95	-2,12	-2,16	<0.001
3.4.22.40	Aminopeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-1,92	-1,94	-1,95	-2,12	-2,16	<0.001
3.4.22.40	Aminopeptidase	Amino acid metabolism	*Bifidobacterium mongoliense*	-4,48	0,00	-4,75	-3,79	-4,44	<0.001
3.4.22.40	Aminopeptidase	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,64	-2,41	-4,37	-1,66	-1,54	<0.001
3.4.22.40	Aminopeptidase	Amino acid metabolism	*Lactococcus lactis*	-3,46	-5,06	-3,27	-4,45	-4,33	<0.001
3.4.22.40	Aminopeptidase	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,63	-5,33	-3,28	-4,83	-5,93	<0.001
3.4.23.36	Aspartyl proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,90	-1,88	-1,93	-2,20	-2,19	<0.001
3.4.23.36	Aspartyl proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,90	-1,88	-1,93	-2,20	-2,19	<0.001
3.4.23.36	Aspartyl proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,07	-2,06	-2,11	-2,39	-2,37	<0.001
3.4.23.36	Aspartyl proteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,07	-2,06	-2,11	-2,39	-2,37	<0.001
3.4.23.36	Aspartyl proteinases	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,55	-2,97	-4,75	-2,18	-2,16	<0.001
3.4.23.36	Aspartyl proteinases	Amino acid metabolism	*Bifidobacterium mongoliense*	-4,44	0,00	-4,69	-3,59	-4,31	<0.001
3.4.24.11	Metalloproteinases	Amino acid metabolism	*Bifidobacterium mongoliense*	-4,54	0,00	-4,93	-3,66	-4,41	<0.001

3.4.24.13	Metalloproteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,98	-1,90	-1,94	-2,17	-2,19	<0.001
3.4.24.57	Metalloproteinases	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,67	-2,90	-4,91	-2,09	-2,01	<0.001
3.4.24.57	Metalloproteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,86	-1,95	-1,82	-2,00	-2,31	<0.001
3.4.24.57	Metalloproteinases	Amino acid metabolism	*Streptococcus thermophilus*	-1,86	-1,95	-1,82	-2,00	-2,31	<0.001
3.4.24.57	Metalloproteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,17	-2,08	-2,47	-2,61	-2,18	<0.001
3.4.24.57	Metalloproteinases	Amino acid metabolism	*Streptococcus thermophilus*	-2,17	-2,08	-2,47	-2,61	-2,18	<0.001
3.4.25.2	Threonine peptidase	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,62	-2,97	-5,15	-2,13	-2,06	<0.001
4.1.1.5	Acetolactate decarboxylase	Citrate fermentation	*Bifidobacterium mongoliense*	-4,42	0,00	-4,94	-3,76	-4,55	<0.001
4.1.1.5	Acetolactate decarboxylase	Citrate fermentation	*Lactococcus lactis*	-3,32	-4,78	-2,99	-4,00	-4,06	<0.001
4.1.1.5	Acetolactate decarboxylase	Citrate fermentation	*Leuconostoc pseudomesenteroides*	-3,36	-5,23	-3,24	-4,63	-4,94	<0.000
4.1.1.5	Acetolactate decarboxylase	Citrate fermentation	*Streptococcus thermophilus*	-2,19	-2,15	-2,10	-2,34	-2,43	<0.001
4.1.1.5	Acetolactate decarboxylase	Citrate fermentation	*Streptococcus thermophilus*	-2,19	-2,15	-2,10	-2,34	-2,43	<0.001
4.1.1.5	Acetolactate decarboxylase	Citrate fermentation	*Streptococcus thermophilus*	-1,95	-1,96	-2,04	-2,28	-2,25	<0.001
4.1.1.5	Acetolactate decarboxylase	Citrate fermentation	*Streptococcus thermophilus*	-1,95	-1,96	-2,04	-2,28	-2,25	<0.001
4.1.3.6	Citrate lyase	Citrate fermentation	*Hafnia paralvei*	-5,04	-3,88	-5,78	-3,62	-4,89	0,002
4.1.3.6	Citrate lyase	Citrate fermentation	*Lactobacillus helveticus*	0,00	0,00	0,00	-3,04	-3,36	<0.001
4.1.3.6	Citrate lyase	Citrate fermentation	*Leuconostoc mesenteroides*	-3,78	-2,70	-3,26	-3,95	-4,09	0,03
4.1.3.6	Citrate lyase	Citrate fermentation	*Leuconostoc pseudomesenteroides*	-3,42	-6,00	-3,76	-4,32	-5,09	<0.000
4.2.1.22	Cystathionine-beta-synthase	Cysteine methionine degradation	*Lactobacillus brevis*	-4,89	-4,98	-4,65	-3,33	-4,49	<0.001
4.2.1.22	Cystathionine-beta-synthase	Cysteine methionine degradation	*Lactobacillus delbrueckii*	-3,61	-3,43	-4,66	-3,04	-3,04	<0.001
4.2.1.22	Cystathionine-beta-synthase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,15	-1,93	-1,98	-2,20	-2,22	<0.001
4.2.1.22	Cystathionine-beta-synthase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,15	-1,93	-1,98	-2,20	-2,22	<0.001
4.2.1.22	Cystathionine-beta-synthase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,46	-2,46	-2,50	-2,67	-2,71	<0.001

4.2.1.22	Cystathionine-beta-synthase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,46	-2,46	-2,50	-2,67	-2,71	<0.001
4.4.1.1	Cystathionine-Gamma-lyase	Cysteine methionine degradation	*Leuconostoc pseudomesenteroides*	-3,56	0,00	-3,35	-4,20	-4,96	<0.000
4.4.1.16	Cystathionine-Gamma-lyase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,90	-1,91	-1,96	-2,12	-2,16	<0.001
4.4.1.16	Cystathionine-Gamma-lyase	Cysteine methionine degradation	*Streptococcus thermophilus*	-1,90	-1,91	-1,96	-2,12	-2,16	<0.001
4.4.1.16	Cystathionine-Gamma-lyase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,13	-2,14	-2,11	-2,34	-2,36	<0.001
4.4.1.16	Cystathionine-Gamma-lyase	Cysteine methionine degradation	*Streptococcus thermophilus*	-2,13	-2,14	-2,11	-2,34	-2,36	<0.001
4.4.1.8	Cystathionine beta-lyase	Cysteine methionine degradation	*Lactobacillus delbrueckii*	-4,58	-2,90	-4,93	-2,06	-2,01	<0.001
4.4.1.8	Cystathionine beta-lyase	Cysteine methionine degradation	*Leuconostoc pseudomesenteroides*	-3,92	-6,14	-3,82	-4,73	-5,47	<0.001
6.3.4.13	Dipeptidase	Amino acid metabolism	*Lactobacillus delbrueckii*	-4,68	-3,05	-5,29	-2,23	-2,18	<0.001
6.3.4.13	Dipeptidase	Amino acid metabolism	*Bifidobacterium mongoliense*	-4,51	0,00	-4,97	-3,68	-4,47	<0.001
6.3.4.13	Dipeptidase	Amino acid metabolism	*Lactococcus lactis*	-3,41	-4,99	-3,25	-4,57	-4,24	<0.001
6.3.4.13	Dipeptidase	Amino acid metabolism	*Leuconostoc pseudomesenteroides*	-3,68	-5,96	-3,47	-4,71	-5,40	<0.001
6.3.4.13	Dipeptidase	Amino acid metabolism	*Streptococcus thermophilus*	-1,65	-1,67	-1,70	-1,86	-1,91	<0.001

Table S12. Correlation between the three most abundant volatile organic compounds (VOC), genes and associated microrganism. Only the strongest significant correlations are reported.

Family	VOC	Gene	Associated m.o.	Enzyme	Metabolism	Correlation	Significance
Carboxylic acids	Acetic acid	2.3.1.20	*Acinetobacter johnsonii*	Keto acid dehydrogenase	Generic amino acid degradation	-0.6593	0.04
	Acetic acid	3.1.1.24	*Acinetobacter johnsonii, Lacticaseibacillus rhamnosus , Pseudomonas fluorescens*	Arylesterase	Lipids metabolism	-0.5070	0.04
	Acetic acid	1.1.1.18	*Lactiplantibacillus plantarum,Propionibacterium freudenreichii, Lactobacillus diolivorans*	Alcohol dehydrogenase	Generic amino acid degradation	0.3185	0.04
	Propionic acid	1.1.1.219	*Lactococcus lactis, Pseudomonas fluorescens,*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.4131	0.05
	Propionic acid	1.1.1.31	*Lactococcus lactis, Pseudomonas simiae, Lacticaseibacillus rhamnosus*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.3086	0.03
	Propionic acid	3.4.25.1	*Propionibacterium freudenreichii*	Threonine peptidase	Amminoacid metabolsim	0.7279	0.008
	Propionic acid	2.3.1.189	*Propionibacterium freudenreichii, Pseudomonas simiae, Escherichia coli*	Keto acid dehydrogenase	Generic amino acid degradation	0.6441	0.01
	Butanoic acid	2.3.1.88	*Lactobacillus otakiensis*	Keto acid dehydrogenase	Generic amino acid degradation	-0.5385	0.05
	Butanoic acid	1.2.4.4	*Lactiplantibacillus plantarum,Propionibacterium freudenreichii, Enterococcus faecalis*	Keto acid dehydrogenase	Generic amino acid degradation	-0.5253	0.01
	Butanoic acid	4.4.1.11	*Lactobacillus delbrueckii, Citrobacter braakii, Raoultella terrigena*	Cystathionine-Gamma-lyase	Cyst meth degradation	0.6321	0.03
Alcohols	Butan-2-ol	2.3.1.189	*Propionibacterium freudenreichii, Pseudomonas fluorescens*	Keto acid dehydrogenase	Generic amino acid degradation	-0.7333	0.04
	Butan-2-ol	1.1.1.313	*Propionibacterium freudenreichii, Pseudomonas fluorescens*	D-hydroxy acid dehydrogenase	Generic amino acid degradation	-0.7000	0.03
	Butan-2-ol	2.3.1.242	*Enterobacter cloacae, Raoultella ornithinolytica, Hafnia alvei*	Keto acid dehydrogenase	Generic amino acid degradation	0.5139	0.03
	Butan-2-ol	2.3.1.118	*Lactococcus lactis, Escherichia coli, Klebsiella michiganensis*	Keto acid dehydrogenase	Generic amino acid degradation	0.5903	0.01
	Ethanol	1.1.1.291	*Acinetobacter johnsonii*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.5944	0.04
	Ethanol	3.4.24.11	*Bifidobacterium mongoliense*	Metalloproteinases	Amminoacid metabolsim	-0.5114	0.005

	Ethanol	3.4.22.49	*Limosilactobacillus fermentum*	Cysteine Proteinases	Amminoacid metabolsim	0.8095	0.02
	Ethanol	2.3.1.109	*Pseudomonas simiae, Acinetobacter johnsonii*	Keto acid dehydrogenase	Generic amino acid degradation	0.7176	<0.001
	3-methyl-butan-1-ol	1.1.1.291	*Acinetobacter johnsonii*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.8112	0.009
	3-methyl-butan-1-ol	4.4.1.11	*Lactobacillus delbrueckii, Escherichia coli*	Cystathionine-Gamma-lyase	Cyst meth degradation	-0.6250	0.02
	3-methyl-butan-1-ol	4.1.1.59	*Lactiplantibacillus plantarum*	Acetolactate decarboxylase	Citrate fermentation	0.6948	0.005
	3-methyl-butan-1-ol	1.1.1.405	*Lactiplantibacillus plantarum*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.5654	0.01
	Butan-2-one	1.1.1.105	*Lactococcus lactis*	Alcohol dehydrogenase	Generic amino acid degradation	-0.7357	<0.001
	Butan-2-one	1.1.1.202	*Lactiplantibacillus plantarum,Lentilactobacillus parabuchneri, Pseudomonas simiae*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.5989	<0.001
	Butan-2-one	1.1.1.91	*Bifidobacterium mongoliense*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.6018	<0.001
	Butan-2-one	2.3.1.86	*Bifidobacterium mongoliense*	Keto acid dehydrogenase	Generic amino acid degradation	0.5762	<0.001
	3-Hydroxybutan-2-one	4.4.1.11	*Lactobacillus delbrueckii, Escherichia coli, Raoultella terrigena*	Cystathionine-Gamma-lyase	Cyst meth degradation	-0.5874	0.05
Ketones	3-Hydroxybutan-2-one	1.1.1.301	*Enterococcus gilvus*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.8929	0.01
	3-Hydroxybutan-2-one	2.7.2.7	*Enterococcus gilvus, Enterococcus faecalis*	Acylkinase	Generic amino acid degradation	0.6848	0.003
	3-Hydroxybutan-2-one	1.1.1.62	*Lactococcus lactis*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.6059	0.005
	Butane-2,3-dione	1.1.1.346	*Bifidobacterium mongoliense, Enterobacter cloacae, Pediococcus pentosaceus*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.5128	0.03
	Butane-2,3-dione	4.4.1.13	*Lactococcus lactis*	Cystathionine-Gamma-lyase	Cyst meth degradation	0.9364	<0.001
	Butane-2,3-dione	1.1.1.62	*Lactococcus lactis*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.6818	0.04
	Butane-2,3-dione	1.1.1.18	*Propionibacterium freudenreichii,*	Alcohol dehydrogenase	Generic amino acid	0.6313	0.005

			Streptococcus parauberis, Klebsiella michiganensis		degradation		
Esters	Ethyl acetate	1.1.1.369	*Lactiplantibacillus plantarum,Lacticaseibacillus rhamnosus , Lactobacillus diolivorans*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.6424	0.04
	Ethyl acetate	3.1.1.24	*Acinetobacter johnsonii, Lacticaseibacillus rhamnosus , Hafnia alvei*	Arylesterase	Lipids metabolism	-0.6333	0.01
	Ethyl acetate	3.4.22.49	*Limosilactobacillus fermentum*	Cysteine Proteinase	Amminoacid metabolsim	0.7381	0.05
	Ethyl acetate	4.4.1.11	*Lactobacillus delbrueckii, Citrobacter braakii, Escherichia coli*	Cystathionine-Gamma-lyase	Cyst meth degradation	0.6036	0.03
	Ethyl butanoate	1.1.1.369	*Lactobacillus diolivorans, Lacticaseibacillus rhamnosus , Streptococcus parauberis*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.7517	0.02
	Ethyl butanoate	3.1.1.24	*Lacticaseibacillus rhamnosus , Pseudomonas fluorescens, Acinetobacter johnsonii*	Arylesterase	Lipids metabolism	-0.5551	0.02
	Ethyl butanoate	1.1.1.47	*Leuconostoc pseudomesenteroides, Enterococcus faecalis, Leuconostoc citreum*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.1730	0.02
	Ethyl butanoate	4.4.1.16	*Streptococcus thermophilus, Raoultella terrigena*	Cystathionine-Gamma-lyase	Cyst meth degradation	0.1706	0.02
	Ethyl hexanoate	1.1.1.369	*Lacticaseibacillus rhamnosus , Lactiplantibacillus plantarum,Lactobacillus diolivorans*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.7608	0.01
	Ethyl hexanoate	1.1.1.62	*Lactococcus lactis*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.5530	0.01
	Ethyl hexanoate	1.1.1.373	*Citrobacter braakii, Hafnia paralvei*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.3828	0.04
	Ethyl hexanoate	3.4.24.13	*Streptococcus thermophilus*	Metalloproteinases	Amminoacid metabolsim	0.3140	0.02
Others	Cyclopentane	1.1.1.313	*Propionibacterium freudenreichii, Pseudomonas fluorescens*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.4719	0.05
	Cyclopentane	1.1.1.18	*Lacticaseibacillus rhamnosus , Propionibacterium freudenreichii, Bifidobacterium mongoliense*	Alcohol dehydrogenase	Generic amino acid degradation	-0.3149	0.05
	Cyclopentane	1.1.1.53	*Lentilactobacillus parabuchneri, Lactobacillus parafarraginis, Limosilactobacillus fermentum*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.2934	0.01
	Cyclopentane	1.1.1.276	*Lactobacillus brevis, Bifidobacterium mongoliense*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.2365	0.01

	Toluene	2.3.1.189	*Glutamicibacter sp BW77, Propionibacterium freudenreichii*	Keto acid dehydrogenase	Generic amino acid degradation	-0.5490	0.03
	Toluene	3.1.1.10	*Lacticaseibacillus rhamnosus , Acinetobacter johnsonii*	Esterase A	Generic amino acid degradation	-0.5276	0.03
	Toluene	2.3.1.109	*Raoultella terrigena, Pseudomonas fluorescens, Escherichia coli*	Keto acid dehydrogenase	Generic amino acid degradation	0.6467	0.002
	Toluene	1.1.1.105	*Lactococcus lactis*	Alcohol dehydrogenase	Generic amino acid degradation	0.5053	0.03
	Xylene	3.1.1.11	*Escherichia coli, Klebsiella michiganensis*	Esterase A	Generic amino acid degradation	-0.6256	0.02
	Xylene	2.3.1.189	*Propionibacterium freudenreichii, Pseudomonas fluorescens*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.5711	0.02
	Xylene	1.1.1.271	*Enterobacter cloacae complex*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.8571	0.03
	Xylene	1.1.1.105	*Lactococcus lactis*	Alcohol dehydrogenase	Generic amino acid degradation	0.5421	0.02
Fatty acids	Hexanoic acid	1.1.1.405	*Lactiplantibacillus plantarum*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.6015	0.02
	Hexanoic acid	3.4.25.1	*Propionibacterium freudenreichii*	Threonine peptidase	Amminoacid metabolsim	-0.5851	0.04
	Hexanoic acid	4.4.1.11	*Lactobacillus delbrueckii, Escherichia coli, Raoultella terrigena*	Cystathionine-Gamma-lyase	Cyst meth degradation	0.7500	0.02
	Hexanoic acid	1.1.1.291	*Acinetobacter johnsonii*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.6923	0.04
	Octanoic acid	1.1.1.90	*Lentilactobacillus parabuchneri, Lactobacillus brevis, Lactobacillus kefiranofaciens*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.3301	0.04
	Octanoic acid	3.1.1.11	*Klebsiella michiganensis, Escherichia coli, Raoultella ornithinolytica*	Esterase A	Generic amino acid degradation	0.6256	0.04
Aldehydes	Benzaldehyde	1.1.1.83	*Raoultella terrigena*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.8571	0.02
	Benzaldehyde	3.1.1.10	*Lacticaseibacillus rhamnosus , Klebsiella michiganensis, Acinetobacter johnsonii*	Esterase A	Generic amino acid degradation	-0.5423	0.03
	Benzaldehyde	1.1.1.53	*Lentilactobacillus parabuchneri, Limosilactobacillus*	D-hydroxyacid dehydrogenase	Generic amino acid	0.3329	<0.001

		fermentum, Lactobacillus parafarraginis		degradation		
Benzaldehyde	1.1.1.42	*Streptococcus thermophilus, Lactococcus lactis*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.3172	<0.001
Nonanal	2.3.1.189	*Propionibacterium freudenreichii, Pseudomonas fluorescens, Glutamicibacter sp BW77*	Keto acid dehydrogenase	Generic amino acid degradation	-0.6348	0.01
Nonanal	1.1.1.91	*Bifidobacterium mongoliense*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.5384	0.007
Nonanal	2.3.1.79	*Enterococcus faecalis, Lactobacillus sakei, Lentilactobacillus buchneri*	Keto acid dehydrogenase	Generic amino acid degradation	0.3684	<0.001
Nonanal	1.1.1.53	*Lentilactobacillus parabuchneri, Limosilactobacillus fermentum , Lactobacillus parafarraginis*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.3089	<0.001
3-Methylbutanal	1.1.1.303	*Leuconostoc mesenteroides, Lactobacillus brevis, Leuconostoc pseudomesenteroides*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	-0.5176	0.03
3-Methylbutanal	1.1.1.271	*Enterobacter cloacae*	D-hydroxyacid dehydrogenase	Generic amino acid degradation	0.7857	0.05
3-Methylbutanal	2.3.1.189	*Propionibacterium freudenreichii, Glutamicibacter sp BW77, Pseudomonas fluorescens*	Keto acid dehydrogenase	Generic amino acid degradation	0.5686	0.02

Chapter 3: Comparison of analytical methods for cheese origin authentication

Dairy products hold a prominent position among the most common commodities vulnerable to food fraud, and within this category, protected land- and tradition-related labelled cheeses, such as those with a Protected Designation of Origin (PDO) or labelled as "mountain products," have become particularly susceptible to fraudulent activities. The reason behind this heightened risk lies in the high economic value of such cheeses. Consumers are willing to pay a premium for these types of traditional and typical cheeses due to their distinctive flavors, appealing appearance, and the perception of more natural and animal-friendly production attributes (Menozzi et al., 2022). The demand for these premium cheeses has created a lucrative market, making them an attractive target for unscrupulous actors seeking to capitalise on the opportunity presented by fraudulent practices. Mislabelling and fraudulent documentation are among the common methods used to deceive consumers and profit from the increased prices associated with authentic, region-specific cheeses.

In response to this challenge, regulatory authorities, as well as cheese producers and manufacturers, have been compelled to take measures to ensure the authenticity of these valuable cheeses. Geographical origin authentication methods have emerged as important tools in this endeavour. These methods aim to verify the geographic origin of cheeses, thereby providing consumers with the assurance of authentic traditional products. Among the range of analytical techniques explored for cheese origin authentication, chemico-physical analyses involving stable isotope ratios, trace elements, and fatty acid profiles have been traditionally employed. Additionally, newer approaches based on DNA-based methodologies and infrared analysis have garnered attention in recent years, offering promising avenues for enhancing the precision and efficiency of cheese origin authentication (Cardin et al., 2022).

