NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

SCHOOL OF SCIENCES

DEPARTMENT OF CHEMISTRY

MASTER OF SCIENCE IN 'FOOD CHEMISTRY'

MASTER THESIS

Milk authenticity assessment using High Resolution Mass Spectrometry - a metabolomics approach to investigate milk origin

KATSOGIANNI NEKTARIA

Chemist

ATHENS

OCTOBER 2023

MASTER THESIS

Milk authenticity assessment using High Resolution Mass Spectrometry - a metabolomics approach to investigate milk origin

KATSOGIANNI NEKTARIA

Registration Number: 7111132100007

SUPERVISING PROFESSOR:

Marilena Dasenaki, Professor NKUA

THREE-MEMBER EXAMINATION COMITTEE:

Proestos Charalampos, Associate Professor NKUA

Marilena Dasenaki, Assistant Professor NKUA

Thomaidis Nikolaos, Professor NKUA

DEFENDING DATE

OCTOBER 2023

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αξιολόγηση αυθεντικότητας γάλακτος με τη χρήση Φασματομετρίας Μαζών Υψηλής Διακριτικής Ικανότητας - μια μεταβολομική μελέτη για τη διερεύνηση της προέλευσης του γάλακτος

Κατσογιάννη Νεκταρία

Α.Μ.: 7111132100007

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:

Δασενάκη Μαριλένα**,** Επίκουρη Καθηγήτρια ΕΚΠΑ

ΤΡΙΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Προεστός Χαράλαμπος, Αναπληρωτής Καθηγητής ΕΚΠΑ

Δασενάκη Μαριλένα**,** Επίκουρη Καθηγήτρια ΕΚΠΑ

Θωμαΐδης Νικόλαος, Καθηγητής ΕΚΠΑ

ΗΜΕΡΟΜΗΝΙΑ ΕΞΕΤΑΣΗΣ

ΟΚΤΩΒΡΙΟΣ 2023

ABSTRACT

In recent years, food authenticity has become increasingly significant as a result of illegal practices, such as the labelling of low-quality products with misleading labels, so that the production cost can be minimized. Consequently, there is an imperative need for effective food control systems to protect consumers from adulterated food products and ensure their authenticity. Dairy product authenticity is a major public issue due to the significant nutritional and economic value of dairy products. Metabolomics is now becoming an emerging approach for the detection and identification of various quality parameters and chemical reactions that take place in food. The growing interest in the quality and safety issues of milk and dairy products, as well as the requirements for compliance with safety issues, have accelerated the need for the application of advanced analytical techniques and sophisticated metabolomic workflows.

The primary goal of the current study was to develop an analytical workflow for the detection of metabolites in Greek milk samples of different animal origin. To achieve that, target and suspect screening using Reversed-Phase Liquid Chromatography coupled to Quadruple Time-of-Flight Mass Spectrometry (LC-QTOFMS) and two different ionization modes (positive and negative) were applied for the analysis of 15 authentic samples of cow's milk, 10 authentic samples of sheep's milk and 10 authentic samples of goat's milk. More than 100 metabolites were detected and identified, belonging to the categories such as amino acids and their derivatives, organic acids, fatty acids, vitamins e.t.c. In combination with chemometrics, potential authenticity markers were investigated. Using PCA and PLS-DA models, the three types of milk were successfully distinguished according to their animal origin.

SUBJECT AREA: Food Chemistry, Analytical Chemistry, Food Authenticity.

KEYWORDS: Milk, authenticity, metabolomics, LC-QTOFMS, target screening, suspect screening.

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια, η αυθεντικότητα των τροφίμων γίνεται όλο και σημαντικότερη λόγω της εφαρμογής παράνομων πρακτικών, όπως η επισήμανση προϊόντων χαμηλής ποιότητας με ψευδείς ετικέτες, έτσι ώστε να ελαχιστοποιηθεί το κόστος παραγωγής. Συνεπώς, υπάρχει επιτακτική ανάγκη για αποτελεσματικά συστήματα ελέγχου τροφίμων για την προστασία των καταναλωτών από νοθευμένα τρόφιμα και τη διασφάλιση της αυθεντικότητάς τους. Η αυθεντικότητα των γαλακτοκομικών προϊόντων αποτελεί μείζον δημόσιο ζήτημα λόγω της οργανοληπτικής, διατροφικής και οικονομικής αξίας τους. Η μεταβολομική γίνεται πλέον μια αναδυόμενη μελέτη για την ανίχνευση και τον προσδιορισμό διαφόρων παραμέτρων και χημικών αντιδράσεων που λαμβάνουν χώρα στα τρόφιμα. Το αυξανόμενο ενδιαφέρον για τα θέματα ποιότητας και ασφάλειας του γάλακτος και των γαλακτοκομικών προϊόντων, καθώς και οι απαιτήσεις για συμμόρφωση με τα θέματα ασφάλειας, έχουν επιταχύνει την ανάγκη εφαρμογής αναλυτικών τεχνικών για τη μεταβολομική τους ανάλυση.

Ο στόχος της παρούσας μελέτης ήταν η ανάπτυξη ενός αναλυτικού πρωτοκόλλου εργασίας για την ανίχνευση μεταβολιτών σε δείγματα ελληνικού γάλακτος διαφορετικών ζωικών προελεύσεων. Για να επιτευχθεί αυτό, εφαρμόστηκε στοχευμένη και ύποπτη σάρωση με τη χρήση Υγροχρωματογραφίας Αντίστροφης Φάσης συζευγμένη με Φασματομετρία Μαζών υψηλής διακριτικής ικανότητας (LC-QTOFMS) και δύο διαφορετικούς τύπους ιοντισμού (θετικό και αρνητικό) και αναλύθηκαν 15 αυθεντικά δείγματα αγελαδινού γάλακτος, 10 αυθεντικά δείγματα πρόβειου γάλακτος και 10 αυθεντικά δείγματα γίδινου γάλακτος. Ανιχνεύθηκαν και ταυτοποιήθηκαν περισσότεροι από 100 μεταβολίτες, που ανήκουν σε κατηγορίες όπως αμινοξέα και παράγωγά τους, οργανικά οξέα, λιπαρά οξέα, βιταμίνες κ.ο.κ. Σε συνδυασμό με χημειομετρία, διερευνήθηκαν πιθανοί δείκτες αυθεντικότητας. Χρησιμοποιώντας μοντέλα πολυπαραμετρικής ανάλυσης PCA και PLS-DA, οι τρεις τύποι γάλακτος διακρίθηκαν επιτυχώς με βάση τη ζωική τους προέλευση.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Χημεία Τροφίμων, Αναλυτική Χημεία, Αυθεντικότητα Τροφίμων.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Γάλα, αυθεντικότητα, μεταβολομική μελέτη, LC-QTOFMS, στοχευμένη σάρωση, ύποπτη σάρωση

CONTENTS

INDEX OF FIGURES

INDEX OF TABLES

PREFACE

This master's thesis was carried out in the Laboratory of Food Chemistry and the Laboratory of Analytical Chemistry in the Department of Chemistry of the National and Kapodistrian University of Athens,under the supervision of the Assistant Professor Dasenaki Marilena.

I would like to thank my supervisor, Assist. Professor Dasenaki Marilena, for her guidance and support during my master. I am very happy to be part of her research group and grateful for all the opportunities and trust that I received. Also, I especially thank the members of the examination committee, Professors C. Proestos and N. S. Thomaidis, for their comments and remarks regarding my master thesis, and also professor V. Valdramidis and E. Kolia for developing my theoretical knowledge and practical skills in the field of Food Chemistry.

Special thanks to Dr. Anastasia Kritikou, who has been by my side throughout my study for the preparation of my thesis, guiding me in everything, giving me advice and fully supporting me. At the same time, I would like to thank Dr. Sofia Drakopoulou, who was also by my side for everything I needed.

I would like to acknowledge Dr. Konstantina Pasvanka and PhD candidate Konstantina Papastavropoulou for all the contribution in the laboratory and for the help I have received all the years that I have been in the Laboratory of Food Chemistry, which provides a friendly working environment for its students.

Last but not least, I would like deeply to thank my parents and my friends for their support and understanding.

CHAPTER 1

INTRODUCTION

1.1. Food Authenticity

Authenticity has become an issue of major concern in recent years and it refers to an agreement between producers and consumers. Food authenticity is the process of verifying that a product complies with its label description. Origin (species, geographical or genetic), production method (conventional, organic farming) and processing technologies (such as freezing, microwave heating etc.) are included in determination of authenticity. The perceived quality of a food is a very important factor for the consumer, who often associates it with the concept of authenticity and food safety [1].

1.2. Milk composition - authenticity assessment

The authenticity of milk is an important public issue because of the great nutritional and economic value of dairy products. Milk has a high nutritional value and is an excellent source of protein, fat, carbohydrates, minerals and vitamins. In particular, the average chemical composition of milk is as follows:

- Water: 87.5%
- Carbohydrates: 4.6%
- Fat: 3.7%.
- Proteins: 2.8%
- Minerals: 0.8%
- Vitamins: 0.6%

Water: Water is the dispersal medium for the other components of milk. It contributes to the transport of various substances and metabolic products, removing from the body anything that is not needed.

Carbohydrates: Lactose, the main sugar in milk, is the most abundant. It is a disaccharide (glucose and galactose) that gives milk its slightly sweet taste and also provides milk and its products nutritional value.