In this study, the effectiveness of several analytical methods for cheese origin authentication, including DNA shotgun metagenomics (bacterial and viral community

profiling), volatilome analysis, near infrared spectroscopy, element metabolomics, and stable isotope ratio analysis were evaluated. To achieve this, mountain Caciotta cheese and sPLS-DA models, a statistical method capable of handling high-dimensional datasets and classifying observations based on their features (Le Cao et al., 2011), were employed.

Application of omics-based technologies for authenticating the origin of typical Italian mountain cheese with sPLS-DA modeling

Marco Cardin[1,2], Jérôme Mounier[2], Emmanuel Coton[2] , Barbara Cardazzo[1], Nicola Perini[3], Luana Bontempo[3], Severino Segato[4], Barbara Di Camillo[5], Marco Cappellato[5], Monika Coton[2], Lisa Carraro[1], Sarah Currò[1], Rosaria Lucchini[6] and Enrico Novelli[1]

[1] *Department of Comparative Biomedicine and Food Science, University of Padova, Viale Università 16, 35020 Legnaro, PD, Italy*

[2] *Univ Brest, INRAE, Laboratoire Universitaire de Biodiversité et Écologie Microbienne, F-29280 Plouzané, France*

[3] *Fondazione Edmund Mach, Via E. Mach, 1, 38098 San Michele all'Adige, Italy*

[4] *Department of Animal Medicine, Production and Health, University of Padova, Legnaro, Viale Università 16, 35020 Legnaro, PD, Italy*

[5] *Department of Information Engineering, University of Padova, Via Gradenigo 6/b, 35131 Padova, Italy*

[6] *Italian Health Authority and Research Organization for Animal Health and Food Safety (Istituto zooprofilattico sperimentale delle Venezie), Viale Università 10, 35020 Legnaro, PD, Italy*

Corresponding author: Enrico Novelli

Keywords: geographical origin, authentication, shotgun metagenomics,Near Infrared Spectroscopy, element metabolomics, stable isotopes

Abstract

Origin authentication methods are pivotal in counteracting frauds and provide evidence for certification systems. For these reasons, geographical origin authentication methods are used to ensure product origin. This study focuses on the origin authentication of a typical mountain cheese origin using multi-omics approaches, including shotgun metagenomics, volatilome, near infrared spectroscopy, stable isotopes, and elemental analysis. DNA-based analysis revealed that viral communities achieved a higher classification accuracy rate (97.42 ± 2.58%) than bacterial communities (96.13 ± 4.02%). Notable non-starter lactic acid bacteria and phages specific to each origin were identified. Volatile organic compounds exhibited potential clusters according to cheese origin, with a classification accuracy rate of 90.0 ± 11.11%. Near-infrared spectroscopy was used for cheese authentication, yielding a 76.0 ± 31.57% classification accuracy rate. The models' performances were influenced by specific regions of the infrared spectrum, possibly associated with fat content, lipid profile, and protein characteristics. Furthermore, we analyzed the elemental composition of mountain Caciotta cheese and identified significant differences in elements related to dairy equipment, macronutrients, and rare earth elements among different origins. The combination of elements and isotopes showed a decrease in authentication performance (97.00 ± 3.09%) compared to the original element models, which were found to achieve the best classification accuracy rate (99.00 ± 0.01%). Overall, our findings emphasise the potential of multi-omics techniques in cheese origin authentication and highlight the complexity of factors influencing cheese composition.

1. Introduction

Dairy products are among the most common products concerned by food frauds. In particular, protected land- and tradition-related labelled cheeses (*e.g.* Protected Designation of Origin or "mountain product") are subjected to food fraud (mainly mislabelling and fraudulent documentation) due to their high economic value. Indeed, consumers are more willing to pay higher prices for traditional and typical mountain cheeses due to their distinctive flavours and appearance but also their more natural and animal-friendly production attributes (Menozzi et al., 2022). For these reasons, geographical origin authentication methods are used to ensure product origin. In addition to chemical analyses, including stable isotope ratios, trace elements, and fatty acid profiles, emerging methods such as DNA-based methodologies and Near Infrared Spectroscopy are also investigated in this field (Cardin et al., 2022). However, to our knowledge, no comparative study has been conducted on the performances of chemico-physical analysis and DNA-based methods in authenticating cheese origin.

The stable isotope ratios of hydrogen ($\delta^{2}D$), carbon ($\delta^{13}C$), nitrogen ($\delta^{15}N$), oxygen ($\delta^{18}O$) and sulphur ($\delta^{34}S$) found in animals and their derived products like milk and cheese are reflective of those present in their diet and drinking water (Camin et al., 2008; Chesson et al., 2010; Pianezze et al., 2020). Due to the correlation between the isotopic ratios of diet constituents and water with the local environment, the analysis of isotope ratios has been widely used to differentiate cheese originating from distinct geographical areas (Bontempo et al., 2011; Camin et al., 2004; Camin et al., 2008; Camin et al., 2012). Element analysis or elemental metabolomics is related analysis in origin authentication, that involves the quantification and characterisation of the concentration of chemical elements in biological samples (Zhang et al., 2018). The origin of cheese has a strong correlation with its elemental composition, primarily influenced by geological and pedological traits, rather than being solely determined by factors like animal breed, feed vegetation, and mineral supplementation (Bontempo et al., 2011; Camin et al., 2012; Danezis et al., 2020). This is supported by the high performances obtained by this method in origin authentication such

as 92.1% of correct classification rate for Greek Graviera (Danezis et al., 2020) that made this analysis and its combination with stable isotope ratio the reference method for cheese origin authentication (Camin et al., 2012; Camin et al., 2015).

An emerging analysis in the origin authenticating field is Near Infrared Spectroscopy. Infrared analysis generates spectral data ranging from 800–2500 nm (near infrared) to 25–1000 μm (far-infrared) that contains diverse chemical and physical information (Lei and Sun 2019). This enables the use of Near Infrared Spectroscopy for authenticating the origin of cheeses, such as geographical origin, type of milk used, manufacturing process, quality parameters, composition data, and detecting any potential fraud through adulteration (Abbas et al., 2018; Medina et al., 2019).

Lastly, the other two prominent analytical methods in origin authentication are volatilome and microbiota analysis. Volatile organic compound (VOC) analysis has emerged as a prominent method for authenticating cheese origin due to the significance of volatile compounds in defining cheese typicity (Pillonel et al., 2003). These compounds are produced by the metabolic activities of cheese microbiota during glycolysis, proteolysis, and lipolysis, resulting in a diverse array of VOCs.

The metabolic activities of cheese microbiota are vital in shaping the distinct characteristics of cheese types. Cheese harbours a diverse microbiome that is composed of distinct and complex bacterial and viral communities. The diversity of microbial communities is influenced by the type of ecosystem in which they reside (Fierer and Jackson, 2006). Although cheese style is the primary predictor of rind microbiota (Wolfe et al., 2014), dairy farms and cheese-producing plants also play critical roles in defining cheese microbiota, which ultimately affects the quality of traditional cheeses (Goerges et al., 2008; Vacheyrou et al., 2011; Frétin et al., 2018). Microbial ecology studies have also highlighted how the combinations of different environmental factors, and cheese-making conditions and traditional know-how select specific microorganisms. Thus, DNA-based methods applied to microbiome analysis have also been suggested as potential tools to authenticate cheese geographical origin (Kamilari et al., 2019).

sPLS-DA (sparse partial least squares discriminant analysis) is a statistical method that can be used to analyze datasets with high dimensionality (large number of features) and identify patterns that can be used to differentiate between different groups or classes (Lê Cao et al., 2011). It is based on partial least squares (PLS) regression, which is a technique that is used to model the relationship between a set of predictor variables (also known as features or characteristics) and a response variable. sPLS-DA is particularly useful when there are more predictor variables than observations, which can be a common problem in -omics datasets. sPLS-DA works by projecting the data into a lower-dimensional space (called a latent space) in a way that maximizes the separation between the different classes or groups. This is achieved by minimizing the residuals between the observed response variables and the predicted response variables, while also maximizing the variance explained by the latent variables (Chin, 1998) . The resulting model can then be used to classify new observations into one of the pre-defined groups based on their feature values. sPLS-DA is a relatively simple and computationally efficient method that has been applied to a wide range of problems, including metagenomics, metagenetics and multi -omics data (Le Cao et al., 2016). sPLS-DA has the advantage of being able to handle high-dimensional data and missing values, and is relatively easy to interpret compared to some other machine learning techniques (Chung and Keles, 2010).

In this context, the aim of this study was to evaluate the discriminative power of DNA shotgun metagenomics (both bacterial and viral community profiling), volatilome, Near Infrared Spectroscopy, element metabolomics and stable isotopes ratio for cheese origin authentication. To do so, we used sPLS-DA models and adopted a case study approach to typical semi-hard raw milk Italian mountain cheeses (Caciotta).

2. Methods

2.1 Cheese sampling

In total, 42 Caciotta cheeses were collected in triplicate from five closely situated producers (within a range of 51 ± 26 km) located in the mountainous regions of Trentino Alto-Adige and Veneto. These regions include Alti Pascoli della Lessinia, Giudicarie esteriori, and Trento province areas (Figure 1). In particular, producer 1 was 38 km, 30 km, 55 km, and 53 km away from producers 2, 3, 4, and 5, respectively. The longest distances were observed for producer 2, who was 60 km, 92 km, and 94 km away from producers 3, 5, and 4, respectively. Lastly, producer 3 was 40 km and 42 km away from producers 4 and 5. Considering the origin altitudes, producers 2 and 3 were located at 1162 m and 1169 m above sea level, respectively, while producers 1, 4, and 5 were located at 628 m, 640 m, and 588 m above sea level, respectively.

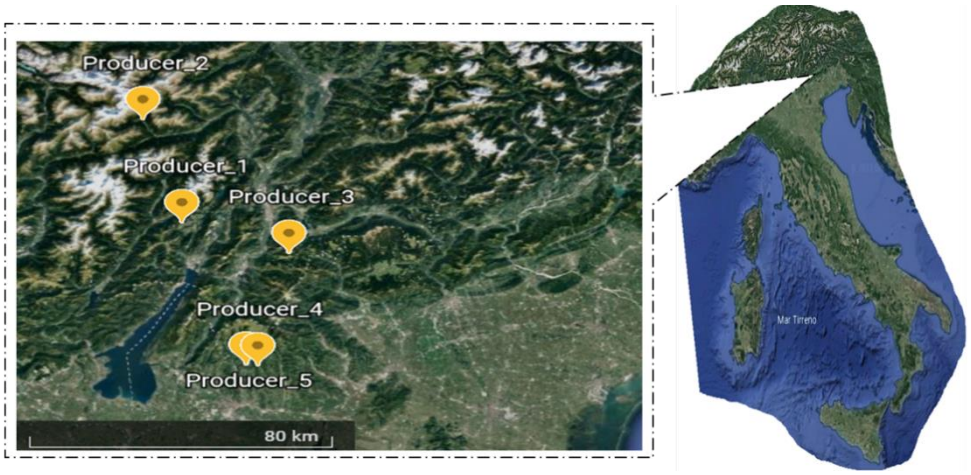


Figure 1. Map of the origin of mountain Caciotta producers. Producer 1 is located in Giudicarie Esteriori area while Producer 2 and Producer 3 are located in Trento province. Producer 4 and Producer 5 are located in Alti Pascoli della Lessinia area.

Cheeses, of 2020 and 2021, were sampled at the end of the ripening period (60±14 days). Approximately 500 g of cheese were sampled from each producer. The obtained sample was divided into aliquots; 300 g were employed for NIR spectroscopy analysis, 150 g were employed for stable isotope and trace elements analysis. The remaining core cheese was homogenised using sterile equipment and 2.5 g and 150 mg weighed for volatile fatty acid

and DNA analyses, respectively. Three technical replicates were obtained for each of the above-mentioned analyses. The samples were stored at -80°C until analysis.

2.2 Shotgun metagenomics

Shotgun metagenomics data obtained from the previous characterization of typical mountain Caciotta cheese were employed to develop predictive sPLS-DA model to authenticate cheese origin. Briefly, for both Caciotta cheese and starters, the total DNA was extracted from the samples using the DNeasy PowerSoil kit according to the manufacturer's instructions (Qiagen, Hilden, Germany). The quantity of DNA was assessed using a Qubit dsDNA HS Assay (Invitrogen, Life Technologies, Italy). Libraries were constructed with the Nextera XT DNA Sample Preparation Kit (Illumina, Inc., San Diego, USA) and IDT for Illumina Nextera DNA UD Indexes. The libraries were combined in equimolar amounts and assessed for quality and quantity using the Agilent 2100 Bioanalyzer and Qubit Assay Kit HS, respectively. The sequencing was performed by UC DAVIS Genome Center (California, US) on a NOVASEQ Sp500 platform, generating 1.1 billion reads for cheese and 340 million reads for starters. The FASTQC software was used to assess the quality of the raw reads (v.0.11.9, Brown et al., 2017), which were then processed with the bioBakery3 platform for quality control, contaminant depletion, and taxonomic assignment using KneadData and MetaPhlAn 3 (Beghini et al., 2021). KneadData was used to remove low-quality, repetitive sequence, and adapter sequences with a quality score cut-off of 35. High quality microbial reads were taxonomically profiled using MetaPhlAn3, an assembly free taxonomic profiler (Segata et al., 2012; Beghini et al., 2021). All raw sequence data in read-pairs format were deposited in the National Centre for Biotechnology Information (NCBI) in the Sequence Read Archive (SRA) under the project PRJNA922379 and PRJNA922380, for cheese samples and starter cultures, respectively.

2.3 Volatilome analysis

The HiSorb probes were used in conjunction with UNITY-xr, both from Markes International (UK), to perform headspace thermal desorption coupled with gas

chromatography mass spectrometry (5977B GC-MS Agilent Technologies, US). The samples of cheese were placed in 2.5 ml vials and headspace was sampled using an HiSorb Agitator (Markes International UK) at 40°C and 200 rpm for one hour. The probes were thermodesorbed using UNITY-xr at 280°C for 12 minutes, and a purge flow of 50 ml/min for 1 minute was used. The flow path was set at 200°C with a trap low of 25°C and a trap high of 290°C, and injection in the GC was performed using a low split 5 ml/min flow. A DB-5ms capillary column 60 m × 250 μm × 0,25 μm (Agilent Technologies, US) was used for the analysis. The oven temperature program was set to initial 40°C held for 2 min, then ramped 3°C/min up to 180°C, and again ramped 20°C/min up to 260°C for 5 min, and finally held for 6 min. The constant flow rate of helium carrier gas was set to 1 ml/min. The MS analyses were done in a full scan mode (TIC mode), with a scan range of 33 to 350 amu. To validate detected peaks, forty-four standard molecules were injected. The MassHunter quantitative analysis workstation (v.11.1, Agilent Technologies, US) was used for semi-quantitative analysis, with total peak area used for statistical analyses.

2.4 Near infrared spectral acquisition

A slice of cheese of 300 g was ground with a Retsch Grindomix (Retsch GmbH, Haan, Germany) at 4000 rpm for 10 s after removal of 2 cm of crust all around. Ground cheese samples were analysed in triplicate using a FOSS DS-2500 scanning monochromator (FOSS NIRSystem, Hillerød, Denmark). Scans were recorded in reflectance mode (850–2500 nm at 0.5-nm intervals) using a slurry cup with a quartz window (12.6 cm^2 area) in 30 g aliquots. Spectral data were recorded as absorbance (A) calculated as log (1/R), where R represents reflectance, using WinISI4 software V4.10.0.15326 (FOSS Analytical A/S, Hillerød, Denmark). Before statistical analysis, spectra were exported to an Excel (Microsoft Office®, USA) spreadsheet and averaged before further chemometric modelling.

2.5 Trace element analysis

About 0.5 g of fresh cheese were acid digested using an UltraWAVE System (Milestone, Shelton, CT, USA) equipped with PTFE vials and following the method reported in Muňoz-

Redondo et al. (2022). The mineral element profile (Li, Be, B, Na, Mg, Al, P, K, Ca, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As, Se, Rb, Sr, Y, Mo, Pd, Ag, Cd, Sn, Sb, Te, Cs, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Dy, Ho, Er, Tm, Yb, Re, Hg, Tl, Pb, Bi and U) was quantified using an ICP-MS (Agilent 7800, Agilent Technologies, Tokyo, Japan). Helium was used as collision gas in the Octopole Reaction System for the effective analysis of Na, Mg, K, Ca, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As and Pd with a flow of 5 ml/min while for Se, Sn, Eu a flow of 10 ml/min was used. Instrumental parameters were optimized to maximize sensitivity and reduce spectral interferences at each analytical batch following manufacturing guidelines. A solution of Sc, Rh and Tb was used as internal standard online for the correction of signal drift whereas a solution of in a known concentration was added to each sample before mineralization for volume correction. Each batch included, together with samples, a blank sample (only reagents) to assure cleanliness. The accuracy was verified using samples spiked with a known amount of a standard solution was used. The calculated recoveries ranged from 83 to 117% and were considered acceptable for the aim of this research. For the determination of the limits of quantification (LOQ), 10 blank samples were prepared and analyzed in a sequence and the calculated standard deviation obtained for each element was multiplied by 10. All the materials in contact with standard or samples during mineralization and analysis were washed with a 5% HNO_3 solution and rinsed with ultrapure water (18.2 MΩ-cm, Millipore, Bedford, MA, USA). The moisture of each sample was quantified as the loss of weight in oven at 105°C. Elemental contents were expressed as percentage of dry matter (d.m.).

2.6 Stable isotopes ratio

The fresh cheeses were lyophilised. 4 g of freeze-dried cheese was extracted 3 times with 30 mL of a petroleum ether: diethyl ether (2:1) mixture, homogenising with an Ultraturrax device (T25, IKA, IKA GmbH, Staufen, Germany) (11500 rpm for 3 minutes) and using a centrifuge (e.g. 4100 rpm for 6 min) to separate the ether from the residue. A Soxhlet extractor (Merck KGaA, Darmstadt, Germany) was used as an alternative to extract fat. After lipid extraction, the skimmed cheese was warmed to 40 °C to remove any possible

residual ether. Then the residue was washed twice with 20 mL of water, centrifuging each time (e.g. 4100 rpm for 3 min). The residue, made up mainly of casein, was lyophilised and conserved at room temperature until analysis.

All samples were weighted in silver and tin capsules for OH− and CNS− isotope measurements, respectively. The $^{15}N/^{14}N$, $^{13}C/^{12}C$ and $^{34}S/^{32}S$ ratios were determined using an isotope ratio mass spectrometer (IRMS) (Isoprime, AP2003, GV Instruments Ltd, Manchester, UK) equipped with an elemental analyser (Vario EL III Elementar Analysensysteme GmbH, Hanau, Germany), while the $^{18}O/^{16}O$ and $^{2}H/^{1}H$ ratios were determined with an IRMS (Flash EA1112, Thermo Fisher Scientific, Bremen, Germany) equipped with a pyrolizer (TC/EA, Thermo Fisher Scientific, Bremen, Germany).

In agreement with the IUPAC protocol (Brand et al., 2014) , the isotopic values are expressed in delta in relation to the international standard V−PDB (Vienna−Pee Dee Belemnite) for $\delta(^{13}C)$, V−SMOW (Vienna−Standard Mean Ocean Water) for $\delta(^{2}H)$ and $\delta(^{18}O)$, V−CDT (Vienna−Canyon Diablo Troilite) for $\delta(^{34}S)$ and Air (atmospheric N_2) for $\delta(^{15}N)$, following equation 1:

$$\delta_{ref}({}^{i}E/{}^{j}E, sample) = \left[\frac{R({}^{i}E/{}^{j}E,\ sample)}{R({}^{i}E/{}^{j}E,\ ref)}\right] - 1 \quad (1)$$

where ref is the international measurement standard, sample is the analysed sample and $^{i}E/^{j}E$ is the isotope ratio between heavier and lighter isotopes. The delta values are multiplied by 1000 and expressed commonly in units "per mil" (‰) or, according to the International System of Units (SI), in unit 'milliurey' (mUr).

The isotopic values were calculated against two standards through the creation of a linear equation. The standards that have been used in the isotopic analyses were international reference materials or in−house working standards that have been calibrated against them. In particular, the international standards that have been used are: for $^{13}C/^{12}C$, fuel oil NBS−22 ($\delta(^{13}C)$=−30.03±0.05‰), sucrose IAEA−CH−6 ($\delta(^{13}C)$=−10.45±0.04‰) (IAEA−International

Atomic Energy Agency, Vienna, Austria), and L−glutamic acid USGS 40 ($\delta(^{13}C)$=−26.39±0.04‰) (U.S. Geological Survey, Reston, VA, USA); for $^{15}N/^{14}N$, L−glutamic acid USGS 40 ($\delta(^{15}N)$=−4.52±0.06‰) (U.S. Geological Survey, Reston, VA, USA), ammonium sulfate salts IAEA−N−1 ($\delta(^{15}N)$=+0.43±0.07‰) and IAEA−N−2 ($\delta(^{15}N)$=+20.41±0.12‰) and potassium nitrate IAEA−NO3 ($\delta(^{15}N)$=+4.7±0.2‰); for $^{34}S/^{32}S$, USGS 42 ($\delta(^{34}S)$=+7.84±0.25‰), USGS 43 ($\delta(^{34}S)$=+10.46±0.22‰), Barium sulphate IAEA−SO−5 ($\delta(^{34}S)$=+0.5±0.2‰) and NBS 127 ($\delta(^{34}S)$=+20.3±0.4‰); for $^{2}H/^{1}H$ fuel oil NBS−22 $\delta(^{2}H)$=−119.6±0.6‰) and Keratins CBS (Caribou Hoof Standard $\delta(^{2}H)$=−157±2‰) and KHS (Kudu Horn Standard $\delta(^{2}H)$=−35±1‰) from U.S. Geological Survey; for $^{18}O/^{16}O$ benzoic acid IAEA 601 ($\delta(^{18}O)$=+23.14±0.19‰) and benzoic acid IAEA 602 (δ(18O)=+71.28±0.36‰) from IAEA.