Fat: Composed of more than 98% triglycerides, it is also a carrier of the fat-soluble vitamins A, D, E and K. Finally, it influences the texture and organoleptic characteristics of dairy products.

Proteins: These are mainly caseins, serum proteins and enzymes. Caseins are responsible for the characteristic 'albumen' of milk and have a high digestive and biological value as they contain all the essential amino acids. Whey proteins are classified according to their solubility: proteases, peptones, albumin and globulin [2,3].

Minerals: These are mostly trace elements (eg. Calcium, Selen, Zinc, etc.).

Vitamins: Milk contains almost all vitamins (A, B6, B12, C, D, K, E, thiamine, niacin, biotin, riboflavin, pantothenic acid, etc.), some in sufficient quantities and others in trace amounts [2].

Growing interest in the quality and safety issues of milk and milk products and regulatory requirements to comply with safety issues have accelerated the need for effective food control systems to protect consumers from adulterated food products and ensure its authenticity.

1.3. Legislative frameworks in milk

Milk and dairy products are traditional foods in Europe and have been widely consumed for centuries. Because of the need to protect the identity of these and other agricultural products, Council Regulation (EC) No 2081/92 of 14 July 1992 introduced the following labels:

- Protected Designation of Origin (PDO),
- Protected Geographical Indication (PGI) and
- Traditional Speciality Guaranteed (TSG).

These labelled products (PDO, PGI and PGI) are particularly vulnerable to adulteration due to their high commercial and nutritional value, texture, flavour and limited raw materials. [4].

Figure 1: Protected Geographical Indication (PGI) and Protected Designation of Origin (PDO) trademarks [5].

- ***Regulation (EC) No 178/2002 of the European Parliament and of the Council of January 28, 2002:*** It refers to the definition of the general principles and requirements of food law, the establishment of the European Food Safety Authority and the definition of food safety procedures [6].
- ***Commission Implementing Regulation (EU) 2018/150 of January 30, 2018*****:** It relates to methods for analyzing and assessing the quality of milk and milk products [7].
- ***Regulation (EC) No 853/2004 of the European Parliament and of the Council of April 29, 2004*****:** Laying down specific hygiene rules for food of animal origin [8].
- ***Regulation (EU) No 2017/625 of the European Parliament and of the Council of March 15, 2017:*** It refers to the establishment of official controls and official activities carried out to ensure the enforcement of food and feed law, animal health and animal welfare rules, plant health, and plant protection products [9].

1.4. Milk fraud – Illegal practices

Food fraud is a serious issue that has come under increasing scrutiny as a potential food safety and public health concern. All foods have the potential to be adulterated, but those that are more expensive or produced in climatically unstable harvesting conditions are at greater risk. Adulteration is a more specific definition of food fraud. Adulterated food is food that has either been deliberately altered or may have been spoiled by adding an unauthorized chemical ingredient. Generally, adulteration of a food is defined as the deterioration of its quality either by the addition or removal of any substance to or from it and can occur at any stage of the production, processing and storage of the food [10,11].

There are many types of milk adulteration, such as diluting milk with water, upgrading low quality milk by adding solids, proteins or harmful chemicals to increase the nutritional content, or even completely disguising low quality milk as a higher quality product [11]. Increasing fraud in milk and dairy products is having a profound impact on market order and reducing consumer confidence in dairy companies. The three most commonly consumed types of milk are:

- cow's milk,
- goat's milk
- sheep's milk and buffalo milk.

These types of milk have similar organoleptic characteristics, appearance and chemical composition and play an important role in human nutrition worldwide. For this reason, fraud and adulteration are common among these species, due to various factors such as production and feeding costs. For example, the prices of buffalo and goat milk are higher than those of cow's milk in order to make a higher profit. It is therefore clear that the development of methods to detect milk adulteration is essential [12].

1.5. Food analysis using Omics Technologies

The agro-food chain in the European Union is subject to ever stricter standards and quality control and monitoring methods due to constantly changing legislation. In risk assessment, profiling and 'omics' technologies are used to corroborate and enrich data rather than to replace current analysis [13]. The use of omics technologies in food analysis (foodomics) is growing rapidly. In the midst of this evolution in food analysis, the term 'foodomics' has been defined to integrate the use of advanced omics technologies together with biostatistics, chemometrics and bioinformatics, to enable the assessment of complex biological systems and the mechanisms of bioactive food compounds that may affect them [14]. Also, foodomics is a powerful tool for investigating many nutritional and health issues in food production, processing, storage, transport and marketing. Further research in this area will help to better characterize the biological, functional and safety properties of food.

Foodomics is a new, powerful approach for studying the food sector together with the nutrition sector to improve human health and well-being. Omics technologies include:

- Genomics: DNA-based approaches have become commonplace in food analysis, allowing to detect food fraud at low concentrations in complicated and highly processed food matrices,
- Transcriptomics: transcriptomics is defined as the study of the transcriptome - the complete set of RNA, also known as expression profiling as it is the study of the expression levels of mRNAs in a given cell population,
- Proteomics: as proteins can be markers for different characteristics of food components and products, proteomics-based tools have been used for food quality purposes such as food authentication,
- Metabolomics: metabolomics consists on the identification and quantification of the metabolites present in a biological system [11,15].

Metabolomics is now becoming an emerging study for the detection and identification of various parameters and chemical reactions taking place in food.

Metabolomics applications:

- Molecular characterization of food (identification of food components)
- Assesment of their quality - authenticity
- Tracking changes in food composition during processing
- Monitoring of metabolic markers in humans after consumption of specific foods (dietary interventions).

In the present study, the authenticity of milk was investigated as an application of metabolomics.

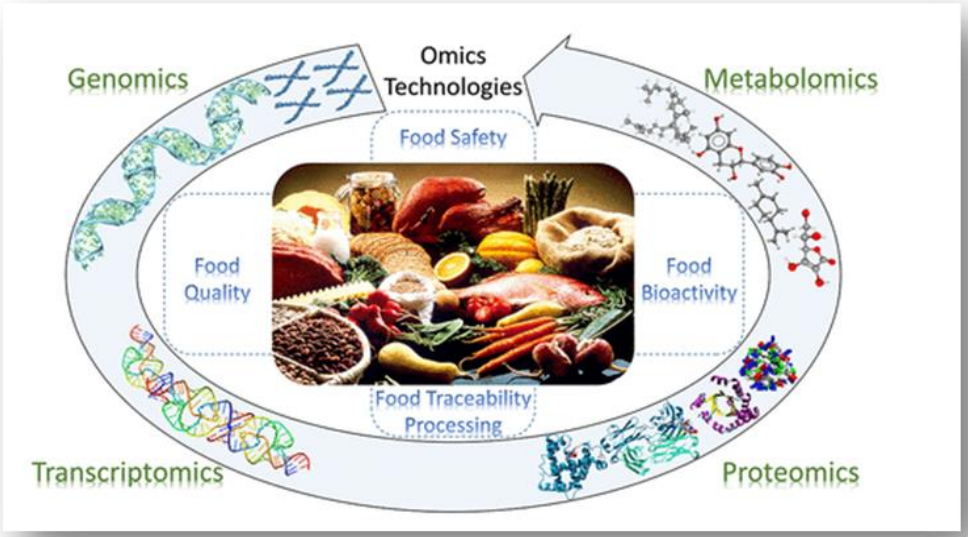


Figure 2: Schematic representation of the omics technologies and areas of food science covered by Foodomics [14]**.**

Foodomics provides information about the origin of the product (geographic, production) and its composition (illegal additives, other impurities). In this way, it ensures the safety, quality and global traceability of food products.

1.6. Milk metabolites – Variability factors

The identification of milk metabolites is important in both animal science and dairy science for the assessment of milk composition and quality. A metabolite

is an intermediate or end product of a metabolic process. This term is usually used to describe molecules of low molecular weight (< 1500 Da) with biological activity. Certain bioactive molecules such as peptides, nucleotides, oligosaccharides, immunoglobulins and immune proteins have received much attention in dairy science. Metabolites also include nutrients derived either from dairy animals or from milk processing, fermentation, etc. Nutrients are chemicals found in food that are essential for growth, maintenance and metabolism, including proteins, carbohydrates, fats, minerals and vitamins. Small molecules such as amino acids, fatty acids, sugars, etc. are metabolites of nutrients [16].

Metabolites represent, to some extent, the quality of milk and dairy products in terms of nutritional value, authenticity and safety, as the levels and types of metabolites can vary during the 'farm to fork' journey and are influenced by many factors. The species and breeds of dairy animals are considered to be a primary factor in determining milk quality, as a large proportion of metabolites are derived from their cellular activities and metabolism [17].

The profile of each metabolite is highly variable and influenced by several factors, including:

- the breed of the animal
- the stage of lactation,
- diet,
- seasonal variations,
- the number of somatic cells,
- processing parameters, etc.

In addition to nutritional value, these factors influence various technological and physico-chemical properties of the milk, such as thermal stability, coagulation properties and fermentation quality. Several microorganisms present in the mammary gland produce metabolites in the milk where they exhibit metabolic activity. Inflammatory conditions such as mastitis, pregnancy, etc. alter the paracellular pathway in the mammary gland, resulting in changes in the metabolic profile of the milk. Metabolic diseases, microbial secretions and enzyme reactions can also affect milk synthesis and yield [17–19].