Each reference material was measured in duplicate at the start and end of each daily group of analyses of samples (each sample was also analysed in duplicate). A control material was also included in the analyses of each group of samples, to check the efficiency of the measure. The maximum standard deviations of repeatability accepted were 0.3‰ for $\delta(^{13}C)$ and $\delta(^{15}N)$, of 0.4‰ for $\delta(^{34}S)$, 0.5‰ for $\delta(^{18}O)$ and of 3‰ for $\delta(^{2}H)$.

2.7 Statistical analysis

Statistical analysis of microbial communities, volatile organic compounds, Near Infrared Spectroscopy, trace elements analysis and stable isotopes ratio were performed in R version 4.0.5 (R Foundation for Statistical Computing, Vienna, Austria). Significance of the median was obtained with Kruskal-Wallis Rank Sum Test while multiple comparison was performed with pairwise Wilcoxon Rank Sum Tests. Obtained *p*-values were adjusted with Benjamini & Hochberg (1995) correction method. Trace elements distribution presented challenging problems connected with LOQ. In this instance we employed missing value substitution according to the results presented by Farnham et al. (2002). LOQ was divided by two to insert missing values.

Unsupervised principal component analysis was used to investigate the datasets. sPLS-DA was employed to evaluate authentication performance (Le Cao et al., 2011). Centred log-

ratio transformed relative abundance for bacterial and viral communities, peak abundances for volatilome analysis, wavelength absorbance for near infrared spectroscopy, elements and isotope concentration for trace elements and isotopes analysis were used with sPLS-DA model using mixOmics R package (Rohart et al., 2017). To avoid sampling biases and represent the true performances of the models, each dataset was sampled 10 times obtaining representative origins class for both training (75% of samples) and test (25% of samples) dataset. M-fold cross-validation (*i.e.* the process of dividing of dataset into M subsets and then, iteratively, using some of them to train the model while exploiting the others to evaluate its performance) of the training dataset was performed using 10 folds and 100 repeats. The obtained models were used to predict the origin of the tested dataset. Average predictive performances were compared considering true positive (TP), true negative (TN), false positive (FP) and false negative (FN) ratios expressed through recall (Figure 2, Equation 2), precision (Figure 2, Equation 3), specificity (Figure 2, Equation 4) and classification accuracy rate (Figure 2, Equation 5) (Kassambara 2018; Bisutti et al., 2019). Figure 2, summarise the procedure used to obtain cross-validated models.

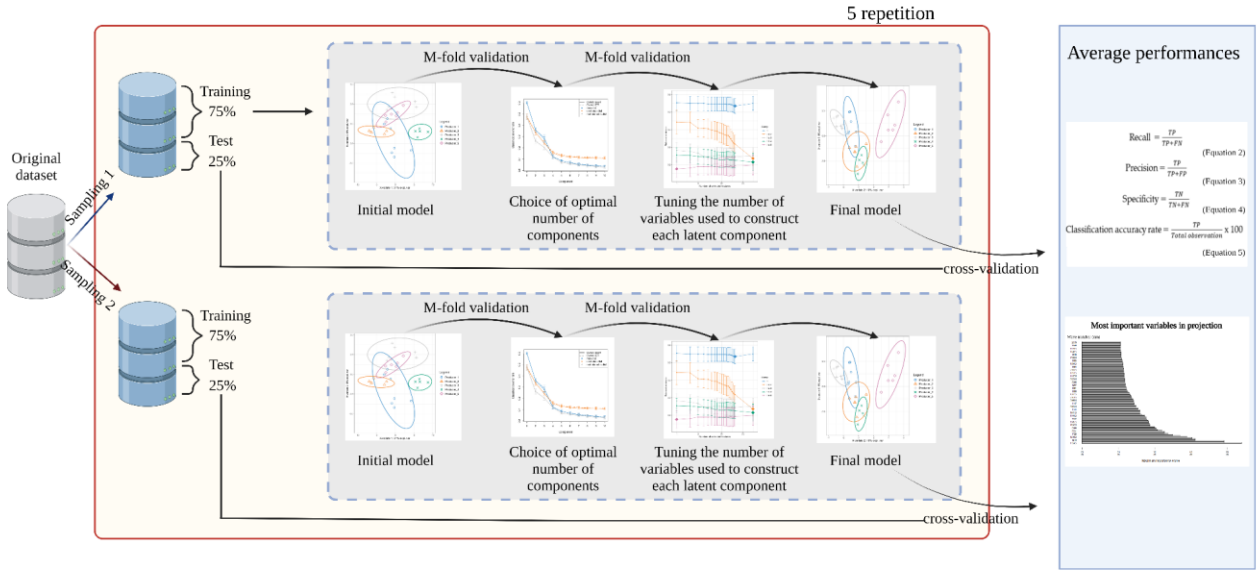


Figure 2. Procedure used to obtain and compare cross-validated sPLS-DA models. Each dataset was sampled 10 times obtaining representative origins class for both training (75% of samples) and test (25% of samples) dataset.

Recall, precision, specificity and classification accuracy rate coefficient are commonly used performance metrics to compare the performance of predictive models. Each metric captures different aspects of the model's performance, and they are used in combination to provide a comprehensive evaluation of the model's effectiveness. Recall measures the proportion of true positive cases that were correctly identified by the model. Precision evaluates the proportion of true positive cases among all cases predicted as positive by the model. Specificity measures the proportion of true negative cases that were correctly identified by the model. It is a useful metric when the cost of a false positive error is high. Classification accuracy rate measures the proportion of correct predictions made by the model.

In sPLS-DA models, the loading is a vector that represents the relationship between the predictor variables (*e.g.* taxa relative abundance) and the response variable (*e.g.* origin) that defines the class labels. The loading was calculated during the model fitting process and used to identify the variables that contribute the most to the separation of the classes in the data. The average loading importance was computed by extracting the loading value from

each model and then calculating the absolute value of each. Finally, the average of these absolute values was used to represent the 25 most important variables for the prediction.

3. Results

3.1 Bacterial and viral communities

Caciotta microbiota showed a complex profile for both bacterial and viral communities. The univariate and multivariate analysis of the microbiota of typical Caciotta cheese has been previously outlined, including the results of the Kruskal-Wallis test, visual representation through non-metric multidimensional scaling, and permutational multivariate analysis of variance. Briefly, 45 bacterial and 44 viral species had significantly different mean relative abundance according to their origin. Starter (*Steptococcus thermophilus* and *Lactobacillus delbrueckii*) and non-starter lactic acid bacteria (*e.g. Lactococcus raffinolactis, Lentilactobacillus parabuchneri, Lactiplantibacillus paraplantarum,* and *Propionibacterium freudenreichii*) showed significant relative abundance differences. Some secondary microbiota components were only found in specific origins and a clear link between the cheese and its producer for several *S. thermophilus* strains was observed. Starter lactic acid bacteria were the most abundant, with associated bacteriophages like Streptococcus phage TP 778L, Streptococcus virus DT1, Streptococcus virus phiAbc2, and Lactobacillus phage A2 exhibiting the higher abundance and significant variations based on Caciotta origin. Both bacterial and viral communities formed clusters according to producer location. However, viral communities showed narrower clusters for each cheese origin and demonstrated significant differences in permutation analysis of variance. Bacterial communities tended to cluster based on regional area (*i.e.* producer 1, 2, 3, were separated from producer 4 and 5) as confirmed by permutation analysis.

The potential of bacterial communities for authenticating Caciotta origin was evaluated using sPLS-DA models. The training models for the bacterial communities utilised a variable number of components, ranging from 9 to 15 and achieved a correct classification accuracy rate 99.89 ± 0.20%. The cross-validated outcomes are presented in Table 1.

Table 1. Summary table of bacterial communities' sPLS-DA models performance based on true positive (TP), true negative (TN), false positive (FP) and false negative (FN) outputs. Recall, precision, specificity, and classification accuracy rate have been calculated according to Kassambara (2018).

Origin/Parameter	Recall	Precision	Specificity	Classification accuracy rate (%)
Producer 1	0.92	0.88	0.98	
Producer 2	0.97	0.98	0.99	
Producer 3	0.91	0.98	0.98	96.13 ± 4.02
Producer 4	1.00	0.98	1.00	
Producer 5	1.00	0.96	1.00	

Cross-validated models exhibit a slight decrease in performances compared to the training ones. Generally, recall, precision, and specificity presented high values for bacterial communities of Caciotta cheese. Producer 1 showed the lowest performances in origin authentication driven by the lowest value of precision while producer 3 had the lowest recall. Overall bacterial communities yielded a high value of classification accuracy rate 96.13 ± 4.02%. The most important bacterial species in Caciotta origin authentication are shown in Figure 3.

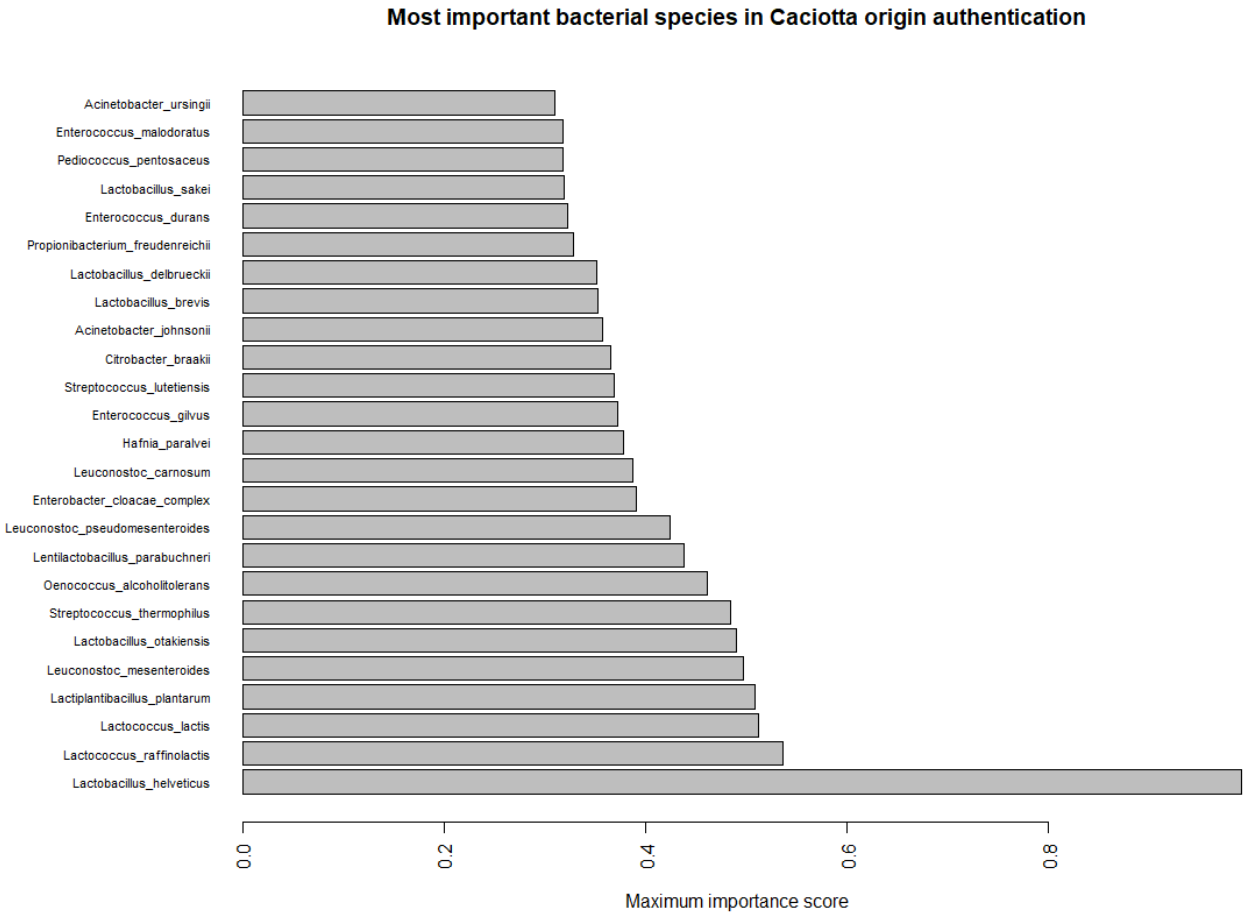


Figure 3. Mean maximum importance score of the most important 25 bacterial species used in the sPLS-DA models.

Most of the taxa with the highest scores were attributed to lactic acid bacteria. Among them the majority showed a significant difference in relative abundance connected to origin. Moreover, as previously noticed, origin specific non-starter lactic acid bacteria like *Lactobacillus helveticus*, *Lactococcus raffinolactis* and *Propionibacterium freudenreichii* presented the highest score for origin authentication.

Viral communities constituted the second portion of the microbiota investigated to authenticate mountain Caciotta cheese. The training models for viral communities presented a lower number of components, from 7 to 12, and higher classification accuracy rate (99.92±0.16%) than the bacterial one. Also the cross-validated performances showed

higher values of recall, precision, specificity and classification accuracy rate than the bacterial one (Table 2).

Table 2. Summary table of viral communities' sPLS-DA models performance based on true positive (TP), true negative (TN), false positive (FP) and false negative (FN) outputs. Recall, precision, specificity, and classification accuracy rate have been calculated according to Kassambara (2018).

Origin/Parameter	Recall	Precision	Specificity	Classification accuracy rate (%)
Producer 1	1.00	0.91	1.00	
Producer 2	0.95	0.98	0.98	
Producer 3	0.97	0.98	0.99	97.42 ± 2.58
Producer 4	1.00	0.98	1.00	
Producer 5	0.95	1.00	0.98	

Overall, viral communities models showed the highest performances for the authentication of producer 1, 4 and 5, while the lowest performances were observed for producer 2. The obtained classification accuracy rate was higher than bacterial communities yielding a correct classification of 97.42 ± 2.58% of the samples. The most important viral species used in the loadings of the model are shown in Figure 4.

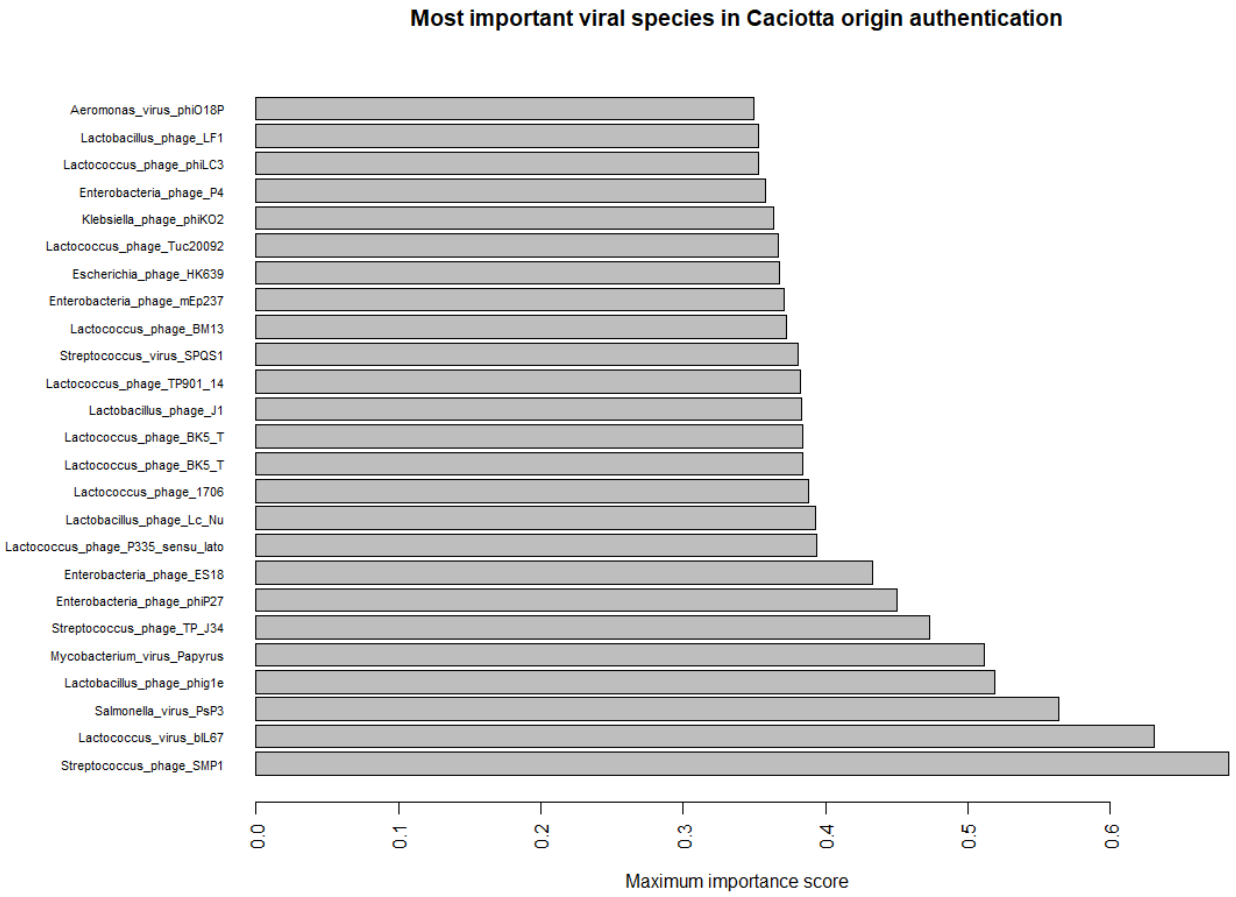


Figure 4. Mean maximum importance score of the most important 25 bacterial species used in the sPLS-DA models.

Generally, Streptococcus phages, Lactoccocus virus, Lactobacillus phages and Lactococcus phages were the taxa obtaining the highest score for origin prediction. Other important taxa were Salmonella virus, Mycobacterium virus and viruses affecting bacteria from *Enterobacteriaceae* family.

3.2 Volatilome

The analysis of volatile organic compounds in typical Caciotta cheese (both univariate and multivariate approaches) were already described in our previous work. These analyses included outcomes from the Kruskal-Wallis test, visual representations via principal components analysis, and permutational multivariate analysis of variance. Mountain Caciotta cheese presented a complex volatilome from which prominent levels of alcohols

and ketones, accompanied by lower proportions of terpenes were observed. Cheese origin led to significant variations in the relative abundances of most investigated volatile organic compounds (*i.e.* ethanol, butan-2-ol, 3-methylbutan-1-ol, benzaldehyde, ethyl acetate, 3-methylbutyl acetate, butan-2-one, heptan-2-one, octanoic acid, D-limonene) with the exclusion of 3-hydroxybutan-2-one and 1-acetophenyletanone. Additionally, certain terpenoids like *p*-cymene and 3-carene were not consistently found across all origins. The pairwise permutational analysis of variance revealed that each producer of typical mountain Caciotta cheese exhibited a distinctive VOC profile, varying to some degree from one another.

The training models for the volatile organic compounds used a variable number of components, ranging from 9 to 15 and achieving a correct classification accuracy rate 99.87 ± 0.18%. The cross-validated outcomes are presented in Table 3.

Table 3. Summary table of VOCs sPLS-DA models performance based on true positive (TP). true negative (TN), false positive (FP) and false negative (FN) outputs. Recall, precision, specificity, and classification accuracy rate have been calculated according to Kassambara (2018).

Origin/Parameter	Recall	Precision	Specificity	Classification accuracy rate (%)
Producer 1	0.98	0.82	0.99	
Producer 2	0.9	0.98	0.97	
Producer 3	0.86	0.91	0.96	90.0±11.11
Producer 4	0.93	0.82	0.98	
Producer 5	0.86	0.98	0.96	

Producer 2 showed the highest values of recall, precision and specificity compared to the other origins. Producers 1 and 4 had the lowest values of precision while producers 5 and 3 had the lowest values of recall. Overall, high specificity values were observed for all the origins. The cross-validated model obtained a high value of correct accuracy rate 90.0±11.11% but it was characterized by a high accuracy error as well. Figure 5 reports the

mean maximum score of the most important volatiles used in Caciotta origin authentication.