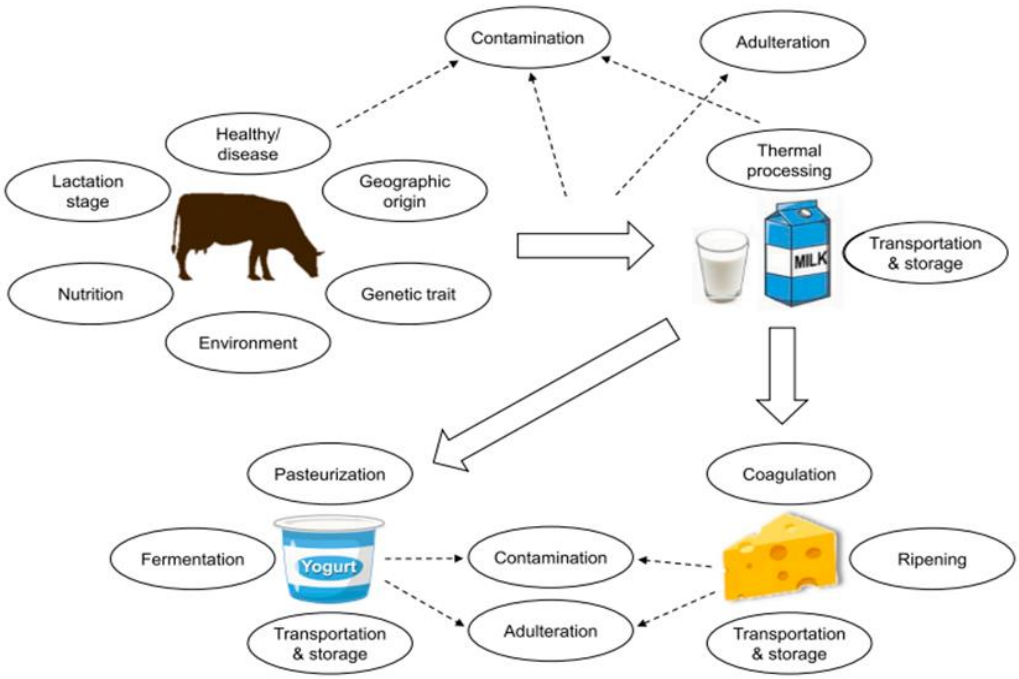


Figure 3: Factors affecting the quality of milk and milk products (from farm to table products)[17]**.**

1.7. Metabolomic analysis of milk

The use of metabolomics in dairy science has developed significantly in recent years. Important aspects and biomarkers related to animal health, production, authentication and contributors to the techno functional qualities of milk have been reported [20]. As mentioned above, metabolomics focuses on the analysis of small molecules, metabolites, in biological systems - the study of the overall metabolite profile in a system (cell, tissue or organism) under specific conditions. It involves the identification and quantification of small molecules involved in metabolic reactions. Metabolomics has recently been used to understand the nutritional quality and authenticity of milk and dairy products [18].

Some applications of metabolomics include:

- molecular characterization of food products (identification of food components),
- food quality - authenticity documentation
- monitoring changes in the composition of food during processing

- monitoring of metabolic indices in humans after consumption of specific foods (dietary interventions).

Metabolomics analyses a wide range of compounds at different concentrations with different physical and chemical properties using different analytical technologies. It often uses analytical instruments such as mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. Due to its high throughput and comprehensive coverage of metabolites, it has gained widespread application in clinical and pharmaceutical research and is steadily expanding into other areas including food, nutrition and plant research [17,18].

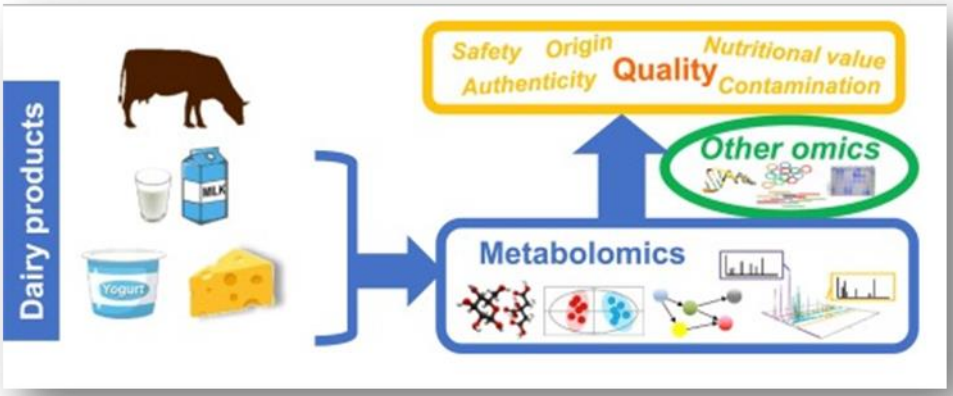


Figure 4: Graphical summary of the evaluation of the quality of milk and dairy products by metabolomic analysis[17] **.**

CHAPTER 2

HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC) - HIGH RESOLUTION MASS SPECTROMETRY (HRMS)

2.1. High Performance Liquid Chromatography (HPLC)

HPLC (High-Performance Liquid Chromatography) is the most widely used of all analytical separation techniques. It is a separation technique that can be used to analyse compounds with different properties from low to very high molecular mass. Its high sensitivity makes it suitable for separating non-volatile or thermally sensitive compounds and it is used to determine substances of industrial interest. In this chromatographic separation, a mixture dissolved in a mobile phase is transferred to another material called a stationary phase, and both mobile and stationary phases are immiscible. HPLC is an advanced form of column chromatography in which the mobile phase flows by a pump rather than by gravity. This speeds up the analysis and allows the use of chromatographic columns with small filler particles. The use of small particles increases the surface area of the static phase available to interact with the molecules carried by the mobile phase. This improves the separation of analyte molecules and significantly reduces the column size required for separation. Depending on the nature, chemical structure and molecular weight of the analytes, we can choose the type of liquid chromatography to be used. The HPLC method is used in food analysis not only for quantitative analysis of substances present in food, but also to verify the quality of the product with higher productivity. For milk and dairy products, HPLC methods could be considered as effective tools to perform measurements of: i) dairy product constituents; ii) amounts of additives used; and iii) environmental contaminants and xenobiotics [21]. A variety of detectors can be used in HPLC, including ultraviolet-visible (UV-Vis), fluorescence, electrochemical, diffractometer and others. Also, different types of HPLC have been developed to provide qualitative and quantitative information on the individual components of the sample under investigation. Ultra-High Performance Liquid Chromatography is a promising method that achieves fast, high-resolution separation while

minimising matrix interference by using small-diameter particles in the stationary phase and short columns.[22–24].

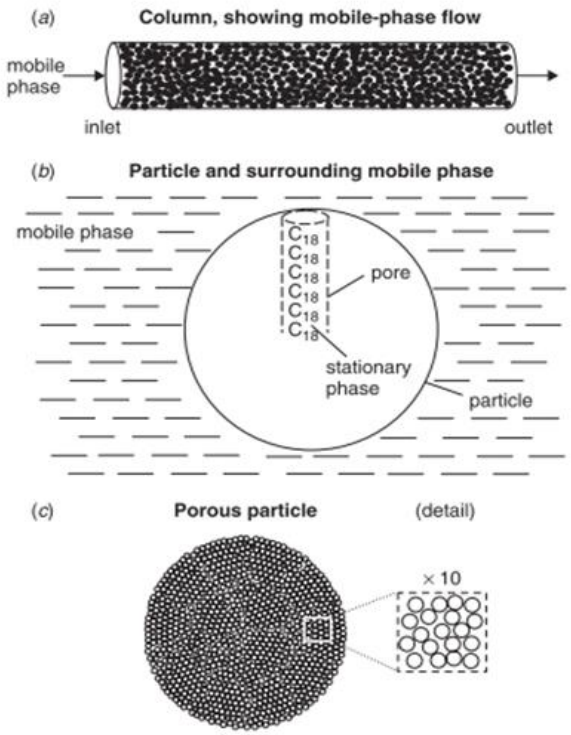


Figure 5:The HPLC column. (a) Column packed with spherical particles; (b) schematic of an individual particle, showing an idealized pore with attached C18 groups; (c) more realistic picture of a spherical, porous particle, showing detail (10× expansion) [23]**.**

2.1.1. Reversed Phase High-Performance Liquid Chromatography (RP-HPLC)

In the field of biological separation and purification, reversed phase chromatography has both analytical and preparative applications. Reversed-phase chromatography can separate molecules with a certain degree of hydrophobicity, such as proteins, peptides and nucleic acids, with high recovery and resolution. Because of its wide range of applications, reversed-phase chromatography is now the most widely used separation technique in HPLC. Over 65% of all HPLC separations are performed in reversed-phase mode. This is mainly due to the simplicity, versatility and breadth of the reversed-phase approach, which can handle compounds of different polarity and molecular mass [25].

The stationary phase (column, usually a C8 or C18 hydrocarbon) is non-polar, while the mobile phase is a polar mixture of water and an organic solvent (e.g., acetonitrile, methanol, water, etc.). The filler material used in reversed phase chromatography is usually chemically modified silica with hydrophobic groups such as alkyl and aryl groups. In reversed phase chromatography, the polar component is eluted first. In particular, the separation mechanism depends on the hydrophobic interaction between the molecule in the mobile phase and the immobilized hydrophobic ligand, i.e., the stationary phase. The initial mobile phase binding conditions used in reversed phase chromatography are predominantly aqueous, indicating a high degree of organized structure surrounding both the solute molecule and the immobilized ligand. As the solute binds to the hydrophobic ligand, the hydrophobic area exposed to the solvent is minimized. Therefore, the degree of organized water structure is reduced with a corresponding beneficial increase in system entropy. In this way it is energetically advantageous for the hydrophobic moieties, i.e., solute and ligand, to associate [25,26].

Reversed-phase chromatography is widely used in metabolomics because it is reliable, robust and repeatable, and has a well-understood separation process that covers a wide spectrum of molecular structures. Due to their ability to distinguish semi-polar compounds such as phenolic acids, flavonoids, glycosylated steroids, alkaloids and other glycosylated species, non-polar columns such as C18 are more suitable for metabolomics analysis. Normal phase separation techniques are uncommon for metabolomics research due to their low separation efficiency, poor repeatability and, more importantly, the use of solvents such as hexane that make them incompatible with mass spectrometry [27].

2.2. Principles of Mass Spectrometry

Mass spectrometry is certainly the technique that can find the greatest variety of applications, compared to the other analytical techniques available to a scientist. This is due to the fact that this technique provides information about:

- the elemental composition of the sample tested,
- the structures of inorganic, organic and biological molecules,

- the qualitative and quantitative composition of complex mixtures,
- the structure and composition of solid surfaces and finally
- the isotopic ratios of atoms in samples.

In the field of food analysis, mass spectrometry has gained a very central position, not only in research, but also in the routine food analysis laboratory, which is increasingly equipped with a number of MS systems. Mass spectrometry is a well-known detection technique that has many advantages, such as selectivity, sensitivity and the ability to analyze multiple samples. Therefore, It is a powerful analytical technique used to quantify known materials, detect unidentified compounds in samples, and elucidate the structure and chemical properties of various molecules. Compounds in a sample can be identified using the mass-to-charge ratio (m/z). By calculating the molecular weight of a compound and examining its isotopic abundance, the technique can identify a particular substance. By measuring the mass-to-charge ratio and relative abundance of the sample, a mass spectrometer ionizes it into gaseous ions, with or without fragmentation, which are then detected. The first step in the mass spectrometric investigation of a substance is the generation of gas phase ions of the molecule, primarily by electron ionization. This molecular ion is fragmented. Fragmentation occurs continuously for each major product ion of the molecular ion, etc. The mass spectrometer separates the ions according to their mass-to-charge ratio and they are detected in proportion to their abundance. The result is the mass spectrum of the molecule and it is shown as a plot of ion abundance against mass-to-charge ratio. The ions reveal details about the composition and structure of their precursor molecule. In the spectrum of a pure compound, the molecular ion, if present, appears at the highest m/z value (followed by ions containing heavier isotopes) and gives the molecular mass of the compound [28,29].

2.2.1. Mass Spectrometry Instrumentation

A mass spectrometer is an analytical instrument typically used to determine the molecular weight of chemical compounds or biomolecules by separating their ions in a vacuum according to their mass-to-charge ratio (m/z).

The three main sections of a mass spectrometer are:

- an **ion source**, which converts analyte molecule into gas-phase ions,
- a **mass analyzer**, which separates ionized analyte according to their mass-to-charge ratio, and
- a **detector system**, that detects the ions and measures the relative abundance of each resolved ion species (conversion dynode with secondary electron multiplier, multichannel plate).

In the **ion source**, the sample molecules are first introduced and ionized. The newly produced ions are then introduced into a **mass analyzer** where they are sorted according to their mass-to-charge ratio (m/z) in a vacuum using an electric or magnetic field. The next step is to move the gaseous ions to the **detector**, where an electric current is generated and amplified. The information from the amplified signals is then sent to the attached **computer** so that a **spectrum** can be generated and further analysis can be performed [30].

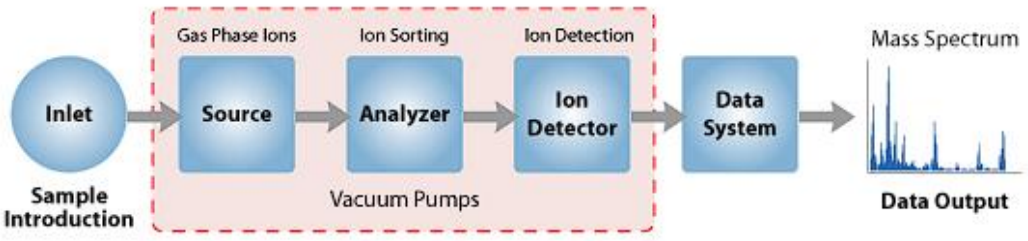


Figure 6:Components of a Mass Spectrometer [30]**.**

2.2.2. Ion source

The only type of substance or molecule that a mass spectrometer can detect is an ion. Radicals and neutral compounds cannot be detected. Ionization of

materials is therefore a crucial first step in mass spectrometry (MS) analysis. The introduction of sample molecules into the ion source initially causes ionization. In an ion source, neutral molecules can be ionized in a number of ways, including proton transfer, electron transfer, protonation or deprotonation, adduct formation, charge transfer or ion pair formation. Ionized molecules are propelled towards the mass analyzer by their newly formed ions. Several methods have been developed to ionize the analyte in mass spectrometry [31].

The ion sources used in mass spectrometry can be divided into two main categories:

- **desorption sources**: the solid or liquid sample is directly converted into gaseous ions. Gas ion sources are limited to thermally stable compounds.
- **gas-phase sources**: the sample is first evaporated and then ionized. Desorption ion sources can be used with non-volatile or thermally unstable compounds.

Ion sources are also classified into **soft** and **hard** ion sources.

- **Soft ion sources:** soft ion sources cause limited fragmentation. Consequently, the mass spectrum obtained with soft ionization sources often consists of the peak of the molecular ion and only a few additional peaks.
- **Hard ion sources:** In order to keep the analyte molecules in a high activation energy state during deactivation, they transmit enough energy to them, causing additional bond breaking and the formation of ionic fragments with ratios lower than the molecular ion [28].

2.2.2.1. Electrospray Ionization (ESI) Process

Electrospray ionization mass spectrometry (ESI-MS) has become a key tool in clinical laboratories over the past decade. It provides a sensitive, robust and reliable tool for analyzing non-volatile and thermally unstable biomolecules in microliter sample volumes that cannot be studied by other traditional methods.

Electrospray belongs to the sources of desorption ionization and it is classified as a soft ionization source, so the fragmentation of the molecules is small. Ionization takes place at atmospheric pressure and temperature. During electrospray ionization, electrical energy is used to move ions from solutions into the gas phase, where they are then subjected to mass spectrometry analysis.

There are several main steps in the process of electrospray ionization:

a. the liquid solution of the sample is sprayed through a capillary tube to which a high voltage of 2.5 - 6.0 kV is applied. High voltage droplets with the same polarity as the capillary voltage produce a mist. Higher sample flow rates can be achieved by using nebulizing gases, such as nitrogen, which shear around the eluted sample solution,

b. the solvent is then evaporated with the aid of an elevated ESI source temperature and/or another stream of nitrogen drying gas, forming tiny droplets that simultaneously charge the molecules of the analyte.,

c. the charged droplets are continuously reduced in size by evaporation of the solvent, leading to an increase of surface charge density and a decrease of the droplet radius and finally,

d. when the repelling electrostatic forces between the ions in the droplet are greater than the surface tension, it is kinetically and energetically possible for ions at the surface of the droplets to be ejected into the gaseous phase [32].

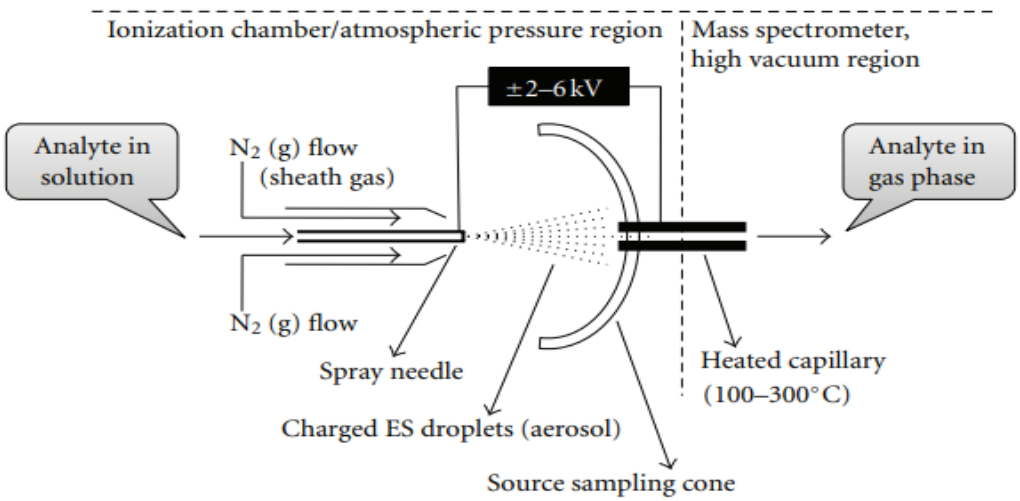


Figure 7: A schematic representation of the ESI-ion source [33]**.**

Depending on the ionization, ions have positive or negative charge. When ions and molecules collide, the H atom is transferred to the ion, forming the protonated pseudomolecular ion [M + H] +, as well as some addition ions after cation attachment, such as [M + NH4] +, [M + Na] +, and [M + K] +. Conversely, in negative ionization, when ions and molecules collide, the H atom is detached from the ion to form the deprotonated pseudomolecular ion [M-H] -. To obtain structural information, the precursor ions of interest can be mass selected and further fragmented in a collision cell. The fragment ions can then be mass analyzed by a second mass analyzer in a tandem mass spectrometer system [34].