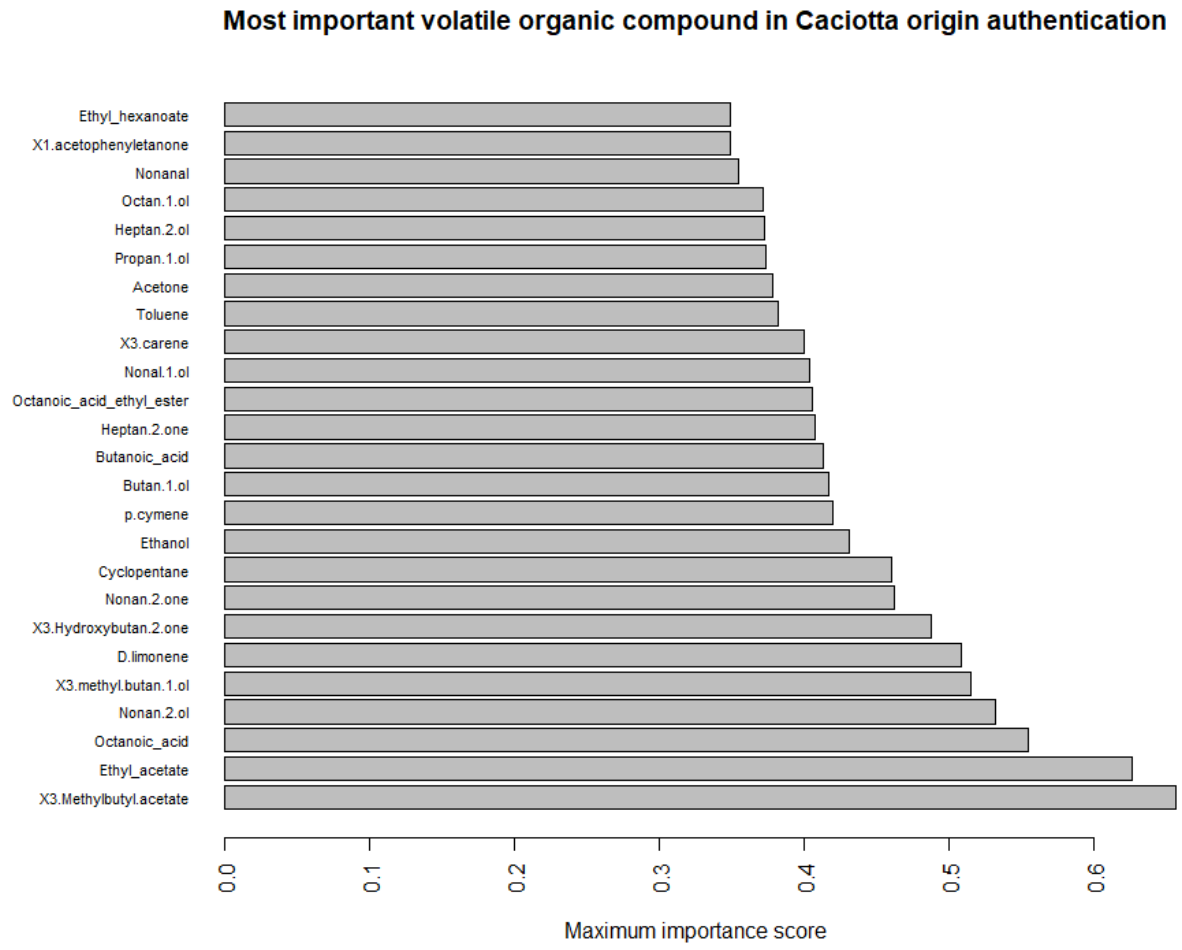


Figure 5. Mean maximum importance score of the most important 25 VOCs used in the sPLS-DA models.

Among the most important VOCs for authenticating mountain Caciotta origin, alcohols and ketones were the most prevalent classes, followed by esters, terpenes, fatty acids, and hydrocarbons. Esters such as 3-methylbutyl acetate and ethyl acetate showed the highest scores. Terpenes like D-limonene, p-cymene, and 3-carene significantly contributed to the model's performance, along with alcohols such as nonan-2-ol, 3-methylbutan-2-ol, and ethanol. Surprisingly, also 3-hydroxybutan-2-one and 1-acetophenyletanone were identified as important compounds for the authentication models.

3.3 Near infrared spectroscopy

Absorbance values of typical Caciotta cheese were firstly investigated with principal components analysis (Figure 6).

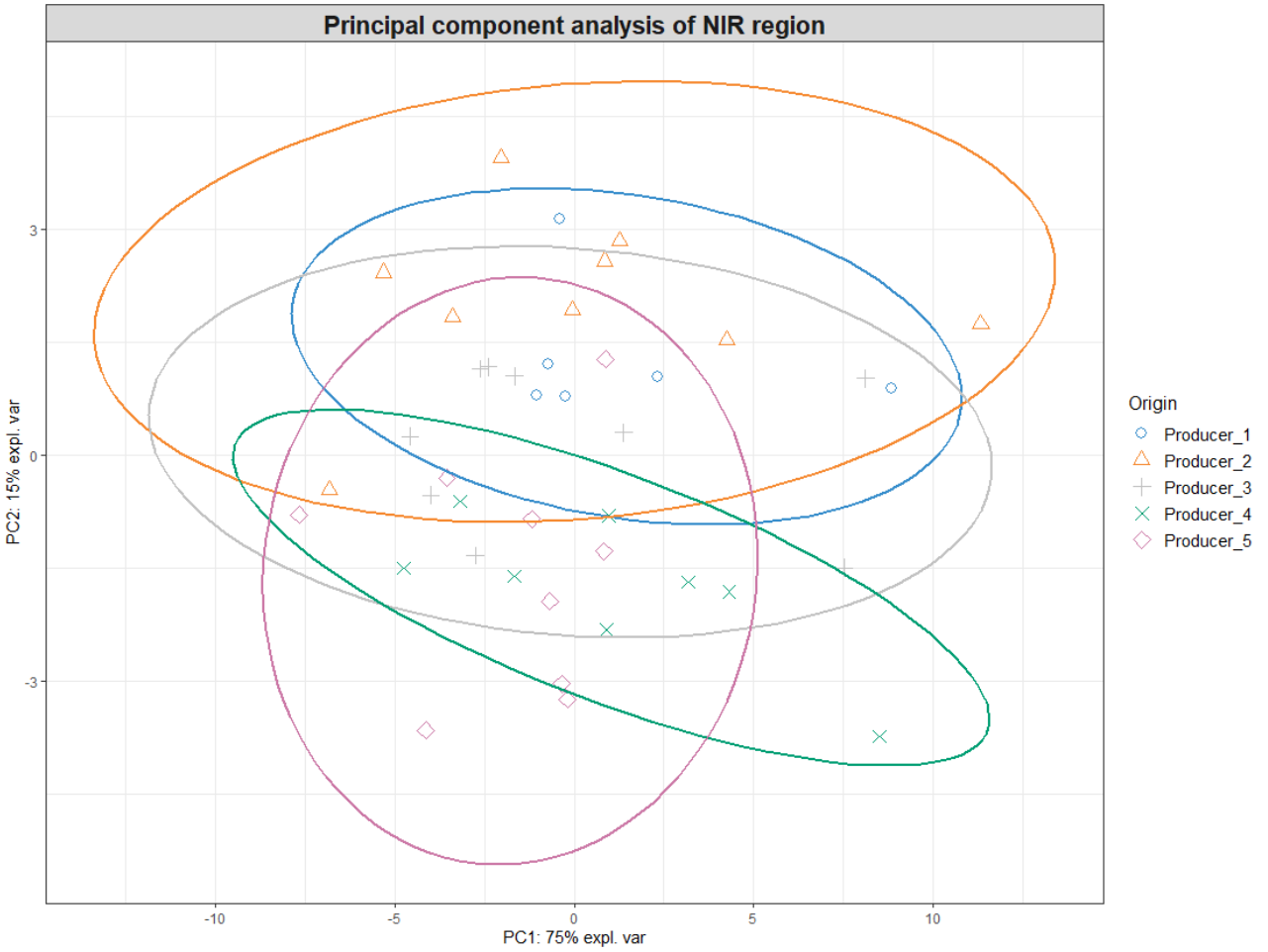


Figure 6. Principal component analysis of Caciotta's absorbance in the Near Infrared region (850–2500 nm).

Even though the unsupervised analysis of Caciotta cheese explained 90% of the variance in the first and second components, it did not reveal any clusters of samples related to their origin. The ability of NIRs data to authenticate Caciotta origin was further investigated with sPLS-Da models. The training models exhibited a wide range of components, varying from 3 to 14. Nonetheless, they displayed encouraging classification performance, resulting in an average classification accuracy rate of 92.67±7.93%. Table 4 reports the average performance of the cross-validated models.

Table 4. Summary table of NIRs sPLS-DA models performance based on true positive (TP), true negative (TN), false positive (FP) and false negative (FN) outputs. Recall, precision, specificity, and classification accuracy rate have been calculated according to Kassambara (2018).

Origin/Parameter	Recall	Precision	Specificity	Classification accuracy rate (%)
Producer 1	0.95	0.68	0.98	
Producer 2	0.95	1.00	0.98	
Producer 3	0.55	0.65	0.88	76.0 ± 31.57
Producer 4	0.7	0.64	0.91	
Producer 5	0.65	0.93	0.9	

Cross-validated models exhibit a considerable decrease in performances compared to the training one. Producer 3 reported the lowest values of recall, precision and specificity, that connected to the precision and recall of producer 4 and 5 lead to a classification accuracy rate of 76 ± 31.57%. This model exhibited a high error in accuracy, which could be attributed to the low precision values observed for producer 1, 3, and 4. Given that a singular wavelength lacks significant information in comparison to NIR analysis regions, we assessed the 100 most crucial wavelengths for predicting the origin of Caciotta (Figure 7).

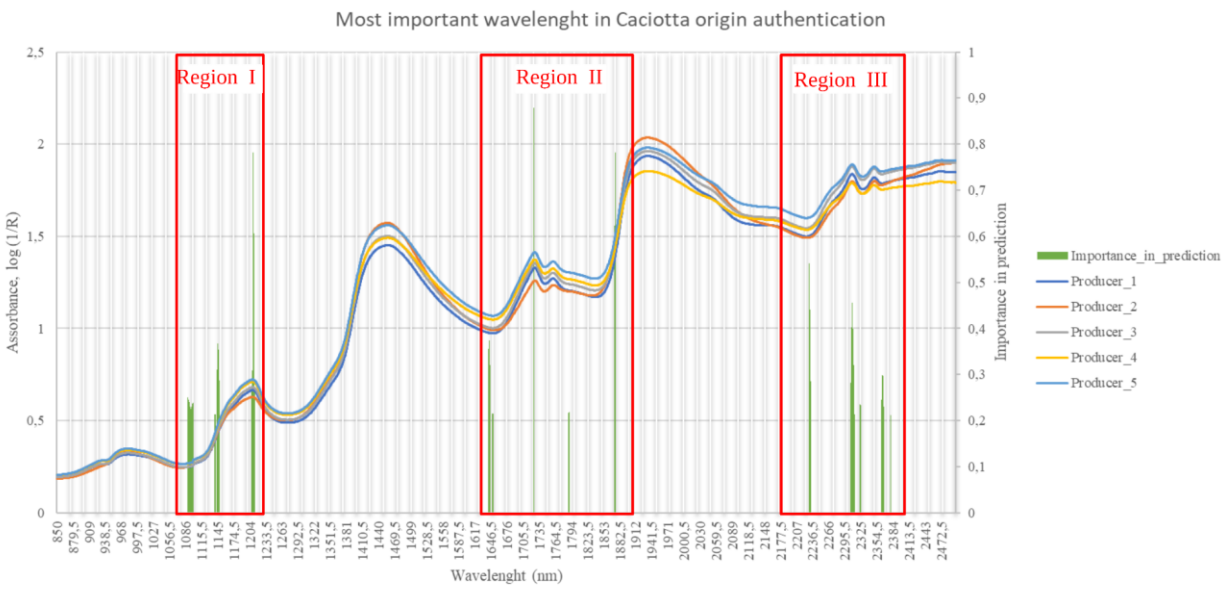


Figure 7. Average absorbance spectra of typical Caciotta cheese based on origin. Green bars, indicating the average importance value in prediction, were grouped in Region I, II and III.

The wavelengths with the highest mean maximum importance score in the prediction of Caciotta origin formed three regions. Region I and III were characterized by three clusters around 1090, 1145, 1204 and 2230, 2310 and 2360 nm, respectively. Finally, Region II presented scattered important wavelengths around 1646, 1735, 1780 and 1870 nm.

3.4 Element analysis

The elemental analysis of Caciotta cheese showed different concentration profiles for 39 elements while Be, Ga, Ge, Pd, Ag, Sn, Te, Dy, Ho, Tm, Hg and Tl presented concentrations inferior to LOQ for 90% of observations were excluded from the analysis. Table 5 reports elements concentration expressed as dry matter to correct the effects of different degrees of moisture and ripening. The results of the univariate analysis showed significant variations in element concentrations across different producers of Caciotta cheese.

Table 5. Pairwise Wilcoxon Rank Sum of element concentration in Caciotta cheese.

Element	Producer 1	Producer 2	Producer 3	Producer 4	Producer 5
Al (ug/kg dm)	324.92±21,62^{a}	504.15±81,75^{ab}	143.47±12.88^{c}	372.86±26.33^{ab}	938.35±300.06^{b}
As (ug/kg dm)	0.5±0^{a}	1.05±0.28^{a}	0.5±0^{a}	0.5±0^{a}	1.20±0.34^{a}
B (ug/kg dm)	134.28±35.58^{ab}	223.79±27.22^{a}	143.52±11.94^{ab}	125.74±23.08^{b}	177.41±28.89^{ab}
Ba (ug/kg dm)	1556.72±14.52^{a}	1366.10±193.27^{ab}	975.13±87.36^{bc}	963.81±23.08^{c}	1045.59±111.25^{bc}
Bi (ug/kg dm)	0.1±0^{a}	0.1±0^{a}	3.11±0.86^{b}	0.1±0^{a}	0.1±0^{a}
Ca (g/kg dm)	12.31±0.08^{a}	10.84±0.29^{ab}	12.42±0.11^{a}	11.59±0.13^{a}	10.60±0.26^{b}
Cd (ug/kg dm)	0.31±0.02^{a}	0.37±0.02^{a}	0.32±0.01^{a}	0.30±0.01^{a}	0.32±0.01^{a}
Ce (ug/kg dm)	0.25±0.04^{a}	0.36±0.06^{ab}	0.04±0^{c}	0.32±0.04^{ab}	0.69±0.22^{b}
Co (ug/kg dm)	0.752±0.01^{a}	1.52±0.09^{b}	1.40±0.12^{b}	1.85±0.09^{c}	1.35±0.17^{b}
Cr (ug/kg dm)	5.32±0.46^{a}	53.00±4.48^{b}	10.56±1.76^{cd}	7.38±0.76^{ac}	16.78±0.74^{d}
Cs (ug/kg dm)	8.25±1.16^{a}	10.98±1.43^{a}	6.35±0.77^{a}	1.62±0.07^{b}	3.24±0.74^{b}
Cu (ug/kg dm)	12240.97±1096^{a}	496.19±131.38^{b}	307.96±21.92^{b}	6078.50±221.11^{c}	250.47±23.46^{b}
Er (ug/kg dm)	0.02±0^{a}	0.03±0.006^{a}	0.02±0^{a}	0.02±0^{a}	0.04±0.01^{a}
Eu (ug/kg dm)	0.09±0.006^{a}	0.06±0.01^{ab}	0.03±0^{b}	0.03±0^{b}	0.07±0.01^{ab}
Fe (ug/kg dm)	1036.84±36.94^{a}	1687.78±120.29^{b}	1078.63±82.21^{a}	1148.50±50.54^{a}	1511.33±122.53^{b}
Gd (ug/kg dm)	0.03±0^{a}	0.03±0^{a}	0.03±0^{a}	0.03±0^{a}	0.08±0.02^{a}
K (mg/kg dm)	1651.22±44^{ab}	2424.96±180^{c}	2107.02±159^{ac}	1529.64±23^{d}	1637.53±146^{bd}

La (ug/kg dm)	0.12±0.01[a]	0.21±0.04[ab]	0.07±0.47[c]	0.23±0.03[ab]	0.39±0.11[b]
Li (ug/kg dm)	2.75±0.16[a]	18.99±6.14[b]	1.36±0.47[a]	1.65±0.01[a]	3.03±0.87[ab]
Mg (mg/kg dm)	0.43±0.01[ab]	0.49±0.01[a]	0.49±0.01[a]	0.44±0.01[a]	0.40±0.01[b]
Mn (ug/kg dm)	198.15±5.99[a]	284.46±23.13[b]	262.56±7.00[b]	235.19±13.97[b]	213.28±22.63[ab]
Mo (ug/kg dm)	252.45±12.49[a]	402.35±32.77[b]	249.34±12.00[a]	251.97±14.52[a]	257.83±10.08[a]
Na (g/kg dm)	8.09±0.30[a]	11.69±0.80[b]	8.36±1.25[ab]	9.12±0.12[a]	9.31±0.49[ab]
Nd (ug/kg dm)	0.11±0.01[ab]	0.15±0.02[ac]	0.03±0[b]	0.14±0.17[ac]	0.34±0,12[c]
Ni (ug/kg dm)	1.25±0[a]	20.62±2.18[b]	3.58±1.07[ac]	4.33±0.61[c]	5.73±0.66c
P (g/kg dm)	7.34±0.10[a]	6.58±0.15[ab]	7.31±0.16[ab]	7.20±0.19[ab]	6.83±0.09[b]
Pb (ug/kg dm)	6.64±0.10[a]	2.59±0.27[b]	3.23±0.15[b]	3.55±0.25[b]	10.81±1.96[a]
Pr (ug/kg dm)	0.02±0[a]	0.04±0.007[b]	0.02±0[a]	0.02±0[a]	0.08±0.03[b]
Rb (ug/kg dm)	3375.65±276[a]	3888.21±626[ab]	5207.71±761[b]	1298.42±48[c]	1479.29±259[c]
Re (ug/kg dm)	0.35±0.06[a]	0.57±0.21[a]	0.06±0.004[bc]	0.03±0[b]	0.17±0.05[ac]
Sb (ug/kg dm)	0.21±0.01[ab]	0.21±0.01[ab]	0.18±0.01[a]	0.18±0.01[a]	0.28±0.01[b]
Se (ug/kg dm)	93.40±8.13[ab]	160.86±23.89c	117.88±3.91[a]	94.56±5.71[b]	109.55±13,7[ab]
Sm (ug/kg dm)	0.03±0[a]	0.03±0[a]	0.03±0[a]	0.03±0[a]	0,087±0.02[b]
Sr (ug/kg dm)	3803.73±336[ab]	5359.01±271[c]	2789.52±41[a]	2848.11±52[a]	3641.18±257[b]
U (ug/kg dm)	0.22±0.03[a]	0.09±0.02[ab]	0.15±0.02[b]	0.08±0.01[b]	0.35±0,04[a]
V (ug/kg dm)	0.25±0[a]	2.07±0.71[b]	0.79±0.05[ab]	0.57±0.12[a]	1.43±0.58[ab]
Y (ug/kg dm)	0.11±0.002[a]	0.15±0.03[ab]	0.04±0[c]	0.33±0.05[b]	0.32±0.08[b]

Yb (ug/kg dm)	0.02±0[b]	0.04±0.004[a]	0.02±0[b]	0.02±0[b]	0.04±0.01[a]
Zn (mg/kg dm)	43.03±1.19[a]	43.42±1.24[ab]	43.07±1,37[a]	51.85±1.20[c]	38.29±0.86[b]

Each producer exhibits unique elemental signatures. The major differences were observed for Producers 1, 2, and 5, who presented significantly different concentrations of several elements compared to the other producers. We categorise the elements into four groups according to their potential ties to dairy equipments (*i.e.* Fe, Cu, Al, Ni, Cr), their classification as macronutrient (*i.e.* P, K, Ca, Mg, Na) or rare earth (*i.e.* Ce, Er, Eu, La, Yb), and "other elements". Among the elements associated with dairy equipment, Producer 1 exhibited the highest concentration of copper (Cu), while Producer 2 had the highest concentrations of iron (Fe) and nickel (Ni). Producer 4 showed the highest concentration of chromium (Cr), and Producer 5 had the highest concentration of aluminium (Al). In the macronutrients class, phosphorus (P) showed similar concentrations among producers, except for Producer 1 and Producer 5, which had the highest and lowest concentrations, respectively. Producer 2 had the highest sodium (Na) concentration compared to Producers 1 and 3. Additionally, Producer 5 had the lowest concentrations of magnesium (Mg) and calcium (Ca). Among the rare elements, Producer 3 exhibited the highest cerium (Ce) concentration in comparison to Producers 1, 2, and 4 and Producer 5 showed significantly higher concentrations of europium (Eu) compared to Producers 1, 2, and 3. There were significant differences in concentration for the "other elements" as well. Producer 3 had the highest levels of manganese (Mn) and rubidium (Rb). Producers 1 and 5 exhibited higher concentrations of certain elements such as barium (Ba), lead (Pb), neodymium (Nd) and uranium (U). On the other hand, Producers 4 and 5 showed lower concentrations of cesium (Cs). Additionally, Producers 2 and 5 demonstrated higher concentrations of praseodymium (Pr) and strontium (Sr).

We further examined the elemental composition of Caciotta cheese with unsupervised multivariate analysis based on principal component analysis (Figure 8).