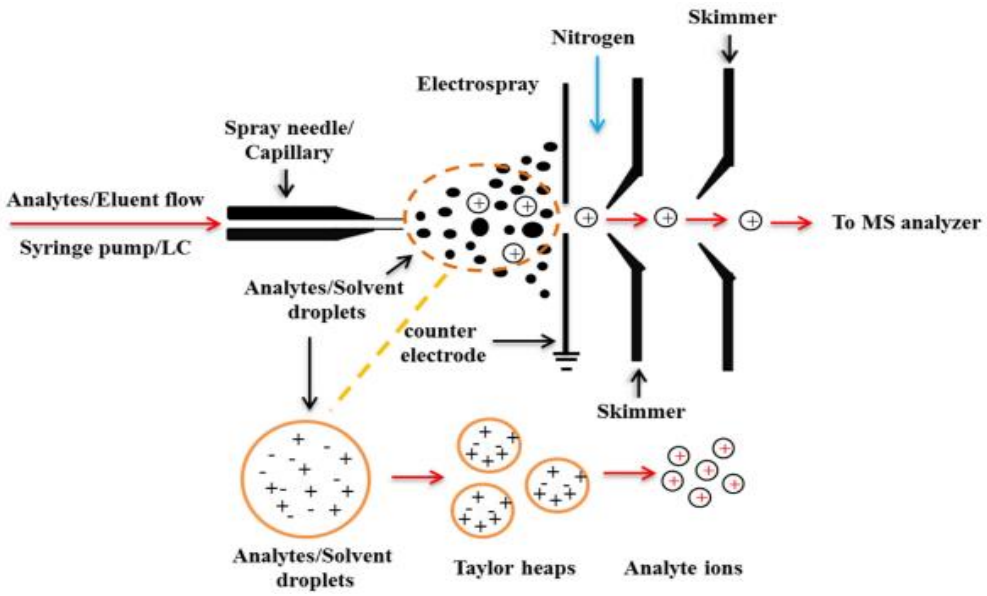


Figure 8: The schematic workflow of typical ESI ionization processes [35]**.**

2.2.3. Mass Analysers

A primary goal of mass spectrometry is the identification of an analyte, particularly in the presence of other analytes, based on the the measurement of the mass of ions. The function of a mass analyzer is to separate ions based on their mass-to-charge ratio. There are now six main categories of mass analyzers, as well as their combinations and hybrids, which measure ions in different ways. The two broad categories include:

- **Beam-type analyzers:** In beam analyzers, such as quadrupole and time-of-flight, Ions leave the ion source in a beam and travel through the analyzing field to the detector.
- **Trap-type analyzers:** In a trapping analyzer, such as a quadrupole ion trap, Fourier Transform Ion Cyclotron Resonance (FT-ICR) or orbitrap, the ions are often trapped in the analyzing field after being generated within the analyzer or injected from an external ion source.

The use of a particular type of mass spectrometer depends on a number of variables, including:

- the desired m/z range to be analyzed,
- the required **resolving power** of the analyzer,
- the **accuracy of the mass** measurement,
- the ability of the analyser to interface with an ion source or chromatographic instrument,
- the **scanning speed**, and
- the limit of detection or quantification required for the analysis [30,36].

In order to provide an instantly measurable ion current, an ideal mass analyzer must discriminate between slightly different masses and allow a number of ions to pass through.

- **Resolution:** Since only one elemental composition can contribute to a given mass spectral peak, accurate mass measurement requires the highest possible mass resolution. Mass resolution is conventionally defined as the closest distinguishable separation between two peaks of

equal height and width, the difference of two adjacent values m / z (m2 - m1), expressed in ppm.

$$\frac{m2 - m1}{m1} = \frac{\Delta m}{m1}$$

- **Resolving Power (R):** The ability of a mass spectrometer to distinguish between adjacent peaks in a mass spectrum is often referred to as "resolving power". The separation of peaks for singly charged ions can be expressed as a mass difference **Δm** (Full Width at Half Maximum, FWHM) and the **ratio m/Δm** is commonly used as a quantitative indicator of how well a mass spectrometer separates ions.

$$R = \frac{m}{\Delta m}$$

Two peaks are said to be separated if their overlap exceeds some fraction of their height, usually 10% [37].

- **Mass Accuracy:** The difference between the measured value (accurate mass) and the true value (exact mass) is the “accuracy” of the “accurate mass measurement” [38].
- **Scan rate/speed:** The scan rate of a mass spectrometer describes how fast it scans a mass spectrum. This is important for chromatography applications where the full mass spectrum must be scanned faster than the elution time of the chromatographic peak. A single chromatographic peak should ideally acquire at least ten full mass spectra [36,37].

Mass spectrometers can also be divided into three categories based on their resolution:

- Low Resolution Instruments,
- High-Resolution Instruments and
- Ultra-High Resolution Instruments,

which will be discussed in more detail in the next section.

2.3. High Resolution Mass Spectrometry

All mass spectrometers measure the mass-to-charge ratio of ions (m/z). However, the precision of this measurement depends on the type of mass spectrometer used. The effectiveness of mass analyzers is often measured in terms of mass resolution, mass accuracy, dynamic range, tandem analysis capabilities and acquisition speed. Focusing on mass resolving power and mass accuracy, low- (LRMS) and high-resolution mass spectrometry (HRMS) can be differentiated. High resolution mass spectrometry is the use of mass spectrometers with resolution (>50,000 at m/z 400) and mass accuracy (≤2 ppm. Since HRMS instruments have a higher resolution, it is possible to identify isotope distributions and generate fragmentation patterns, increasing the accuracy of chemical formula prediction and library matching for compound identification. Compared to LRMS systems, which can only achieve a so-called "nominal mass", HRMS systems are known to be able to discriminate between substances based on their exact masses. A low-resolution mass spectrometer is used for general analysis or simple compound identification. It provides the mass of the molecule with two decimal places. However, there may be more than one chemical compound with the same m/z ratio, which is a limitation of a low-resolution mass spectrometer. In these cases, we need high resolution to separate the peaks of these compounds. With high-resolution MS we get a mass peak with up to 4-5 decimal places) [39,40].

We categorize mass spectrometers according to their resolution as follows:

- Low Resolution instruments: Quadrupole (**Q**), Quadrupole Ion Trap (**IT**),
- High Resolution instruments: Time of Flight (**TOF**) and Magnetic sector analyzers,
- Ultra-High Resolution instruments: Fourier transform ion cyclotron resonance (**FT-ICR**), Orbitrap

In general, high-resolution time-of-flight (TOF) instruments have resolving powers in the tens of thousands, whereas low-resolution mass spectrometers, such as quadrupole filters and ion traps, have unit mass resolution (resolving power ~1000). The resolving power of TOF instruments is affected by molecular mass, with lower resolution found for lighter compounds and improving with

increasing compound molecular mass. Mass spectrometers such as Orbitraps and FT-ICRs can resolve masses in the hundreds of thousands to millions. Massive technological advances have been incorporated into a new generation of hybrid instruments, Q-TOFs and Q-Orbitraps, which allow specific ion fragmentation to be carried out, bringing an additional dimension to the possibilities of identifying unknown compounds (MS/MS spectra or fragment ion spectra by HRMS) [40].

2.4. High-performance liquid chromatography coupled with mass spectrometry (HPLC/MS)

Coupled analytical methods are those that integrate two or more analytical tools to create an entirely new, more efficient instrument. Coupling MS to chromatographic techniques has always been desirable due to the sensitive and highly specific nature of MS compared to other chromatographic detectors [41]. LC/MS is an analytical technique in which analytes are physically separated and then detected by their mass. Its sensitivity, selectivity and accuracy have made it a popular approach for detecting microgram or even nanogram levels of a range of analytes, including drug metabolites, pesticides, food adulterants and natural product extracts. Proteomics, metabolomics and lipidomics are three of the new '-omics' technologies, most of which are based on LC-MS. It provides both structural and quantitative data and can be used broadly or specifically, for example to identify thousands of proteins from a tissue or to detect physiologically active metabolites at levels as low as a few parts per billion [42].

LC separation: Some analytes interact more strongly with the stationary phase than others, causing them to separate as they move through the column. The analytes that interact least with the stationary phase are the first to leave the column. The remaining analytes are eluted sequentially, with the analytes with the strongest interactions eluted last as the mobile phase continues to flow through the column. The remaining time of an analyte in the column is known as its retention time (RT) and is indicative of that analyte. In Liquid Chromatography, the peaks of a sample are identified by comparing their

retention times with those of standards. However, many substances have similar retention times and cannot be identified or separated and as a result they cannot be quantified.

MS separation: After travelling through the mass analyser, the ions are directed to different areas of the detector according to their mass/charge (m/z) ratio. As the ions come into contact with the detector, a computer system generates and records usable signals. The signals are graphically displayed by the computer as a mass spectrum, which shows the relative abundance of the signals based on their m/z ratio. However, if the compounds are part of a mixture, the mass spectrum will be complicated because it would include ions from each component.

As a result, it is impossible to make an accurate identification, especially if the detected chemicals are present in small quantities. Therefore, by combining the two methods, identification can be accurate because molecules with similar retention times will have different mass spectra. Particularly in the field of milk, HPLC coupled to mass spectrometry (MS) has been used to:

i. identify and quantify some milk constituents such as vitamins, lactose, protein, etc.
ii. evaluate the authenticity and traceability of milk and
iii. detect contaminated dairy products [32].