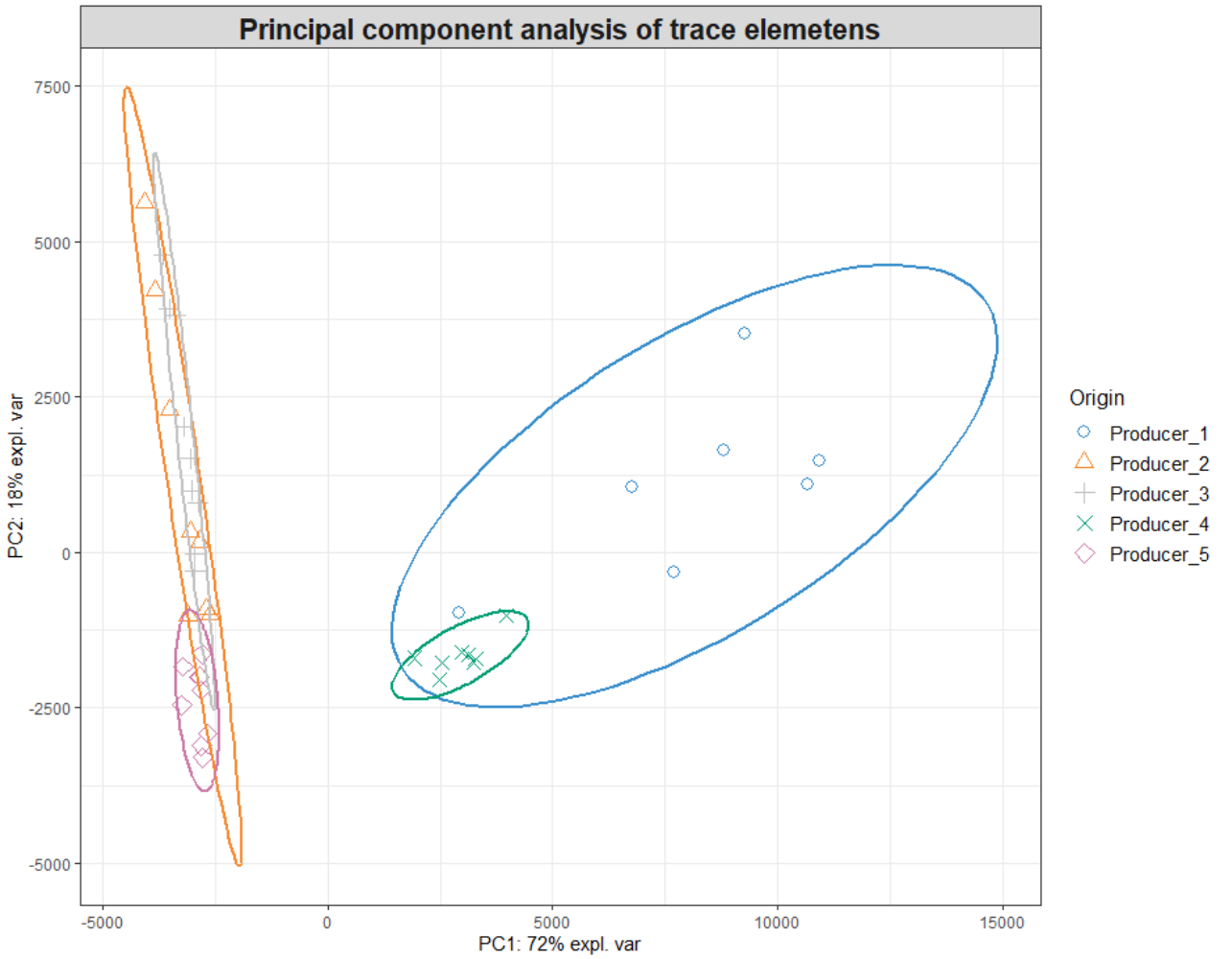


Figure 8. Principal component analysis of Caciotta elemental composition.

Caciotta cheeses from the same origin tended to cluster together. However, the confidence intervals of producer 2, 3 and 5 as well as producer 1 and 4 presented considerable overlapping area. The possibility of authenticating Caciotta origin was tested with the sPLS-DA model for which the cross-validated results are reported in Table 6. The number of components in the training models varied from 9 to 15, suggesting a potential impact of dataset sampling. Nevertheless, this effect did not seem to affect the performances of these models, as they all achieved a 100% correct classification rate.

Table 6. Summary table of elements sPLS-DA models performance based on true positive (TP), true negative (TN), false positive (FP) and false negative (FN) outputs. Recall, precision, specificity, and classification accuracy rate have been calculated according to Kassambara (2018).

Origin/Parameter	Recall	Precision	Specificity	Classification accuracy rate (%)
Producer 1	1.00	1.00	1.00	
Producer 2	0.95	1.00	0.99	
Producer 3	1.00	0.95	1.00	99.00 ± 0.01
Producer 4	1.00	1.00	1.00	
Producer 5	1.00	1.00	1.00	

The cross-validated models (Table 6), on average, demonstrated a slightly lower but still excellent correct classification rate, achieving 99 ± 0.01% accuracy in correctly classifying Caciotta origin. For Producers 1, 4, and 5, the model achieved 100% recall, precision, and specificity, demonstrating its accurate classification of all samples from these areas. However, Producers 2 and 5 had recall, precision, and specificity values ranging from 0.95 to 0.99, which led to a slight decrease in the model's performance. We further investigate which element characterized the performance of the obtained models through maximum importance score (Figure 9).

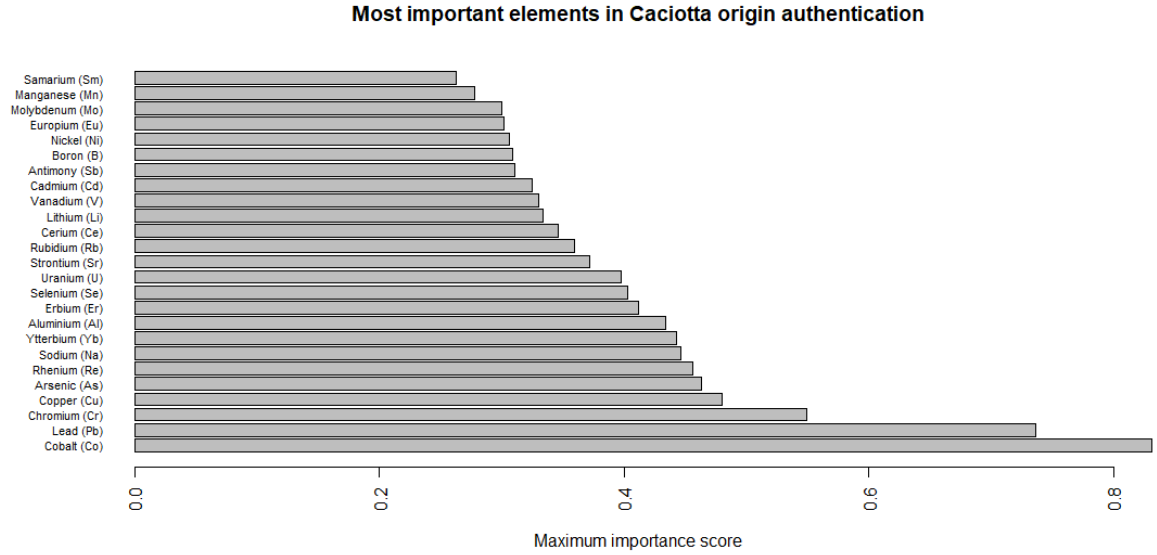


Figure 9. Mean maximum importance score of the most important 25 elements used in the sPLS-DA models.

Among the most important elements for Caciotta origin authentication there were "other elements"such as cobalt, lead and arsenic. Elements associated with dairy equipment like chromium, copper and aluminium. Macronutrients like sodium and rare earth such as rhenium, ytterbium and cerium.

3.5 Stable isotopes

The median values with standard error of $\delta^{13}C$, $\delta^{15}N$, $\delta^{34}S$, $\delta^{18}O$ and $\delta^{2}D$ determined in the mountain Caciotta cheese are shown in Table 7.

Table 7. Pairwise Wilcoxon Rank Sum of isotopes concentration in Caciotta cheese

Origin	^{13}C	^{15}N	^{34}S	^{18}O	^{2}D
Producer 1	-25.70±0.11^{a}	4.15±0.51^{ab}	4.98±1.75^{a}	7.50±1.64^{a}	-114.32±4.64^{a}
Producer 2	-23.46±0.15^{b}	5.05±0.16^{ab}	3.32±0.96^{a}	5.82±1.84^{a}	-120.02±4.15^{b}
Producer 3	-24.13±0.14^{b}	4.11±0.07^{ab}	5.30±1.35^{a}	6.71±1.31^{a}	-114.05±3.66^{c}
Producer 4	-22.95±0.05^{b}	4.72±0.26^{a}	4.84±1.42^{a}	8.21±1.10^{a}	-109.02±2.07^{d}
Producer 5	-24.51±0.18^{b}	4.52±0.09^{b}	4.32±1.51^{a}	7.86±0.87^{a}	-108.60±3.80^{e}

Producer 1 showed a notably lower $\delta^{13}C$ value compared to the rest of the producers. On the other hand, Producer 4 exhibited the higher $\delta^{15}N$ value when compared to Producer 5. Overall, the $\delta^{34}S$ values ranged from 3.32‰ to 5.30‰, but no significant differences were observed. Similarly, $\delta^{18}O$ values, which ranged from 5.82‰ to 8.21‰, did not show significant variations based on cheese origin. Lastly, we noticed significant differences on the $\delta^{2}D$ values for all the origins considered. The unsupervised analysis of stable isotope ratio is shown in Figure 10.

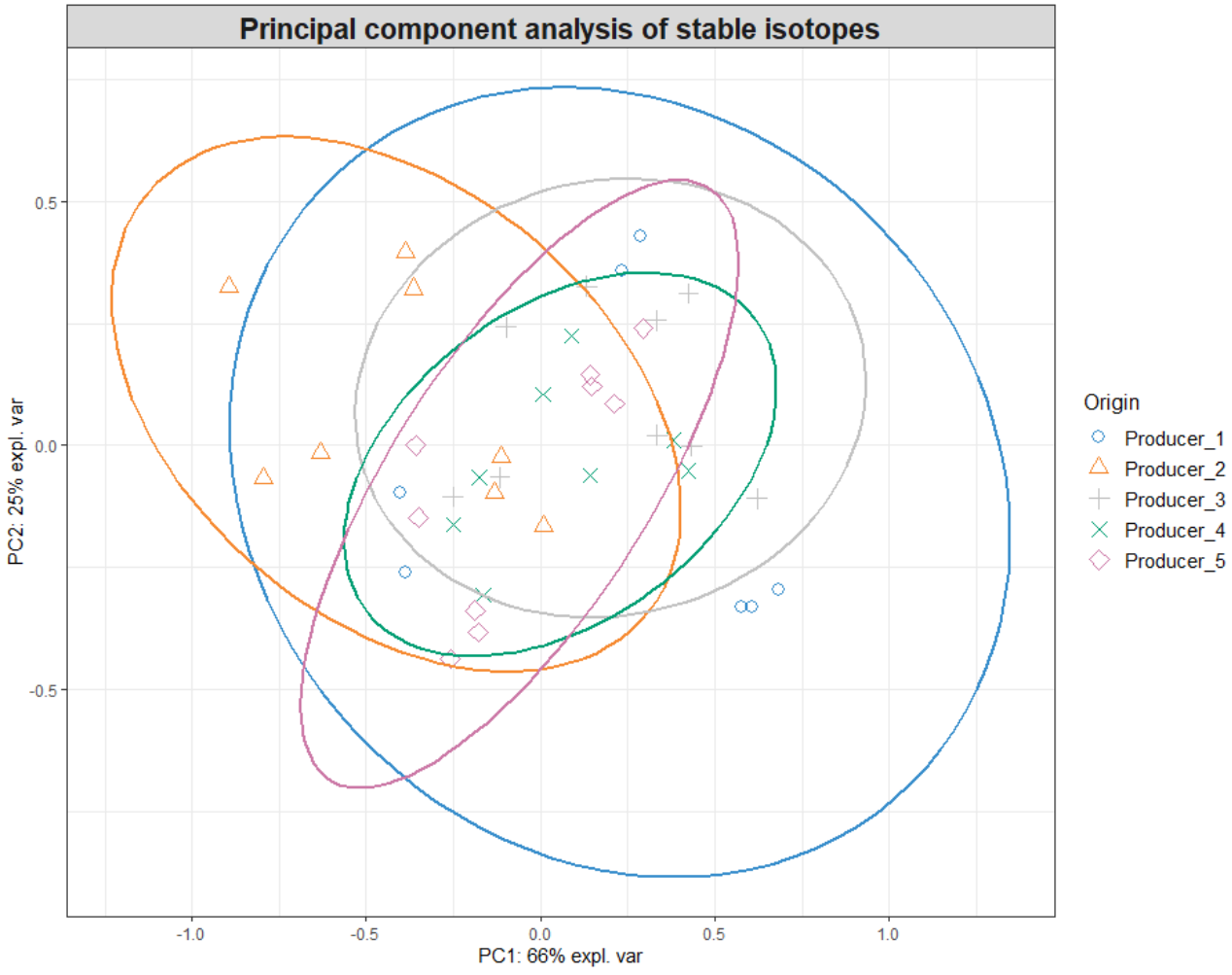


Figure 10. Principal component analysis of Caciotta stable isotopes composition.

The first two components of the principal component analysis accounted for 91% of the variability, but no distinct clusters of Caciotta cheese related to its origin were discernible. We proceeded to assess the potential for authenticating Caciotta origin using sPLS-DA. The

training models showed a slight variation in components, ranging from 3 to 4. However, upon comparison with the previously reported analysis, it became evident that the training model achieved the lowest classification performance, resulting in an average classification accuracy rate of 86.94 ± 5.34%. Table 8 reports the classification performances of the cross-validated models.

Table 8. Summary table of stable isotopes sPLS-DA models performance based on true positive (TP), true negative (TN), false positive (FP) and false negative (FN) outputs. Recall, precision, specificity and classification accuracy rate have been calculated according to Kassambara (2018).

Origin/Parameter	Recall	Precision	Specificity	Classification accuracy rate (%)
Producer 1	0.5	0.33	0.85	
Producer 2	0.90	0.62	0.96	
Producer 3	0.00	0.00	0.76	65.00 ± 53.84
Producer 4	1.00	0.91	1.00	
Producer 5	0.85	1.00	0.94	

We observed a dichotomy in the performance of sPLS-DA for cheese origin authentication. Cheeses from producer 4 and 5 demonstrated high recall, precision, and specificity values, indicating accurate classifications. However, as we moved to producers 2, 1, and 3, the classification performances gradually decreased. In particular, producer 3 exhibited a recall and precision of zero, indicating that the models were unable to correctly identify or predict any cheeses from this producer. However, the specificity was 0.76, suggesting that the model performed comparatively well in correctly identifying cheeses that did not originate from Producer 3. These factors influenced the classification accuracy rate, resulting in the lowest value and the highest accuracy error when compared to the other models. Figure 11 shows the mean maximum importance score of the stable isotopes used in the sPLS-DA model.

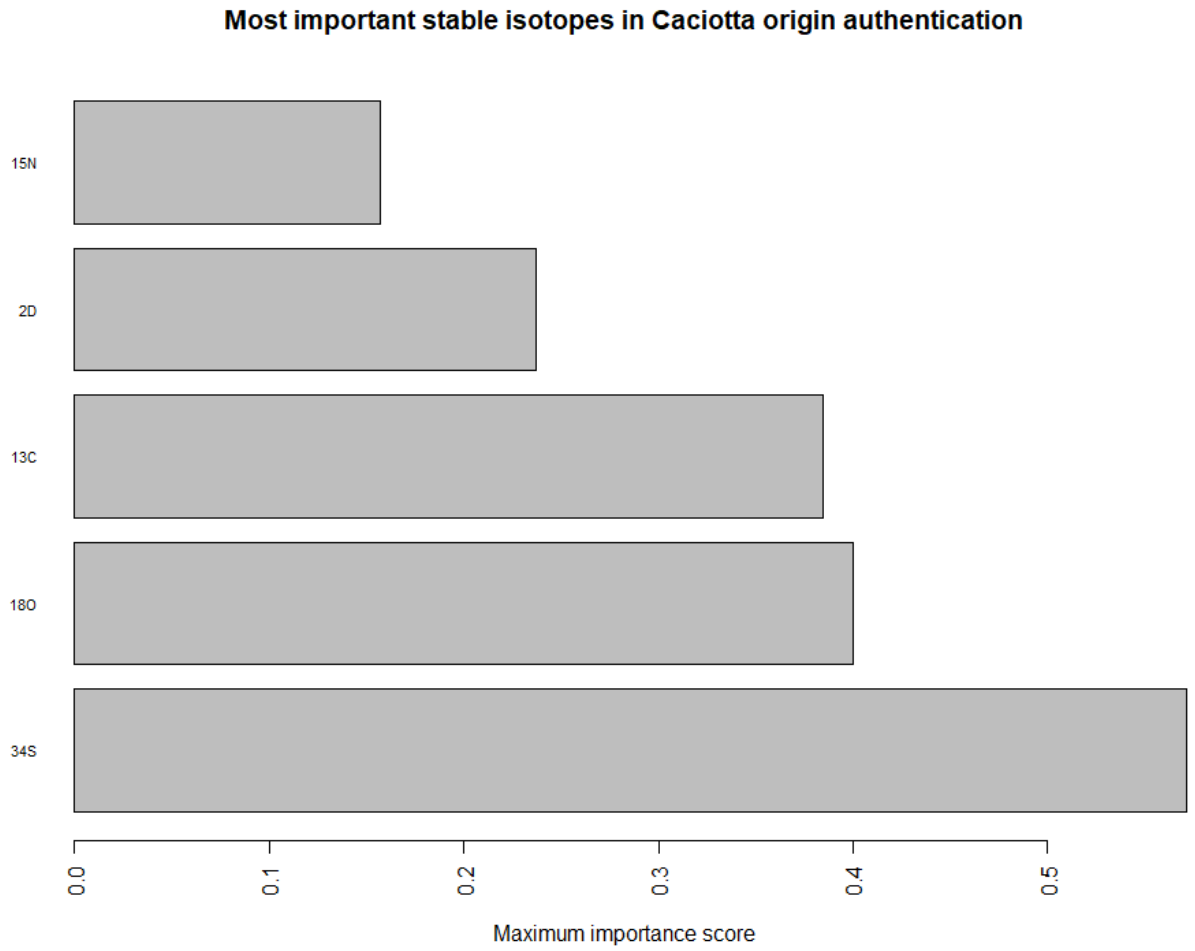


Figure 11. Mean maximum importance score of the stable isotopes used in the sPLS-DA models.

3.6 Combination of element analysis and stable isotopes

Scientific literature reports that the combination of elements and stable isotopes can increase the origin authentication performance of cheese. Consequently, we merged the two datasets and proceeded to assess the performance of the models.

The unsupervised analysis, conducted through principal component analysis, revealed significant differences when compared to individual analyses (Figure 12).

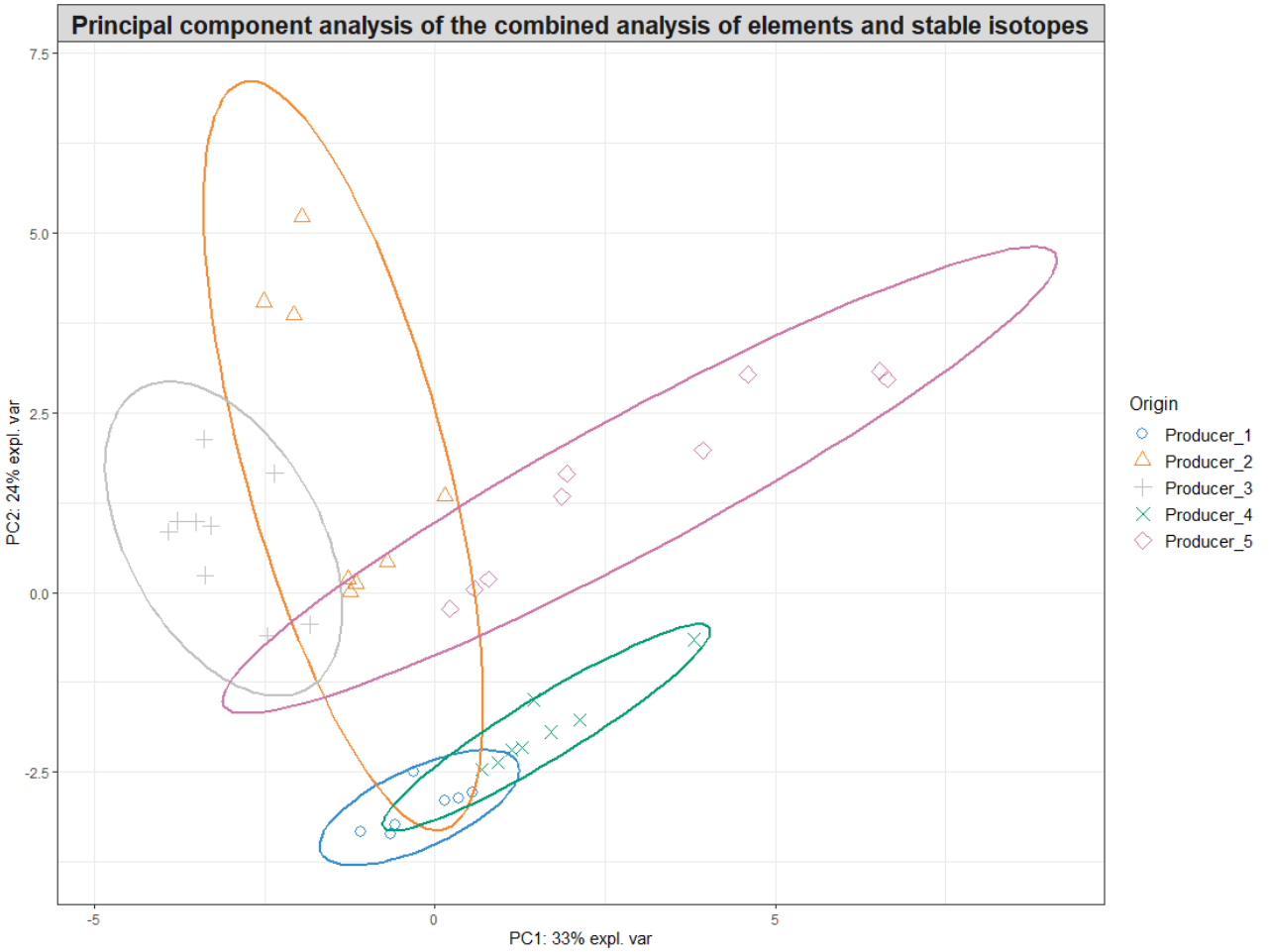


Figure 12. Principal component analysis of Caciotta stable isotopes and elements composition.

The clusters observed in the elemental analysis were more separated. This effect was more appreciable for producers 2, 3 and 5 than producers 1 and 4. Nevertheless, the explained variability of the analysis decreased from 91% for stable isotopes and 90% for elements, respectively, to 57% when both analyses were combined.

In the supervised analysis, the training models showed a decrease in the number of optimal components, ranging from 5 to 8, while achieving the same performance of elements analysis (100%). However, we did notice variations in the cross-validated models (Table 9).

Table 9. Summary table of combined elements and stable isotopes models performance based on true positive (TP). true negative (TN). false positive (FP) and false negative (FN) outputs. Recall. precision. specificity and classification accuracy rate have been calculated according to Kassambara (2018).

Origin/Parameter	Recall	Precision	Specificity	Classification accuracy rate (%)
Producer 1	1.00	1.00	1.00	
Producer 2	1.00	0.90	1.00	
Producer 3	0.95	1.00	0.98	97.00 ± 3.09
Producer 4	1.00	1.00	1.00	
Producer 5	0.9	0.94	0.97	

The origin authentication performance of producer 1 and 4 matched that of the elements models. As for producer 2, there was an improvement in recall and specificity, but the precision of the models decreased from 1 to 0.9. A similar trend was noticed for producer 3. The most significant decrease in performance was observed for producer 5. Overall, there was a slight decline in the classification accuracy rate, which was 97 ± 3.09%. The most important variables for the prediction (Figure 13) were similar to the elements models. We noticed a decrease in the importance of some elements such as lead and cobalt while other elements (*e.g.* copper and selenium) reported higher scores. There were some differences in the importances of $\delta^{15}N$, $\delta^{18}O$ and $\delta^{34}S$ as well. $\delta^{15}N$ achieved a higher score than $\delta^{18}O$, while $\delta^{18}O$ obtained a higher score than $\delta^{34}S$. However, the reverse order was observed for the stable isotopes models.

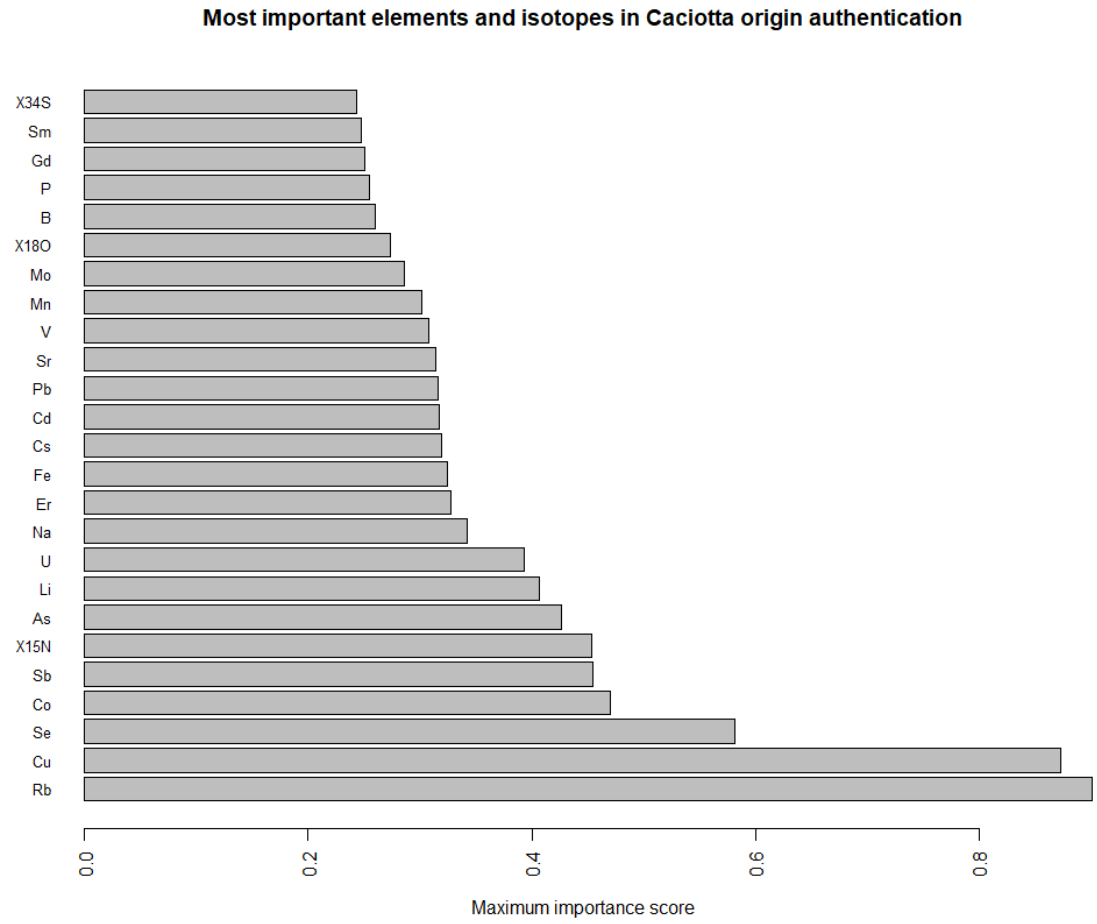


Figure 13. Mean maximum importance score of the stable isotopes and elements sPLS-DA models.

4. Discussion

Cheese origin (*i.e.* the combination of geographical origin and local environmental variables) has been reported to significantly affect cheese microbiota (Sun and D'Amico, 2021; Kamilari et al., 2022; Reuben et al., 2023). In our previous study, we assessed the effect of different biotic and abiotic factors that shaped the microbiome of typical Caciotta cheese, finding that origin was the major contributor to the observed differences. A comparable effect was observed concerning the VOCs, highlighting the importance of microbiota's metabolic activity in the development of Caciotta's typicity. In this study, we further evaluate the microbiome's capability to authenticate cheese origin in a multi-omics environment. To achieve this, we compared it with both established reference analyses and emerging methods. Additionally, we used a two-year sampling period to assess the temporal stability of each method (Riedl et al., 2015).

Among DNA-based methods, the classification accuracy rate of the viral communities outperformed the bacterial communities (97.42 ± 2.58 vs 96.13 ± 4.02%). For both communities, we noticed origin specific non-starter lactic acid bacteria, like *Lactobacillus helveticus, Lactococcus raffinolactis* and *Propionibacterium freudenreichii,* or phage, like Salmonella virus PsP3 and Enterobacteria phage ES18. Corroborating these findings, Dugat-Bony et al. (2016) reported the presence of specific operational taxonomic units in one or several of the 12 analyzed cheese varieties, forming distinct patterns with the cheese production facility. Similarly, the taxa presented in our study showed the highest score in the prediction, suggesting their discrimination ability for origin authentication.

To our knowledge, only another study has evaluated the performances of cheese microbiota in authenticating cheese origin (Kamilari et al., 2022). Our results showed lower prediction performances than those reported using Random Forest algorithm. In particular, the area under the curve (*i.e.* the ratio between specificity and sensitivity) was found to be equal to 1, meaning that all the tested samples were correctly classified according to their origin. The differences between the two studies could be connected to potential overfitting of the

developed model and/or excessively small training dataset. More studies are needed to really assess the performances of DNA-based authentication methods.

Volatile organic compounds showed potential clusters according to cheese origin. Classification accuracy rate of the predicted origin was 90.0 ± 11.11% with secondary alcohols, ketones, esters, terpenes, free fatty acids and hydrocarbons as most important volatile classes. Origin authentication using volatilome data reported different correct classification performances. A study on Pecorino Romano showed that 87.5% of samples were correctly classified after cross-validation while another study on Graviera cheese reported 47.5% of correctly classified samples (Di Donato et al., 2021; Vatavali et al., 2020). Our study was able to correctly classify 90.0 ± 11.11% of mountain Caciotta cheese suggesting that the discriminative power of VOC profile for cheese authentication could vary considerably according to the considered cheese type (Cardin et al., 2022).

In many food fields, spectroscopic techniques have gained popularity since these are fast, solvent-free, automatic, non-destructive, non-invasive, inexpensive, and can be used as a multiparameter analysis. In this context, near infrared analysis is emerging in the food authentication field (Currò et al., 2022; Sammarco et al., 2023; Silva et al., 2022). However, considering cheese, no comparison of the performances of NIR spectrometry and DNA-based methods or reference analysis exist. In our study, we use NIR spectroscopy to authenticate cheese origin obtaining a 76.0 ± 31.57% classification accuracy rate highlighting a considerable error in prediction's accuracy. The performance of these models were primarily influenced by three areas of the IR spectrum, where we observed clustering of important wavelengths for the prediction. In region I and III the importance of the selected wavelength might be associated with different fat content and/or different lipidic profile. Indeed, Silva et al. (2022) reported that vibrational and rotational motions of C-H/C=O groups could be associated with different wavelengths ranging from 1200 to 1214 and 2234 to 2348 nm. On the other hand, region II could be associated with a different protein profile/proteolytic activity, since the range from 1620 to 1700 nm is associated with the N-H bond (Alinovi et al., 2019). Considering the models performances, our results report lower

accuracy compared to the authentication of Emmental PDO (85.7%) or the discrimination of cheeses from different dairy systems (67.1%) (Karoui et al., 2005; Bergamaschi et al., 2020) possibly connected to the limited size of the investigated samples (Niemöller and Holroyd, 2019).

Elements in mountain Caciotta cheese showed characteristic signature for each producer. Significant differences were observed for elements connected with dairy equipment, macronutrient as well as rare earth and other elements. The elemental pathway from farm to the final food product is regulated by the bioavailability of elements, which is influenced by the elemental composition of soil, water, feed, and, subsequently, by animals and transformation process, such as cheesemaking for cheese (Danezis and Georgiou, 2022). In the case of mountain Caciotta cheese, this complex interplay could have been influenced by differences in soil composition between the Veneto and Trentino regions, the proximity of the pastures to roads, the traditional tools employed during cheesemaking, the variations in acidification curves and other factors. For example, the Pedemontana area of the Veneto region lies on limestone rocks and is characterized by normal salinity and alkaline/subalkaline acidity (ARPAV, 2020). These factors are known to affect element availability for plants (Tyler and Olsson, 2001) and could therefore affect the composition of milk and cheese. Among the investigated elements, we noticed concentrations of some toxic metalloids above the limit of quantification (LOQ), which further prompted their selection for their importance in prediction. The maximum tolerable daily intake of inorganic arsenic and antimony are 0.002 and 0.006 mg/kg body weight, respectively (WHO, 2011; WHO, 2022). Caciotta cheese has an average humidity of 37.57%. Considering that the highest concentrations observed for arsenic and antimony were 1.2 and 0.28 ug/kg of dry matter, respectively, we can exclude the potential harmful effects connected to the consumption of these cheeses.

Elements cross-validated models demonstrated an excellent correct classification rate, achieving 99.00 ± 0.01% accuracy in correctly classifying Caciotta origin. Among the most important elements for the prediction we found a majority of “other elements” such as

cobalt, lead and arsenic, elements connected with dairy equipment (*e.g.* copper and aluminium), and rare earths such as cerium and europium which were found to significantly differ among origins. The high correct classification rate and low accuracy error are similar to the studies of Danezis et al. (2019 and 2020) who achieved 95.9% and 92.1% classification accuracy rate. Nevertheless, while rare earth had critical importance in these authors' models, their relevance in Caciotta authentication was lower than other elements.

We investigated mountain Caciotta stable isotope ratios as a reference method for origin authentication. For both $\delta^{13}C$ and $\delta^{15}N$, we found significant differences among some producers, while $\delta^{2}H$ showed significant differences for all the considered origins. We further evaluate the ability of stable isotopes to authenticate Caciotta origin finding that cross-validated sPLS-DA models yielded poor performances in classification accuracy rate (65.00 ± 53.84%), characterized by a high error in accuracy. The ratio of stable isotopes is connected to isotopic fractionation phenomena, such as photosynthetic CO_2 fixation pathways (Boutton et al., 1988) and the water cycle (O'Sullivan et al., 2022). Among the stable isotopes, $\delta^{13}C$ and $\delta^{15}N$ are most affected by animal feed, while $\delta^{2}H$ is heavily influenced by the animal's drinking water. Additionally, the combination of $\delta^{2}H$ with $\delta^{18}O$ is mainly impacted by geographical origin and seasonality (Cardin et al., 2022). On the other hand, $\delta^{34}S$ is primarily linked to geographical origin, specifically soil geology (Pianezze et al., 2020). However, we did not notice any significant difference between the $\delta^{18}O$ and $\delta^{34}S$ stable isotopes. Moreover, we noticed a considerable difference in the performance of our models compared to those described in the scientific literature. For instance, Bontempo et al. (2012) used a similar combination of isotopes and achieved a classification accuracy rate of 80% for the two origins of the investigated cheese types. Lastly, upon combining elements and isotopes, we observed a decrease in the authentication performances (97.00 ± 3.09%) compared to the original element models. These differences could be explained by the close distance between the investigated origins of mountain Caciotta cheese (range of 51 ± 26 km) that could affect isotope fractionation.

Conclusions

DNA-based methods have been suggested as potential tools to authenticate cheese origin. However, no comparative work has assessed their performances in origin authentication. In this study, the discriminative power of shotgun metagenomics, volatilomics, near-infrared spectroscopy, elements metabolomics, and stable isotope ratio were evaluated using sPLS-DA models and a mountain cheese case study obtained from 5 producers in a limited geographic area. The different analytical methods highlighted significant differences connected to origin and/or origin-specific features that could potentially discriminate against Caciotta origin.

The overall classification accuracy rate varied from stable isotope ratio (65.00 ± 53.84%) to element analysis (99.00 ± 0.01%). DNA-based methods obtained appreciable classification accuracy, with the viral communities reaching the second most accurate method (97.42 ± 2.58%), followed by the combination of stable isotopes and element analysis (97.00 ± 3.09%), and the bacterial communities (96.13 ± 4.02%). Average and below-average performances were obtained for near-infrared spectroscopy and stable isotope ratio analysis, possibly connected to the limited number of analyzed samples and the close proximity of cheese origin.

References

Abbas, O., Zadravec, M., Baeten, V., Mikuš, T., Lešić, T., Vulić, A., ... & Pleadin, J. (2018). Analytical methods used for the authentication of food of animal origin. Food Chemistry, 246, 6-17.

Alinovi, M., Mucchetti, G., & Tidona, F. (2019). Application of NIR spectroscopy and image analysis for the characterisation of grated Parmigiano-Reggiano cheese. International Dairy Journal, 92, 50-58.

ARPAV 2020. LEGENDA della CARTA DEI SUOLI IN SCALA 1:50.000. Available at https://gaia.arpa.veneto.it/documents/772 (accessed July 26, 2023).

Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., ... & Segata, N. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife, 10, e65088.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.

Bergamaschi, M., Cipolat-Gotet, C., Cecchinato, A., Schiavon, S., & Bittante, G. (2020). Chemometric authentication of farming systems of origin of food (milk and ripened cheese) using infrared spectra, fatty acid profiles, flavor fingerprints, and sensory descriptions. Food chemistry, 305, 125480.

Bisutti, V., Merlanti, R., Serva, L., Lucatello, L., Mirisola, M., Balzan, S., ... & Capolongo, F. (2019). Multivariate and machine learning approaches for honey botanical origin authentication using near infrared spectroscopy. Journal of Near Infrared Spectroscopy, 27(1), 65-74.

Bontempo, L. U. A. N. A., Lombardi, G., Paoletti, R., Ziller, L., & Camin, F. (2012). H, C, N and O stable isotope characteristics of alpine forage, milk and cheese. International dairy journal, 23(2), 99-104.

Bontempo, L. U. A. N. A., Larcher, R. O. B. E. R. T. O., Camin, F. E. D. E. R. I. C. A., Hölzl, S., Rossmann, A., Horn, P., & Nicolini, G. (2011). Elemental and isotopic characterisation of typical Italian alpine cheeses. International dairy journal, 21(6), 441-446.

Boutton, T. W., Tyrrell, H. F., Patterson, B. W., Varga, G. A., & Klein, P. D. (1988). Carbon kinetics of milk formation in Holstein cows in late lactation. Journal of Animal Science, 66(10), 2636-2645.

Brand, W. A., Coplen, T. B., Vogl, J., Rosner, M., & Prohaska, T. (2014). Assessment of international reference materials for isotope-ratio analysis (IUPAC Technical Report). Pure and Applied Chemistry, 86(3), 425-467.

Brown, J., Pirrung, M., & McCue, L. A. (2017). FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics, 33(19), 3137-3139.

Camin, F., Bertoldi, D., Santato, A., Bontempo, L., Perini, M., Ziller, L., ... & Larcher, R. (2015). Validation of methods for H, C, N and S stable isotopes and elemental analysis of cheese: results of an international collaborative study. Rapid Communications in Mass Spectrometry, 29(5), 415-423.

Camin, F., Wehrens, R., Bertoldi, D., Bontempo, L., Ziller, L., Perini, M., ... & Larcher, R. (2012). H, C, N and S stable isotopes and mineral profiles to objectively guarantee the authenticity of grated hard cheeses. Analytica Chimica Acta, 711, 54-59.

Camin, F., Perini, M., Colombari, G., Bontempo, L., & Versini, G. (2008). Influence of dietary composition on the carbon, nitrogen, oxygen and hydrogen stable isotope ratios of milk. Rapid Communications in Mass Spectrometry: An International Journal Devoted to the

Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry, 22(11), 1690-1696.

Camin, F., Wietzerbin, K., Cortes, A. B., Haberhauer, G., Lees, M., & Versini, G. (2004). Application of multielement stable isotope ratio analysis to the characterization of French, Italian, and Spanish cheeses. Journal of agricultural and food chemistry, 52(21), 6592-6601.

Cardin, M., Cardazzo, B., Mounier, J., Novelli, E., Coton, M., & Coton, E. (2022). Authenticity and Typicity of Traditional Cheeses: A Review on Geographical Origin Authentication Methods. Foods, 11(21), 3379.

Chesson, L. A., Valenzuela, L. O., O'Grady, S. P., Cerling, T. E., & Ehleringer, J. R. (2010). Hydrogen and oxygen stable isotope ratios of milk in the United States. Journal of Agricultural and Food Chemistry, 58(4), 2358-2363.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. Modern methods for business research, 295(2), 295-336.

Chung, D., & Keles, S. (2010). Sparse partial least squares classification for high dimensional data. Statistical applications in genetics and molecular biology, 9(1).

Currò, S., Fasolato, L., Serva, L., Boffo, L., Ferlito, J. C., Novelli, E., & Balzan, S. (2022). Use of a portable near-infrared tool for rapid on-site inspection of freezing and hydrogen peroxide treatment of cuttlefish (Sepia officinalis). Food Control, 132, 108524.

Danezis, G. P., & Georgiou, C. A. (2022). Elemental metabolomics: Food elemental assessment could reveal geographical origin. Current Opinion in Food Science, 44, 100812.

Danezis, G. P., Pappas, A. C., Tsiplakou, E., Pappa, E. C., Zacharioudaki, M., Tsagkaris, A. S., ... & Georgiou, C. A. (2020). Authentication of Greek Protected Designation of Origin cheeses through elemental metabolomics. International Dairy Journal, 104, 104599.

Di Donato, F., Biancolillo, A., Mazzulli, D., Rossi, L., & D'Archivio, A. A. (2021). HS-SPME/GC–MS volatile fraction determination and chemometrics for the discrimination of typical Italian Pecorino cheeses. *Microchemical Journal, 165*, 106133.

Dugat-Bony, E., Garnier, L., Denonfoux, J., Ferreira, S., Sarthou, A. S., Bonnarme, P., & Irlinger, F. (2016). Highlighting the microbial diversity of 12 French cheese varieties. International Journal of Food Microbiology, 238, 265-273.

European Commission, E. (2006). Commission Regulation (EC) No 1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs. Off. J. Eur. Union, 364, 5-24.

Farnham, I. M., Singh, A. K., Stetzenbach, K. J., & Johannesson, K. H. (2002). Treatment of nondetects in multivariate analysis of groundwater geochemistry data. Chemometrics and Intelligent Laboratory Systems, 60(1-2), 265-281.

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences*, *103*(3), 626-631.

Frétin, M., Martin, B., Rifa, E., Isabelle, V. M., Pomiès, D., Ferlay, A., ... & Delbès, C. (2018). Bacterial community assembly from cow teat skin to ripened cheeses is influenced by grazing systems. *Scientific reports*, *8*(1), 200.

Goerges, S., Mounier, J., Rea, M. C., Gelsomino, R., Heise, V., Beduhn, R., ... & Scherer, S. (2008). Commercial ripening starter microorganisms inoculated into cheese milk do not successfully establish themselves in the resident microbial ripening consortia of a South German red smear cheese. *Applied and environmental microbiology*, *74*(7), 2210-2217.

Kamilari, E., Tsaltas, D., Stanton, C., & Ross, R. P. (2022). Metataxonomic mapping of the microbial diversity of Irish and eastern Mediterranean cheeses. Foods, 11(16), 2483.