2.5. QTOFMS Instrument

Analytical techniques such as Q-TOF-MS combine the benefits of two separate mass spectrometers. A special 'hybrid' analyser has been developed by combining the high compound fragmentation efficiency of quadrupole technology with the fast analysis speed and high mass resolution capacity of time-of-flight.

Unlike magnetic sector and quadrupole MS, **Time-Of-Flight (TOF)MS** is a pulsed and non-scanning MS. Two main types of TOF instruments can be found in the market, the LC-TOF-MS and LC-QTOF-MS, which differ mainly in the capacity or not to perform MS/MS experiments. In Time-of-Flight instruments, positive ions are periodically generated by bombarding the sample with short

pulses of electrons, secondary ions or laser photons. The ions produced are accelerated into a field-free tube with an electrical pulse of 10^3 to 10^4 V. The mass separation of the ions takes place as the ions fly towards the detector at the end of the tube. Since all ions entering the tube have the same kinetic energy, their velocity through the tube is inversely proportional to their mass, so that the lighter particles reach the detector faster than the heavier ones. The time of flight is given by the relation:

$$t_F = \frac{L}{v} = L\sqrt{\frac{m}{2zeV}}$$

where L is the distance from the source to the detector. Flight times typically vary from 1 to 50 ms [28].

TOF-MS can operate at very high repetition rates, typically between 5 and 30 kHz, or 5,000 and 30,000 raw mass spectra per second. In TOF-MS there are two possible approaches: instruments with very high resolution (5,000-10,000) but moderate scan speed (10 Hz) and instruments with a speed of 100-500 spectra/sec but unit resolution. In the context of non-target identification in complex matrices, the former approach is best suited. TOF-MS has become an effective tool for the unequivocal detection of non-targets and for compound confirmation because it can accurately capture full scan spectra over the entire acquisition range. TOF-MS allows identification and structural elucidation of target and non-target compounds in a sample, unlike conventional monitoring techniques using SIM or MRM.

Time-of-flight instruments offer several advantages over other types of mass spectrometers. These include:

- ✓ simplicity,
- ✓ stability,
- ✓ high mass accuracy,
- ✓ high acquisition speed,
- ✓ absence of spectral distortions when used in conjunction with fast separations and narrow chromatographic peaks,
- ✓ the ability to easily approach the ion source,

- ✓ a nearly unlimited m/z range and
- ✓ good adaptation to pulsed ionization sources.

However, they have the disadvantage of limited resolution and sensitivity [42].

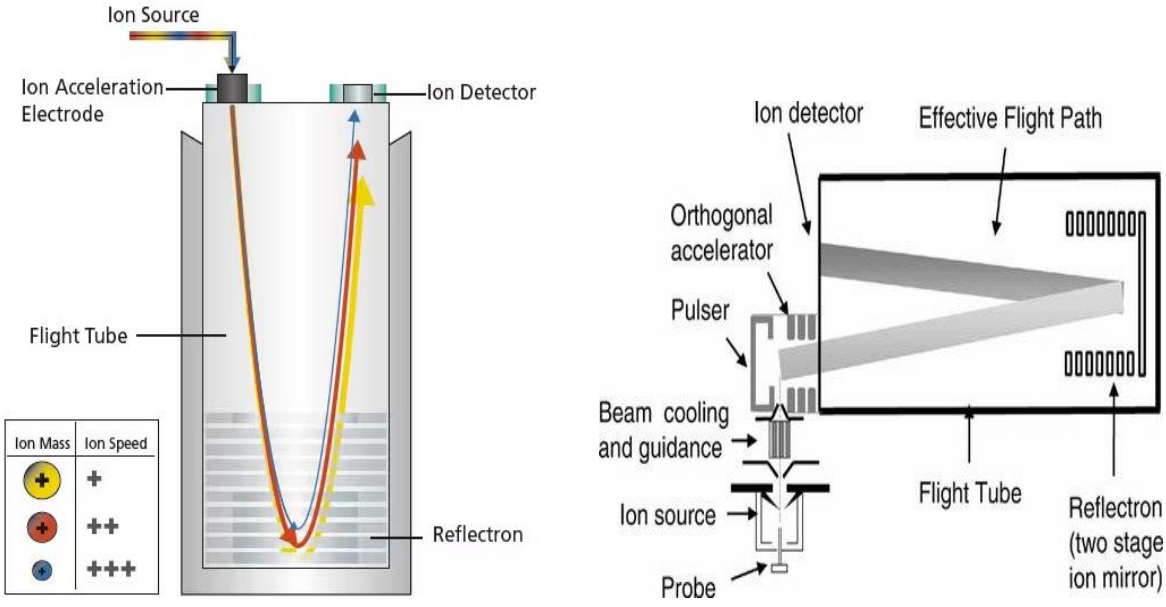


Figure 9: Schematic of a TOFMS [43,44]**.**

The hybrid quadrupole time-of-flight mass spectrometer (QTOF), which combines an MS1 and collision region modified from a triple quadrupole instrument with an MS2 reflectron orthogonal acceleration time-of-flight analyser, provides the most robust detection of target analytes. Although the third quadrupole has been replaced by a time-of-flight tube, the Q-TOF-MS instrument is very similar to a triple quadrupole mass spectrometer. The first quadrupole (Q1) can be used as a mass filter to select particular ions on the basis of their mass-to-charge ratio (m/z), or it can be used in radio frequency (RF) only mode to emit all ions. In a process known as Collision Induced Dissociation (CID), the second quadrupole (Q2) acts as a collision cell where ions are struck by neutral gas molecules such as nitrogen or argon. The Q2 can also operate in RF-only mode without ion fragmentation. The ions are re-accelerated as they leave the quadrupole and enter the ion modulator region of the time-of-flight analyser, where they are pulsed by an electric field and accelerated orthogonally to their initial path. In the flight tube, a region of field-

free drift where mass separation takes place, all ions have reached the same kinetic energy. While heavier ions take longer to travel the flight path to the detector, lighter ions travel further and faster. Using a reflectron device, modern time-of-flight analyzers can also account for the spatial scattering and kinetic energy dispersion of ions with the same m/z but different velocities. This reflectron adjustment allows ions with the same mass/charge ratio to enter the detector simultaneously. The reflectron system extends the flight path and improves mass resolution.

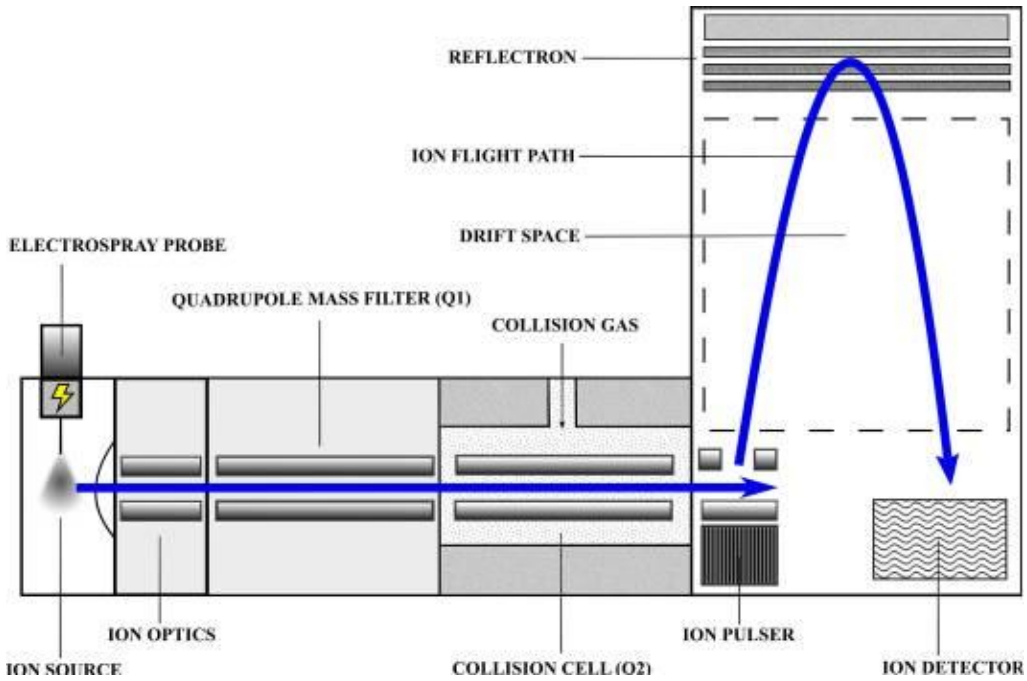


Figure 10: Schematic diagram of a quadrupole time-of-flight mass spectrometer [45]**.**

QTOF has the capability for single MS as well as MS/MS operation modes. In the first mode, known as single MS mode, the first and second quadrupoles are used in RF-only mode to provide an accurate mass scan of the unfragmented precursor ion. The Q1 can also be used to select a specific mass or range of masses to be sent to the TOF analyser. In the second mode (MS/MS) the Q1 can be used in RF-only or mass filter mode to send ions into the collision cell (Q2) where CID occurs. The subsequent product ions and any unfragmented precursor ions are transferred to the TOF analyzer where accurate mass measurement takes place. The Q-TOF-MS can collect information on both precursor and product ions simultaneously by switching between two modes.