Kamilari, E., Tomazou, M., Antoniades, A., & Tsaltas, D. (2019). High throughput sequencing technologies as a new toolbox for deep analysis, characterization and potentially authentication of protection designation of origin cheeses?. International Journal of Food Science, 2019.

Karoui, R., Dufour, É., Pillonel, L., Schaller, E., Picque, D., Cattenoz, T., & Bosset, J. O. (2005). The potential of combined infrared and fluorescence spectroscopies as a method of

determination of the geographic origin of Emmental cheeses. International Dairy Journal, 15(3), 287-298.

Kassambara, A. (2018). Machine learning essentials: Practical guide in R. Sthda.

Le Cao, K. A., Costello, M. E., Lakis, V. A., Bartolo, F., Chua, X. Y., Brazeilles, R., & Rondeau, P. (2016). MixMC: a multivariate statistical framework to gain insight into microbial communities. PloS one, 11(8), e0160169.

Lê Cao, K. A., Boitard, S., & Besse, P. (2011). Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. BMC bioinformatics, 12(1), 1-17.

Lei, T., & Sun, D. W. (2019). Developments of nondestructive techniques for evaluating quality attributes of cheeses: A review. Trends in Food Science & Technology, 88, 527-542.

Medina, S., Perestrelo, R., Silva, P., Pereira, J. A., & Câmara, J. S. (2019). Current trends and recent advances on food authenticity technologies and chemometric approaches. Trends in Food Science & Technology, 85, 163-176.

Menozzi, D., Yeh, C. H., Cozzi, E., & Arfini, F. (2022). Consumer Preferences for Cheese Products with Quality Labels: The Case of Parmigiano Reggiano and Comté. Animals, 12(10), 1299.

Montel, M. C., Buchin, S., Mallet, A., Delbes-Paus, C., Vuitton, D. A., Desmasures, N., & Berthier, F. (2014). Traditional cheeses: Rich and diverse microbiota with associated benefits. *International journal of food microbiology*, *177*, 136-154.

Muñoz-Redondo J.M., Bertoldi D., Tonon A., Ziller L., Camin F., Moreno-Rojas J.M. (2022). Multi-element and stable isotopes characterization of commercial avocado fruit (Persea americana Mill) with origin authentication purposes. Food Control 137 (2022) 108975.

Niemöller, A., & Holroyd, S. (2019). Applications of near infrared spectrometry for the analysis of milk and milk products. Bulletin of the International Dairy Federation, (497).

O'Sullivan, R., Schmidt, O., & Monahan, F. J. (2022). Stable isotope ratio analysis for the authentication of milk and dairy ingredients: A review. International Dairy Journal, 126, 105268.

Pianezze, S., Bontempo, L., Perini, M., Tonon, A., Ziller, L., Franceschi, P., & Camin, F. (2020). δ34S for tracing the origin of cheese and detecting its authenticity. Journal of Mass Spectrometry, 55(7), e4451.

Pillonel, L., Ampuero, S., Tabacchi, R., & Bosset, J. (2003). Analytical methods for the determination of the geographic origin of Emmental cheese: volatile compounds by GC/MS-FID and electronic nose. European Food Research and Technology, 216, 179-183.

Reuben, R. C., Langer, D., Eisenhauer, N., & Jurburg, S. D. (2023). Universal drivers of cheese microbiomes. Iscience, 26(1).

Riedl, J., Esslinger, S., & Fauhl-Hassek, C. (2015). Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. Analytica Chimica Acta, 885, 17-32.

Rohart, F., Gautier, B., Singh, A., & Lê Cao, K. A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS computational biology, 13(11), e1005752

Sammarco, G., Bardin, D., Quaini, F., Dall'Asta, C., Christmann, J., Weller, P., & Suman, M. (2023). A Geographical Origin assessment of Italian Hazelnuts: Gas Chromatography-Ion mobility spectrometry coupled with Multivariate Statistical Analysis and Data Fusion approach. Food Research International, 113085.

Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., & Huttenhower, C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. Nature methods, 9(8), 811-814.

Silva, L. K., Santos, L. S., & Ferrão, S. P. (2022). Application of infrared spectroscopic techniques to cheese authentication: A review. International Journal of Dairy Technology, 75(3), 490-512.

Sun, L., & D'Amico, D. J. (2021). Composition, succession, and source tracking of microbial communities throughout the traditional production of a farmstead cheese. Msystems, 6(5), e00830-21.

Tyler, G., & Olsson, T. (2001). Plant uptake of major and minor mineral elements as influenced by soil acidity and liming. Plant and Soil, 230, 307-321.

Vacheyrou, M., Normand, A. C., Guyot, P., Cassagne, C., Piarroux, R., & Bouton, Y. (2011). Cultivable microbial communities in raw cow milk and potential transfers from stables of sixteen French farms. *International journal of food microbiology*, *146*(3), 253-262.

Vatavali, K., Kosma, I., Louppis, A., Gatzias, I., Badeka, A. V., & Kontominas, M. G. (2020). Characterisation and differentiation of geographical origin of Graviera cheeses produced in Greece based on physico-chemical, chromatographic and spectroscopic analyses, in combination with chemometrics. *International Dairy Journal, 110*, 104799.

World Health Organization. (2011). Evaluation of certain contaminants in food. World Health Organization technical report series, (959), 1.

Guidelines for drinking-water quality: fourth edition incorporating the first and second

addenda. Geneva: World Health Organization; 2022. Licence: CC BY-NC-SA 3.0 IGO.

Wolfe, B. E., Button, J. E., Santarelli, M., & Dutton, R. J. (2014). Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. *Cell*, *158*(2), 422-433.

Zhang, P., Georgiou, C. A., & Brusic, V. (2018). Elemental metabolomics. Briefings in bioinformatics, 19(3), 524-536.

General discussion

In the context of this PhD work, 167 articles published over the last 27 years were reviewed to present methods, either physical and chemical or DNA-based, that are currently used for cheese geographical origin authentication. The review highlighted their principles, applications, discriminative power, advantages, and limitations. Additionally, we identified knowledge gaps, such as the unknown discriminative power of the reference methods for closely distant cheese producers, the lack of quantitative assessment of DNA methods, and the unexplored possibility of using shotgun metagenomics for cheese origin authentication. Subsequently, we employed shotgun metagenomics and volatilomics to understand the biotic and abiotic factors contributing to the typicity and diversity of mountain Caciotta cheese. We characterized bacterial and viral communities, profiled the *Streptococcus thermophilus* and *Lactobacillus delbrueckii* strains, investigated their functional potential, and found correlations between Caciotta's volatile organic compounds (VOCs) and genes involved in five different metabolic classes. Lastly, we compared DNA-based methods with physico-chemical methods for origin authentication. We utilized mountain Caciotta cheese in a multi-omics approach framework, which included shotgun metagenomics, volatilome analysis, near-infrared spectroscopy, stable isotopes, and elemental analysis, to evaluate the authentication performance of these methods using sPLS-DA machine learning models.

As part of this study, artisanal cheese from mountain areas was investigated as a case study to explore the capabilities of shotgun metagenomics and evaluate its ability to accurately authenticate cheese origin. Artisanal cheeses are gaining popularity worldwide, prompting regulatory bodies to develop guidelines and minimum requirements to safeguard their unique qualities and ensure consumer safety (Paxson, 2010; Barron et al., 2017). Consequently, verification of authenticity and standards of identity are being established in various countries and regions, providing support to producers and enabling regulatory inspections of these artisanal products (Barron et al., 2017; Wilkinson et al., 2017).

Geographical characteristics, such as soil, climate, and geography, play a decisive role in shaping the distinctive features of artisanal cheeses, along with the ingredients and starter cultures used (Barron et al., 2017). This combination of geographical factors is commonly referred to as "*terroir*," and its characterization is considered essential in defining the identity of artisanal foods, including cheeses (Paxson, 2010). In this context, we evaluate a specific landscape in the Veneto and Trentino regions' mountain areas. Italy, with 60% of its territory covered by mountains, stands as the biggest producer of EU mountain products (Euromontana, 2020), utilizing 47.50% of its agricultural area for mountain food production, amounting to €7,195 million (Santini et al., 2013). Mountain cheese plays a significant role in the cultural heritage of mountain areas and substantially contributes to the mountain economy (Mazzocchi et al., 2022). In this framework this research could be important for the preservation of the cultural heritage of mountain Caciotta cheese, and for protecting the economy of rural communities through different authentication systems. Nevertheless, our work has some relevant limitations driven by economic and time constraints, which may stimulate future research in this complex field.

One limit of this research is the lack of culture-based microbiological analysis. In the context of metagenetics and metagenomics studies, there is a growing concern about the deficiencies in DNA-based methods, which involve uncharted sequences that could potentially be associated with known or unknown microorganisms. Recently, the concept of culturomics has gained popularity in microbiota studies (Lagier et al., 2012; Seng et al., 2009; Nowrotek et al., 2019). The aim of culturomics is to identify and characterize a wide range of bacteria present in a sample (Lagier et al., 2012). This is achieved by plating serial dilutions of the sample on agar media with various culture conditions. After growth, each bacterial colony is isolated and identified using Matrix Assisted Laser Desorption Ionisation-Time of Flight (MALDI-TOF) mass spectrometry, providing taxonomic identification (Seng et al., 2009). If identification by MALDI-TOF is not possible, 16S rRNA amplification and sequencing are employed for identification. Culturomics allows to obtain living bacteria or intact viruses that could be investigated for different microbiological features (Martellaci et al., 2019); for example, each isolated colony could be tested for

cheesemaking attitudes such as VOCs production or isolated phages could be used against pathogenic bacteria during cheese production. Both shotgun metagenomics and culturomics complement each other (Lagier et al., 2018; Nowrotek et al., 2019). For example, the shotgun sequencing of microbial DNA enables the functional annotation of genes, offering valuable insights into the microbes' functional capabilities, which could be beneficial for their cultivation process. Ultimately, this approach enables the investigation of microbial competition, interactions, and the strategies they employ to thrive and survive within communities (Ferrocino et al., 2022). Combining these methods to study cheese, in particular traditional cheese, can significantly enhance our understanding of cheese microbiome and explore cheese diversity and typicity to unprecedented levels.

Some considerations about our multi-omics geographical origin authentication study may arise from using only one cheese variety. We exclusively utilized mountain Caciotta cheese to compare different authentication methods due to its similar cheesemaking process, traditional tools employed, animals feeding systems, ripening days, sampling years, and close geographical locations. This approach was intended to challenge the employed methods in authenticating cheese origin. Furthermore, we observed no decrease in authentication performance across the investigated sampling years, suggesting stability in the obtained fingerprints for many of the investigated methods. These encouraging results could be further explored with different traditional cheese varieties (*e.g.*, mountain and PDO cheese) and diverse geographical locations to validate the findings more extensively. In this context, a recent study employing 16S metabarcoding on 824 cheese samples spanning 58 cheese types and 16 countries discovered diversity patterns in the cheese microbiota, influenced by geography and local environmental variables (Rueben et al., 2023). However, the limited discrimination power of metagenetics, combined with the analyses, was hindered by the use of different experimental approaches, the unavailability of deposited (sequence) data, a lack of consistent reporting, and the inability to develop authentication models. Future studies may consider investigating multiple cheeses using shotgun metagenomics to further validate bacterial and viral fingerprints.

In this elaborate NIRS yielded average performances in origin authentication, possibly connected to the limited number of investigated cheese (Niemöller and Holroyd, 2019). Considering the relatively low cost per analysis of this method, further studies should consider increasing the sample size of the investigated cheeses to gain a better assessment of these results. Moreover, we did not investigate emerging miniaturized vibrational spectroscopy devices. These devices integrate nano-engineered sensor components, customizable machine learning analysis, and data transmission and processing capabilities (Rodriguez-Saona et al., 2020). The miniaturization of these devices brings several advantages, including improved performance in terms of speed, portability, ruggedness, reduced power consumption, and compactness. Despite the lack of studies on cheese origin authentication, promising performances have been obtained in the authentication of whey protein, oil, spinach, among other foods (Peng and Wang, 2015; Basri et al., 2017; Sánchez et al., 2018). In addition, innovative tools (*e.g.,* radio frequency identification, near field communication, and QR code technologies in combination with blockchain systems) can be commercially implemented to verify the processes and could aid in controlling the quality of foods (Violino et al., 2020).

In our study, we mainly used PCA and sPLS-DA models to investigate and assess the performances of different methods for authenticating cheese origin. However, in the authentication field, different models could be used to classify samples. The aim of classification models is to develop a function that assigns a sample to a predefined class. Classification algorithms can be categorized as either discriminant analysis or class modeling. Discriminant algorithms function as two or multiclass classifiers, while class modeling algorithms are used for one-class classification models. Common class modeling algorithms include soft independent modeling of class analogy (SIMCA) and unequal dispersed classes (UNEQ), both of which assess whether a sample aligns with the characteristics of a single class of interest (Oliveri and Downey, 2012). On the other hand, discriminant algorithms establish boundaries between analyzed classes based on training samples. Well-known discriminant algorithms include linear discriminant analysis (LDA), partial least squares discriminant analysis (PLS-DA), quadratic discriminant analysis

(QDA), k nearest neighbors (kNN), and support vector machines (SVM). Additionally, there are other multivariate classification methods based on decision rules, such as decision trees (DT), random forests (RF), and artificial neural networks (ANN). In the realm of food analysis, Jiménez-Carvelo et al. (2019) categorized commonly used multivariate methods as conventional ones (PCA, PLS-DA, LDA, KNN, PARAFAC, SIMCA) and alternative data mining/machine learning methods from computer science (SVM, ANN, CART, J48, C5.0, Random Forest). The latter group of methods is sought after as data complexity increases due to advanced analytical techniques providing detailed physical-chemical parameter information.Various decision tree algorithms (CART, J48, C5.0, Random Forest), SVM, and ANNs have emerged from computer science, and some of them may be computationally intensive and lack a reproducible solution in certain cases. Despite being widely used in other artificial intelligence applications, their application in food science remains relatively limited (Jiménez-Carvelo et al., 2019; Chung et al., 2022). Future investigations should consider preliminary test using different algorithms in order to evaluate which model could be the most suitable for the method under investigation.

Lastly, some consideration regarding the costs of the different analyses should be addressed. In our study, the cost per sample varied significantly from NIRS to stable isotopes analysis, while volatilomics, shotgun metagenomics, and element analysis fell in the middle of the cost scale. Nevertheless, the average cost per analysis of some methods is constantly decreasing. The cost of elemental screening is rapidly declining, with instruments getting smaller and sample throughput increasing, while at the same time, both the capital and operating costs of instrumentation and measurement are diminishing (Danezis et al., 2016). Also, DNA-based methods are presenting interesting opportunities. In various research areas, a cost reduction per sample was reported (Elliot et al., 2021; Sanders et al., 2022), which could facilitate future advancements in DNA-based methods for cheese origin authentication. For instance, Sanders et al. (2022) introduced a workflow for bacterial whole-genome sequencing using open-source labware and the OpenTrons robotics platform, effectively reducing costs to approximately $10 per genome.

General conclusion

In the framework of this PhD work, aimed at better understanding the microbiological diversity of traditional cheeses and safeguarding consumers against food fraud, we tried to address four main scientific questions: i) What are the main ecological drivers of typical Caciotta cheese produced in different regions with similar cheesemaking know-how? ii) Is the cheese microbiome stable over time? Is it possible to identify considerable changes in the cheese microbiome according to season or year? iii) Can the cheese microbiome be used as an authentication marker? iv) What are the performances of shotgun metagenomics, volatilome, near-infrared spectroscopy, stable isotope ratio, and trace element analysis in authenticating the origin of typical mountain cheese?

We reviewed the scientific literature and found that among the possible methods to study and authenticate traditional cheese, the emerging shotgun metagenomics could offer multiple advantages compared to traditional culture-based methods and metabarcoding. Subsequently, we employed shotgun metagenomics and volatilomics to investigate three of the main questions described herein. Results showed that the geographical origin of cheese played a significant role; however, curd cooking temperature, pH, salt concentration, and water activity also had an impact. Moreover, the two-year sampling period of this study allowed us to observe a stable microbial community in the investigated Caciotta cheese. Additionally, alongside the bacterial communities, we characterized the viral communities, which are another important and poorly described component of the cheese microbiome.

Lastly, we developed sPLS-DA models to authenticate mountain Caciotta origin using a multi-omics approach. To the best of our knowledge, this is the first study that employed shotgun metagenomics, reference analytical methods such as stable isotope ratio and trace element analysis, spectroscopy analysis, as well as volatilomics to quantitatively evaluate and compare the performances of these methods. Overall, DNA-based analysis revealed a high classification accuracy rate for both viral (97.42 ± 2.58%) and bacterial communities (96.13 ± 4.02%), with the former bested only by elemental analysis models (99.00 ± 0.01%).

Our findings emphasize the potential of multi-omics techniques in cheese origin authentication and highlight the complexity of factors influencing cheese composition. Moreover, these results shed light on the world of DNA-based methods that, if future research validates them, could be employed to authenticate the origin of traditional cheeses.

Bibliography

Afshari, R., Pillidge, C. J., Read, E., Rochfort, S., Dias, D. A., Osborn, A. M., & Gill, H. (2020). New insights into cheddar cheese microbiota-metabolome relationships revealed by integrative analysis of multi-omics data. Scientific reports, 10(1), 3164.

Apolito, A. (2018). Programma Rete Rurale-Nazionale 2014-2020-Rapporto di valutazione delle attività. Available at https://www.reterurale.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/17423 (accessed July 18, 2023)

Aprile, M. C., Caputo, V., & Nayga Jr, R. M. (2012). Consumers' valuation of food quality labels: the case of the European geographic indication and organic farming labels. International Journal of Consumer Studies, 36(2), 158-165.

Arfini, F., Cozzi, E., Mancini, M. C., Ferrer-Perez, H., & Gil, J. M. (2019). Are geographical indication products fostering public goods? Some evidence from Europe. Sustainability, 11(1), 272.

Arfini, F., & Bellassen, V. (Eds.). (2019). Sustainability of European food quality schemes: Multi-performance, structure, and governance of PDO, PGI, and organic agri-food systems. Springer Nature.

Bancalari, E., Montanari, C., Levante, A., Alinovi, M., Neviani, E., Gardini, F., & Gatti, M. (2020). *Lactobacillus paracasei* 4341 as adjunct culture to enhance flavor in short ripened Caciotta-type cheese. Food Research International, 135, 109284.

Barcaccia, G., Lucchin, M., & Cassandro, M. (2015). DNA barcoding as a molecular tool to track down mislabeling and food piracy. Diversity, 8(1), 2.

Barron, L. J. R., Aldai, N., Virto, M., & de Renobales, M. (2017). Cheeses with Protected Land- and Tradition-Related Labels: Traceability and Authentication. Global cheesemaking technology: Cheese quality and characteristics, 100-119.

Basri, K. N., Hussain, M. N., Bakar, J., Sharif, Z., Khir, M. F. A., & Zoolfakar, A. S. (2017). Classification and quantification of palm oil adulteration via portable NIR spectroscopy. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 173, 335-342.

Bokulich, N. A., & Mills, D. A. (2013). Facility-specific "house" microbiome drives microbial landscapes of artisan cheesemaking plants. Applied and environmental microbiology, 79(17), 5214-5223.

Cardin, M., Cardazzo, B., Mounier, J., Novelli, E., Coton, M., & Coton, E. (2022). Authenticity and Typicity of Traditional Cheeses: A Review on Geographical Origin Authentication Methods. Foods, 11(21), 3379.

Calasso, M., Ercolini, D., Mancini, L., Stellato, G., Minervini, F., Di Cagno, R., ... & Gobbetti, M. (2016). Relationships among house, rind and core microbiotas during manufacture of traditional Italian cheeses at the same dairy plant. *Food Microbiology*, *54*, 115-126.

Chung, T., Tam, I. Y. S., Lam, N. Y. Y., Yang, Y., Liu, B., He, B., ... & Lau, L. T. (2022). Non-targeted detection of food adulteration using an ensemble machine-learning model. Scientific Reports, 12(1), 20956.

Danezis, G. P., Tsagkaris, A. S., Camin, F., Brusic, V., & Georgiou, C. A. (2016). Food authentication: Techniques, trends & emerging approaches. TrAC Trends in Analytical Chemistry, 85, 123-132.

Deselnicu, O. C., Costanigro, M., Souza-Monteiro, D. M., & McFadden, D. T. (2013). A meta-analysis of geographical indication food valuation studies: What drives the premium for origin-based labels?. Journal of Agricultural and Resource Economics, 204-219.

Elliott, T. M., Harris, P. N., Roberts, L. W., Doidge, M., Hurst, T., Hajkowicz, K., ... & Gordon, L. G. (2021). Cost-effectiveness analysis of whole-genome sequencing during an outbreak of carbapenem-resistant Acinetobacter baumannii. Antimicrobial Stewardship & Healthcare Epidemiology, 1(1), e62.

Euromontana. (2020). Implementation of the EU Optional Quality Term "Mountain Product". Where Do We Stand in the Different Member States?. Available et https://www.euromontana.org/wp-content/uploads/2020/06/2020-05-26-Implementation-of-the-OQT_EN.pdf (accessed July 31, 2023)

Ferrocino, I., Rantsiou, K., & Cocolin, L. (2022). Investigating dairy microbiome: an opportunity to ensure quality, safety and typicity. Current Opinion in Biotechnology, 73, 164-170.

Garavaglia, C., & Mariani, P. (2017). How much do consumers value protected designation of origin certifications? Estimates of willingness to pay for PDO dry-cured ham in Italy. Agribusiness, 33(3), 403-423.

García-Hernández, C., Ruiz-Fernández, J., & Rodríguez-Gutiérrez, F. (2022). Geographical indications in cheese mountain areas: Opportunity or threat to landscape and environmental conservation? The case of Cabrales (Spain). Applied Geography, 146, 102753.

Gobbetti, M., Di Cagno, R., Calasso, M., Neviani, E., Fox, P. F., & De Angelis, M. (2018). Drivers that establish and assembly the lactic acid bacteria biota in cheeses. Trends in Food Science & Technology, 78, 244-254.

Imarc. (February 5, 2022). Market value of cheese worldwide from 2019 to 2027 (in billion U.S. dollars). In Statista. Retrieved February 03, 2023, from https://www.statista.com/statistics/602542/cheese-market-value-worldwide/

ISMEA (Italy). (February 14, 2022). Production value of certified food companies in Italy in 2020, by category (in million euros) [Graph]. In Statista. Retrieved February 02, 2023, from https://www.statista.com/statistics/719148/production-value-certified-food-companies-by-category-in-italy/

Jiménez-Carvelo, A. M., González-Casado, A., Bagur-González, M. G., & Cuadros-Rodríguez, L. (2019). Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity–A review. Food research international, 122, 25-39.

Kamilari, E., Tomazou, M., Antoniades, A., & Tsaltas, D. (2019). High throughput sequencing technologies as a new toolbox for deep analysis, characterization and potentially authentication of protection designation of origin cheeses?. International journal of food science, 2019.

Karoui, R., Bosset, J. O., Mazerolles, G., Kulmyrzaev, A., & Dufour, É. (2005a). Monitoring the geographic origin of both experimental French Jura hard cheeses and Swiss Gruyère and L'Etivaz PDO cheeses using mid-infrared and fluorescence spectroscopies: a preliminary investigation. International Dairy Journal, 15(3), 275-286.

Karoui, R., Dufour, É., Pillonel, L., Schaller, E., Picque, D., Cattenoz, T., & Bosset, J. O. (2005b). The potential of combined infrared and fluorescence spectroscopies as a method of determination of the geographic origin of Emmental cheeses. International Dairy Journal, 15(3), 287-298.

Komarova, N., Barkova, D., & Kuznetsov, A. (2020). Implementation of high-throughput sequencing (HTS) in aptamer selection technology. International journal of molecular sciences, 21(22), 8774.

Lagier, J. C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., ... & Raoult, D. (2012). Microbial culturomics: paradigm shift in the human gut microbiome study. Clinical Microbiology and Infection, 18(12), 1185-1193.

Lagier, J. C., Dubourg, G., Million, M., Cadoret, F., Bilen, M., Fenollar, F., ... & Raoult, D. (2018). Culturing the human microbiota and culturomics. Nature Reviews Microbiology, 16(9), 540-550.

Lê Cao, K. A., Boitard, S., & Besse, P. (2011). Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. BMC bioinformatics, 12(1), 1-17.

Mancini, M. C., & Donati, M. (Eds.). (2012). Local agri-food systems in a global world: Market, social and environmental challenges. Cambridge Scholars Publishing.

Martellacci, L., Quaranta, G., Patini, R., Isola, G., Gallenzi, P., & Masucci, L. (2019). A literature review of metagenomics and culturomics of the peri-implant microbiome: current evidence and future perspectives. Materials, 12(18), 3010.

Medina, S., Pereira, J. A., Silva, P., Perestrelo, R., & Câmara, J. S. (2019). Food fingerprints–A valuable tool to monitor food authenticity and safety. Food Chemistry, 278, 144-162.

Menapace, L., Colson, G., Grebitus, C., & Facendola, M. (2011). Consumers' preferences for geographical origin labels: evidence from the Canadian olive oil market. European Review of Agricultural Economics, 38(2), 193-212.

Menozzi, D., Yeh, C. H., Cozzi, E., & Arfini, F. (2022). Consumer Preferences for Cheese Products with Quality Labels: The Case of Parmigiano Reggiano and Comté. Animals, 12(10), 1299.

Montel, M. C., Buchin, S., Mallet, A., Delbes-Paus, C., Vuitton, D. A., Desmasures, N., & Berthier, F. (2014). Traditional cheeses: Rich and diverse microbiota with associated benefits. International journal of food microbiology, 177, 136-154.

Niemöller, A., & Holroyd, S. (2019). Applications of near infrared spectrometry for the analysis of milk and milk products. Bulletin of the International Dairy Federation, (497).

Nowrotek, M., Jałowiecki, Ł., Harnisz, M., & Płaza, G. A. (2019). Culturomics and metagenomics: In understanding of environmental resistome. Frontiers of Environmental Science & Engineering, 13, 1-12.

Oliveri, P., & Downey, G. (2012). Multivariate class modeling for the verification of food-authenticity claims. TrAC Trends in Analytical Chemistry, 35, 74-86.

Parente, E., Ricciardi, A., & Zotta, T. (2020). The microbiota of dairy milk: A review. International Dairy Journal, 107, 104714.

Paxson, H. (2010). Locating value in artisan cheese: reverse engineering terroir for new-world landscapes. American Anthropologist, 112(3), 444-457.

Peng, Y., & Wang, W. (2015). Application of near-infrared spectroscopy for assessing meat quality and safety. Infrared Spectroscopy—Anharmonicity of Biomolecules, Crosslinking of Biopolymers, Food Quality and Medical Applications.

Pilone, V., De Lucia, C., Del Nobile, M. A., & Contò, F. (2015). Policy developments of consumer's acceptance of traditional products innovation: The case of environmental sustainability and shelf life extension of a PGI Italian cheese. Trends in Food Science & Technology, 41(1), 83-94.

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. Nature biotechnology, 35(9), 833-844.

Ranjan, R., Rani, A., Metwally, A., McGee, H. S., & Perkins, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. Biochemical and biophysical research communications, 469(4), 967-977.

Reuben, R. C., Langer, D., Eisenhauer, N., & Jurburg, S. D. (2023). Universal drivers of cheese microbiomes. Iscience, 26(1), 105744.

Rodriguez-Saona, L., Aykas, D. P., Borba, K. R., & Urtubia, A. (2020). Miniaturization of optical sensors and their potential for high-throughput screening of foods. Current Opinion in Food Science, 31, 136-150.

RoślinnieJemy. (May 25, 2021). What decision do you make when given the choice between buying traditional cheese or its plant-based alternatives (e.g. made from cashew nuts)? [Graph]. In Statista. Retrieved February 02, 2023, from https://www.statista.com/statistics/1248146/poland-non-dairy-cheese-consumption-preferences/

Sánchez, M. T., Entrenas, J. A., Torres, I., Vega, M., & Pérez-Marín, D. (2018). Monitoring texture and other quality parameters in spinach plants using NIR spectroscopy. Computers and electronics in agriculture, 155, 446-452.

Sanders, J. G., Yan, W., Mjungu, D., Lonsdorf, E. V., Hart, J. A., Sanz, C. M., ... & Moeller, A. H. (2022). A low-cost genomics workflow enables isolate screening and strain-level analyses within microbiomes. Genome Biology, 23(1), 212.

Santeramo, F. G., & Lamonaca, E. (2020). Evaluation of geographical label in consumers' decision-making process: A systematic review and meta-analysis. Food Research International, 131, 108995.

Sanjuan, A. I., & Khliji, S. (2016). Urban consumers' response to the EU food mountain labelling: an empirical application in Southern Europe. New Medit, 15(1), 72-81.

Santini, F., Guri, F., & Gomez y Paloma, S. (2013). Labelling of agricultural and food products of mountain farming. Proyecto encargado por la Dirección General de Agricultura.

Sattin, E., Andreani, N. A., Carraro, L., Lucchini, R., Fasolato, L., Telatin, A., ... & Cardazzo, B. (2016). A multi-omics approach to evaluate the quality of milk whey used in ricotta cheese production. Frontiers in microbiology, 7, 1272.

Seng, P., Drancourt, M., Gouriet, F., La Scola, B., Fournier, P. E., Rolain, J. M., & Raoult, D. (2009). Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clinical infectious diseases, 49(4), 543-551.

Silva, L. K., Santos, L. S., & Ferrão, S. P. (2022). Application of infrared spectroscopic techniques to cheese authentication: A review. International Journal of Dairy Technology, 75(3), 490-512.

Somerville, V., Berthoud, H., Schmidt, R. S., Bachmann, H. P., Meng, Y. H., Fuchsmann, P., ... & Engel, P. (2022). Functional strain redundancy and persistent phage infection in Swiss hard cheese starter cultures. The ISME Journal, 16(2), 388-399.

USDA, US Department of Agriculture, & Economic Research Service. (December 22, 2022). Major cheese producing countries in 2022 (in 1,000 metric tons) [Graph]. In Statista. Retrieved February 06, 2023, from https://www.statista.com/statistics/195809/cheese-production-in-selected-countries-2009/

Violino, S., Ortenzi, L., Antonucci, F., Pallottino, F., Benincasa, C., Figorilli, S., & Costa, C. (2020). An artificial intelligence approach for Italian EVOO origin traceability through an open source IoT spectrometer. Foods, 9(6), 834

Wilkinson, J., Cerdan, C., & Dorigon, C. (2017). Geographical indications and "origin" products in Brazil–the interplay of institutions and networks. World Development, 98, 82-92.

Yang, C., You, L., Kwok, L. Y., Jin, H., Peng, J., Zhao, Z., & Sun, Z. (2021). Strain-level multiomics analysis reveals significant variation in cheeses from different regions. LWT, 151, 112043.

Communications

Publications

1. CARDIN M, CARDAZZO B, MOUNIER J, NOVELLI E, COTON M, COTON E (2022) Authenticity and Typicity of Traditional Cheeses: A Review on Geographical Origin Authentication Methods. *Foods*, **11**(21):3379. 10.3390/foods11213379

2. CARDIN M, CARDAZZO B, COTON M, CARRARO L, LUCCHINI R, NOVELLI E, COTON E., MOUNIER J. Ecological diversity and associated volatilome of typical mountain Caciotta cheese from Italy (in preparation)

3. CARDIN M, CARDAZZO B, MOUNIER J, CARRARO L, LUCCHINI R, NOVELLI E, COTON M, COTON E. Application of omics-based technologies for authenticating the origin of typical Italian mountain cheese with sPLS-DA modeling (in preparation)

Posters

1. CARDIN M, CARDAZZO B, MOUNIER J, CARRARO L, LUCCHINI R, NOVELLI E, COTON M, COTON E. Discriminative power of shotgun metagenomic and volatilome analysis for geographical origin authentication of typical Italian mountain cheeses. 10th International Symposium on Recent Advances in Food Analysis (RAFA 2022) September 6–9, 2022, Prague, Czech Republic

Publication out of the direct scope of the PhD

1. BAILONI L, CARRARO L, CARDIN M, CARDAZZO B. Active Rumen Bacterial and Protozoal Communities Revealed by RNA-Based Amplicon Sequencing on Dairy Cows Fed Different Diets at Three Physiological Stages. *Microorganisms* **9**(4):754 10.3390/microorganisms9040754
2. MOHAMMADPOUR H, CARDIN M, CARRARO L, FASOLATO L, CARDAZZO B. (2023). Characterization of the archaeal community in foods: The neglected part of the food microbiota. International Journal of Food Microbiology, 110275. 10.1016/j.ijfoodmicro.2023.110275

Discriminative power of shotgun metagenomic and volatilome analysis for geographical origin authentication of typical Italian mountain cheeses

Marco Cardin[1,2], Barbara Cardazzo[1], Jérôme Mounier[2], Lisa Carraro[1], Rosaria Lucchini[3], Enrico Novelli[1], Monika Coton[2] and Emmanuel Coton[2]

[1]Dipartimento di Biomedicina Comparata e Alimentazione, Università degli Studi di Padova, Viale dell'Università 16,35020 Legnaro, Italy

[2]Univ Brest, INRAE, LUBEM, Technopôle Brest-Iroise, F-29280 Plouzané, France

[3]Istituto Zooprofilattico Sperimentale delle Venezie, Section of Trento, Via Lavisotto, 129, 38121 Trento TN, Italy

E-mail: marco.cardin.4@phd.unipd.it

Overview

•Origin authentication methods are pivotal in cunteracting frauds and provide evidence for certification systems.
•Bacterial and viral communities are heavily affected by origin.
•Cross-validated sPLS-DA models on bacterial and viral communities were able to correctly classify 83.87 and 60.21% of cheese origin.
• Volatilome analysis based on 44 validated molecules was able to correctly classify 67.74% of cheese origin

Introduction

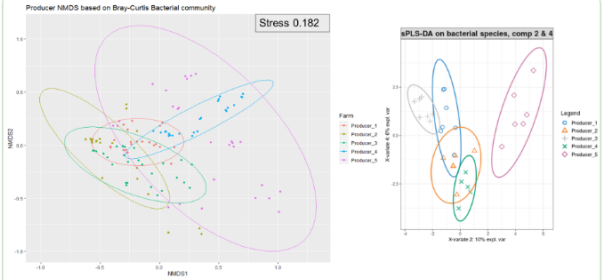
Dairy products are among the most common products concerned by food frauds. In particular, protected land- and tradition-related labeled cheeses (*e.g.* PDO or produced in mountainous regions) are subjected to food fraud (mainly mislabelling and fraudulent documentation) due to their high economic value. Indeed, consumers are more willing to pay higher prices for better quality and typical mountain cheeses due to their distinctive flavors and appearance but also their more natural and animal-friendly production attributes. For these reasons, geographical origin authentication methods are used to ensure the product origin. The main methods rely on chemical analyses such as stable isotope ratios, trace elements and fatty acid profiles. Microbial ecology studies have also highlighted how the combinations of different environmental factors, and cheese-making conditions and traditional know-how select specific microorganisms. Thus, DNA-based methods applied to microbiota definition have also been suggested as potential tools to authenticate cheese geographical origins. The aim of this study was to evaluate the discriminative power of DNA shotgun metagenomics (both bacterial and viral community profiling) but also volatilome for cheese product origin authentication. To do so, a case study approach was applied to typical semi-hard raw milk Italian mountain cheeses (caciotta and caciotta-like).

Results

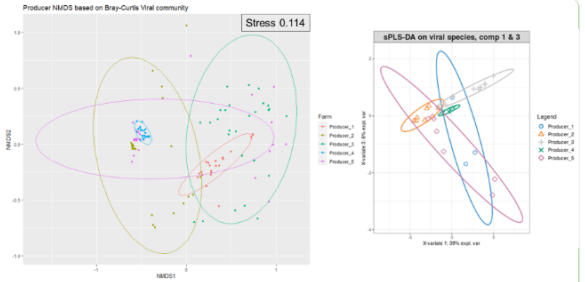
Non-metric multidimensional scaling of bacterial and viral communities at species level showed overlaping clusters (CI=95%) when considering cheese origin. Although dominant starter lactic acid bacteria - *Streptococcus thermophilus and Lactobacillus delbrueckii* - were similar among cheeses from different origins, non-starter species differentiated most cheese origins(*e.g. Loigolactobacillus bifermentans, Moraxella osloensis, Lactobacillus kefiranofaciens).* Stronger **origin effect** was observed for **viral communities** in comparison with bacterial communities for viral communities (PERMANOVA R^2=0.548 Vs R^2=0.452, *p=0.001*). However, **classification accuracy rate** of **bacterial communities** on sparse partial least squares discriminant analysis outperformed viral communities (83.87% Vs 60.21%). Variable importance in the projection (VIP) showed that phages of *Streptococcus, Lactobacillus and Lactoccus* such as Lactococcus phage BK5T, Streptococcus phage TP 778L, Lactococcus phage P335 sensu lato, Lactobacillus prophage Lj771, Streptococcus virus 7201 and Streptococcus virus DT1 presented the highest score. Bacterial communities VIP considered **secondary microbiota** *Lactococcus petauri, Lactococcus piscium, Enterococcus italicus, Companilactobacillus crustorum, Loigolactobacillus bifermentans, Moraxella osloensis, Lactobacillus kefiranofaciens, Raoultella terrigena, Chryseobacterium jejuense…* more important. **Volatile organic compounds** showed potential clusters according to cheese origin. Classification accuracy rate of the predicted origin was 67.74% with secondary alcohols, free fatty acids, esters and idrocarbon as most important class of volatiles. High score VIP were butan-2-one, 2 –phenylethanol, hexanoic acid, cyclopentane, butan-1-ol, 3-methyl-butan-1-ol, toluene, ethyl butanoate, heptan -2-ol, benzaldehyde, nonan -2-one, nonal -1-ol, and octanoic acid ethyl ester.

Methods

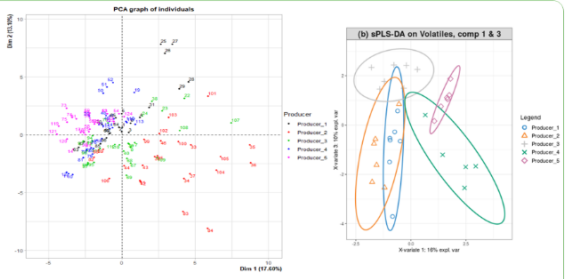
42 typical caciotta and caciotta like cheese were sampled from **5 closely located producers** (51±26 km range) for two years sampling. Core cheese samples of 2.5 g and 150 mg were employed for volatile organic compounds and DNA analyses, respectively. Three biological replicates were obtained for each the above-mentioned analyses. Total DNA was extracted using DNeasy PowerSoil kit. Libraries were prepared using NEXTERA kit and sequenced using NOVASEQ Sp500 PE250 (8 million reads per sample). Raw sequences were quality filtered (phred quality score of 35) and analyzed using bioBakery3 (Beghini et al., 2021).
Headspace thermal desorption coupled with gas chromatography mass spectrometry was performed using HiSorb probes desorbed with UNITY-xr (both from Markes International) combined with 5977B GC-MS (Agilent). The DB-5ms capillary column 60 m × 250 µm × 0,25 µm (Agilent) was used. The oven temperature program was as follows: initial 40°C held for 2 min, then ramped 3°C/min up to 180°C, and again ramped 20°C/min up to 260°C for 5 min, and held for 6 min. The constant flow rate of helium carrier gas was 1 mL/min. The MS analyses were done in a full scan mode (TIC mode), with a scan range of 33 to 350 amu. 44 standard molecules were injected to validate detected peaks.
Semi-quantitative analysis was performed with MassHunter quantitative analysis.
Taxa and peaks abundances were centered log-ratio transformed and analyzed with sPLS-DA model using mixOmics R package (Rohart et al., 2017). M-fold cross-validation of the training (25% of samples) dataset was perfomed using 10 folds and 100 nrepeats.



Unsupervised non-metric multidimensional scaling (NMDS) of bacterial communities at species level (n=124) and supervised sparse partial least squares – discriminant analysis (sPLS-DA) of the training dataset.



Unsupervised non-metric multidimensional scaling (NMDS) of viral communities at species level (n=124) and supervised sparse partial least squares – discriminant analysis (sPLS-DA) of the training dataset.



Unsupervised principal component analysis of volatile oorganic compounds (n=124) and supervised sparse partial least squares – discriminant analysis (sPLS-DA) of the training dataset.

Summary table of cross-validated models (training n=31, prediction n=93)

Parameter	Recall	Precision	Specificity	Classification accuracy rate (%)
Species Bacterial communities				
Producer_1	0.929	0.722	0.985	
Producer_2	1.000	0.880	1.000	
Producer_3	0.500	0.909	0.872	**83.871**
Producer_4	1.000	0.900	1.000	
Producer_5	0.895	1.000	0.967	
Species Viral communities				
Producer_1	0.444	0.889	0.828	
Producer_2	0.063	0.250	0.786	
Producer_3	0.944	0.586	0.975	**63.636**
Producer_4	0.650	0.704	0.974	
Producer_5	0.550	0.458	0.833	
Volatilome (VOC)				
Producer_1	0.714	0.526	0.930	
Producer_2	0.632	0.706	0.879	
Producer_3	0.600	0.522	0.864	**67.742**
Producer_4	0.632	0.800	0.879	
Producer_5	0.810	0.895	0.920	

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Specificity = \frac{TN}{TN + FN}$$

$$Classification\ accuracy\ rate = \frac{TP}{Total\ observations} \times 100$$

Summary table of models performance based on true positive (TP), true negative (TN), false positive (FP) and false negative (FN) outputs. Recall, precision, specificity and classification accuracy rate have been calculated according to Kassambara (2018).

Conclusions

DNA-based methods have been suggested as potentential tool to autheticate cheese origin. In this work, we evaluated the disciminative power of shotgun metagenomics using typical mountain semi-hard raw milk cheese obtained from 5 producers in a limited geographic area. Bacterial and viral communities were heavily affected by cheese origin. We further examinate the ability of shotgun metagenomics to autheticate cheese origin thought cross-validated models finding that, at species level, bacterial commuities are more relieable than viral comunities. Moreover, the classification accuracy rate of bacterial communities was higher than the volatilome. These promessing results suggest futher studies to confirm the ability of shotgun metagenomics to autheticate cheese origin and to compare the its output with accurate methods such as isotope and element fingerprinting.

Beghini, F., McIver, L. J., Blanco-Miguez, A., Dubois, L., Asnicar, F., Maharjan, S., ... & Segata, N. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife, 10, e65088.
Kassambara, A. (2018). Machine learning essentials: Practical guide in R. Sthda.
Rohart, F., Gautier, B., Singh, A., & Lê Cao, K. A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS computational biology, 13*(11), e1005752.