Ion detection is performed by a detector device called a time-digital converter, which converts the time of flight of the ion into a mass signal [45].

2.6. HRMS screening workflows

Three main data processing strategies are currently recognized, including:

- **Target screening:** Targeted analysis focuses on the study and identification of a specific group of analytes and generally refers to the identification and quantification of a set of pre-selected analytes. This strategy is more limited, reaching a relatively smaller number of metabolites. Identification is performed using standard solutions. A reference standard is therefore required to compare and match the experimental retention time (t_R) and MS/MS fragmentation to determine the concentration of the target compound in the sample [46,47].
- **Suspect screening:** Suspect scanning is a data processing strategy to identify 'suspect substances' selected on the basis of prior information such as literature, prediction models and so on, but for which no reference standards are available. It is a list of potential metabolites as it involves the selection of existing analytes and the use of custom databases containing information on accurate mass, adduct ion formation and fragment spectra, MS/MS. Specifically, LC-HRMS suspect screening is based on accurate mass and isotope information for the precursor ion. Compounds expected to be present in the samples can be tested using the accurate mass of their molecular ions in negative ([M-H]-) or positive ([M+H]+) ionization. The exact mass of each suspected compound is extracted from the chromatogram and evaluated [48].
- **Non-targeted screening:** Non-targeted analysis is based on the study to detect and identify metabolites for which there is no prior knowledge. It is usually preceded by a targeted scanning. It covers a wider range of molecules and can be useful for identifying new metabolites. Non-targeted analysis can be divided into two types: fingerprinting and profiling.

Fingerprinting involves the analysis of the largest possible number of compounds in a matrix, including their detection and statistical processing of the results, without the need for identification and quantification. It allows the classification of sample components into different regions based on metabolite patterns to obtain a typical fingerprint sample, which is faster and more convenient.

On the other hand, **profiling** focuses on the analysis of a class of metabolites, often identified and quantified using databases or software tools [46].

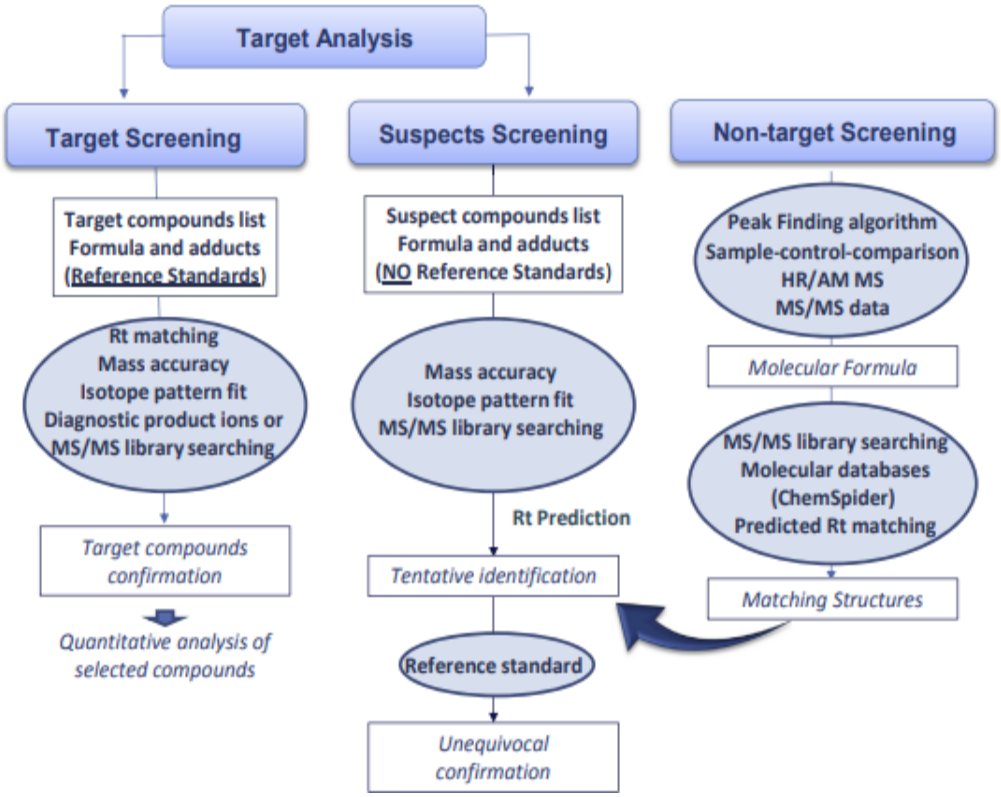


Figure 11: Main workflows used in high-resolution mass spectrometry (HRMS) methods [48]**.**

2.6.1. Levels of identification

The ability to identify and detect compounds in different samples has been greatly enhanced by the increasing use of high-resolution mass spectrometry (HR-MS) in chemical analysis. Because it is not always feasible or useful to synthesize or confirm each substance by complementary methods (e.g. nuclear magnetic resonance), the confidence in these HR-MS-based identifications varies from study to study and compound to compound. It can be very challenging to communicate these different levels of confidence in a clear and accurate manner [49].

Mass spectrometry chemical identification is the most powerful qualitative technique. Five levels of confidence, especially performed for identification have been established.

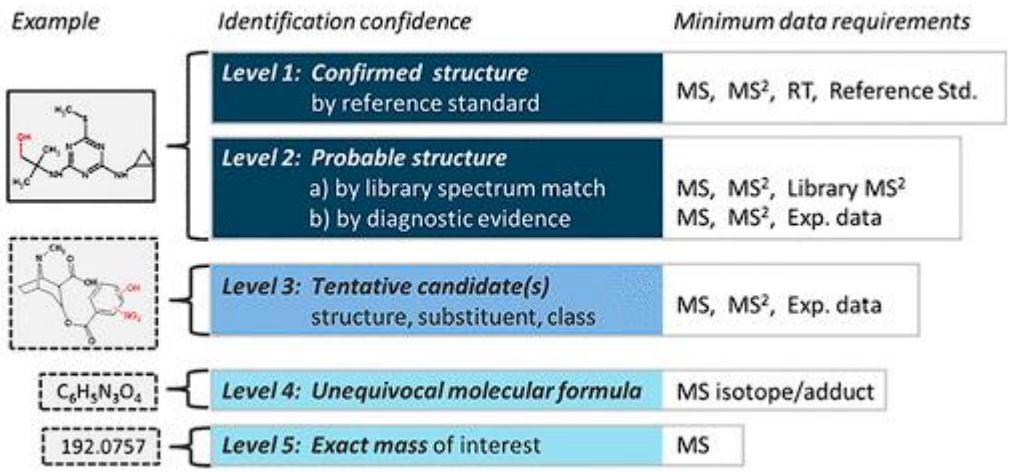


Figure 12: Proposed identification confidence levels in High Resolution Mass Spectrometric analysis [49]**.**

Levels 1 and 2 refer to structure identification

- **Level 1:** A structure that has been confirmed by adequate measurement of a reference standard by MS, MS/MS and retention time matching is optimal. An orthogonal approach should also be adopted if at all possible. Here the means of identification is co-analysis by comparing the properties of an analyte and an authentic standard/reference material in the same experiment. In co-mass spectrometry identification, spiking a sample with a standard of an analyte does not result in significant distortion of the mass spectral signals. Combined with co-

chromatography, this is the strongest evidence for the presence of the analyte [49,50].

- **Level 2:** Probable structure indicates that it is possible to propose an exact structure based on various pieces of evidence. The means of identification here is comparison with experimental reference data.

  2a: *Library*, this involves matching literature or library spectral data where the spectral-structural match is unambiguous. Care should be taken when comparing spectra acquired with different acquisition parameters (e.g. resolution, collision energy, ionization, MS level) to ensure the validity of the match and decision criteria should be clearly stated.

  2b: *Diagnostic*, refers to the situation where no other structure can fit the experimental data, but no available standards or literature material can be used to corroborate it. Evidence may include parent compound information, the experimental diagnostic MS/MS fragments and/or ionization behavior.

The most reliable data originate from the same laboratory and analytical instrument and replicate/successive experiments (successive spectral scans and chromatographic runs) [49,50].

Levels 3 to 5 refer to the substance class, the formula or the "mass of interest".

- **Level 3**: It describes a tentative candidate or candidates that represent a "grey zone" where there is evidence for one or more potential structures but inadequate data for just one specific structure (such as positional isomers). A few candidate molecular/fragment formulas are generated on the basis of a close similarity between experimental and theoretical masses (≤1÷5 ppm) and also that of isotope ratios [49,50].
- **Level 4**: It refers to an unambiguous molecular formula, which is possible when a formula can be uniquely assigned using spectral information (e.g. adduct, isotope and/or fragment information), but there is insufficient evidence to suggest other possible structures.

- **Level 5**: It refers to a precise mass (m/z) that can be determined in a sample and is of particular importance to the investigation, but lacks sufficient detail to even give a formula. These masses can be traced in other investigations by screening and non-target methods, although level 5 indicates that there is no clear knowledge of the structure or formula [49,50].

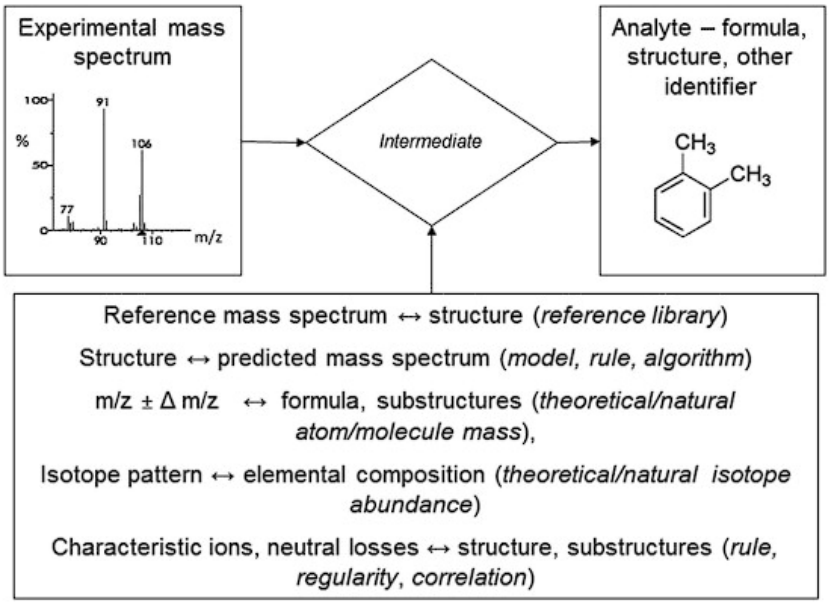


Figure 13: Idea of chemical identification. The chemical nature of an analyte is deduced with the involvement of various intermediates (e.g., data, correlation, regularity, rule, algorithm, and program) [50]**.**

2.7. Chemometrics – Classification Methods

Chemometrics is a fast and efficient way to analyze food matrices beyond univariate dimensionality, revealing hidden information in complex instrumental datasets. Chemometrics is defined as "the chemical discipline" that uses mathematical and statistical methods to:

a) design or select optimal measurement procedures and experiments and
b) provide maximum chemical information through chemical data analysis.

Concerning Food Research and Innovation, chemometrics can be used to help

- ✓ facilitate the exploration and interpretation of chemical and biological information obtained through instrumental analysis,
- ✓ identify chemical, biochemical or biological markers associated with food quality and safety,
- ✓ identify the authenticity, quality, biofunctional and nutritional aspects of food based on geographical and botanical origin, presence of undisclosed food additives or contaminants,
- ✓ construct robust mathematical models to monitor the operation of food processing units and predict the shelf-life of processed foods and
- ✓ understand the interactions between sensory modalities and compositional, physical, textural and microstructural properties of food [51].

Chemometrics, especially multivariate analysis, has various applications in quality control, quantitative and qualitative assessment of chemical properties, and determination of food authenticity. Chemometrics provides powerful tools for the calibration analysis of spectroscopic and chromatographic data, which can be used in both targeted and non-targeted techniques to detect various food fraud situations or to authenticate their geographical or biological origin.

Classification methods are key multivariate techniques that aim to find mathematical models capable of determining the membership of each sample to its correct class based on a set of measures. Once a classification model has been constructed, the membership of unknown samples to one of the defined classes can be approximated. In other words, classification models seek mathematical correlations between a set of descriptive variables resulting from chemical measurements and a qualitative variable (such as category membership, geographical origin, labels, etc.). Classification approaches based on multivariate data analysis might be **supervised** or **unsupervised**. Unsupervised approaches seek to find clusters or correlations between samples without prior knowledge of classifications or groups. In contrast, supervised approaches require knowledge regarding class membership as well as a training phase to construct an appropriate mathematical model. Unknown objects of a test or validation set can then be predicted [52].

The method used to build the mathematical relationship can be categorized into two groups: supervised classification (like PLS-DA or LDA) and unsupervised classification methods (PCA and clustering technique).

2.7.1. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is the most widely utilized unsupervised method in food authenticity investigation. PCA is often used as an initial step in data analysis to uncover or validate patterns in measurable data. PCA transforms measured primary variables into new independent variables, called principal components (PCs), which are linear combinations of the original variables. PCA generates a series of orthogonal axes that indicate the direction of the greatest (remaining) variation in the data. The first principal component (PC1) accounts for the largest proportion of the total variance, the second (PC2) is orthogonal to the first and lies in the direction of the largest remaining variation, and so on until all the variance is accounted for. By reducing the dimensions of the data, PCA allows it to be viewed while retaining as much information as possible from the original data [53].

PCA is probably the most popular multivariate statistical technique and it is widely used for handling LC-HRMS data. Through LC-HRMS, a peak list of thousand masses can be extracted, with their intensities varying across several samples analyzed. PCA analyzes the peak list generated by LCHRMS and its goal is to extract the important information from the peak list (loading plot) and to express this information as a set of new orthogonal variables, called Principal Components (PCs). PCA also represents the pattern of similarity of the observations and the exact masses by displaying them as points in maps (score plot) [54].

2.7.2. Partial Least Square Discriminant Analysis (PLS-DA)

PLS-DA, or Partial Least Squares Discriminant Analysis, is a supervised classification technique and also a linear classification method that combines the characteristics of partial least squares regression with the discriminatory power of the classification technique. PLS-DA is based on the PLS regression algorithm (PLS1 when dealing with a single dependent Y-variable and PLS2 when dealing with multiple dependent Y-variables), which searches for latent variables that have the greatest correlation with the Y-variables. The main advantage of PLS-DA is that it models the relevant sources of variability in the data using so-called latent variables (LVs), which are linear combinations of the original variables. As a result, it allows for graphical visualization and understanding of the various data patterns and relationships through LV scores and loadings. Scores are the coordinates of the samples in the LV projection hyperspace, whereas loadings are the coefficients of the variables in the linear combinations that generate the LVs and can therefore be understood as the influence of each variable on each LV [55]. The optimal number of LVs is usually selected by means of cross validation procedures, by choosing the latent variables which minimize the cross validation error in classification [56]. After the PLS regression model has been built, a linear discriminant classifier is used for classifying unknown samples (spectra). When −1 and +1 are the encoded values of class membership, if the predicted value is above 0, a corresponding object is considered a member of a class and if not it is considered a stranger [57].

CHAPTER 3

LITERATURE REVIEW

This review focuses on recent advances in metabolomics to assess milk quality, safety and authenticity.

3.1. Milk metabolomics approaches

Consumers are increasingly interested in knowing the geographical origin and perceived quality of the products they eat and drink due to the globalization of food markets and the resulting increase in variability and availability of food from other countries. From both commercial and legal perspectives, quality assurance and food authentication methods are of great interest [58].

Omics technologies have the potential to expand the scope of conventional focused analysis and open up new opportunities for unique applications. Foodomics is emerging as a useful method to assess food quality, safety and traceability in a robust, efficient, sensitive and economical manner. The use of omics technologies in food offers the opportunity to investigate previously unsolved concerns and questions in the field of food science [59]. As a consolidated branch of omics strategies, Foodomics can combine the three basic omics branches of genomics, proteomics and metabolomics to investigate the molecular underpinnings of the nutritional and functional properties of food and to understand the relationship between dietary exposure and human health [60].

Metabolomics is concerned with both targeted and non-targeted analysis of endogenous and exogenous small molecule metabolites (<1500 Da), and presents a promising tool for biomarker discovery [61].The use of metabolomics in milk and dairy science has expanded significantly in recent years thanks to key elements and biomarkers related to animal health, production, authenticity, etc. As with macro-components, milk metabolites can be influenced by a number of factors such as complex matrix effects, animal species (e. g. cow, goat and sheep), thermal treatment, cow status and farming practices.

Metabolomics is applicable to dairy breeding programs and management techniques to produce milk with desirable nutrients. During milk processing, transport and storage, metabolomics can provide reliable metabolite indicators for maintaining and improving product quality. Metabolite profiles obtained from metabolomics could also be used to determine whether dairy products are contaminated or adulterated by chemical and biological agents. [20].

Due to the complexity and abundance of the metabolites, high-resolution MS has combined with other advanced high-throughput separation techniques to perform these measurements. These methods include NMR, GC-MS, LC-MS and others, and some positive results have been obtained using these methods. While GC-MS and NMR spectroscopy are particularly well suited to the analysis of bulk and volatile organic compounds and derivatized primary metabolites, respectively, LC-MS is highly applicable to the analysis of a wide range of semi-polar compounds, including many secondary metabolites of interest. LC-MS is a widely used tool because it can avoid chemical derivatization. The high efficiency of these analytical techniques allows the simultaneous study of a wide variety of chemicals, although the most commonly mentioned metabolites include polar molecules such as organic acids, amino acids, sugars and amines, polar and non-polar lipids, terpenoids and sterols. Compared to MS-based techniques, which require a sample preparation step that may result in metabolite loss, MS-based metabolomics offers high selectivity and sensitivity for metabolite identification and quantification, and its combination with sophisticated and high-throughput separation techniques can reduce the complexity of metabolite separation. It is therefore preferable to use different techniques in tandem, such as GC-MS, LC-MS or NMR, to analyze the whole metabolome [61,62].

Metabolomics in the various studies described above showed a wide range of metabolites that can be used to distinguish between the milk of different animal species. A typical example is the studies by Jia et al. (2020) [12] and Zhang et al. (2022) [63] , in which discriminating biomarkers were identified, such as b-carotene only in cow's milk, retinol, retinaldehyde, retinoic acid abscisic acid, nonanoic acid, decanoic acid and octanoic acid only in goat's milk and finally ergocalciferol only in buffalo's milk.

3.2. Analytical approaches applied in the milk field

For food authenticity claims to be possible, it is necessary to develop analytical techniques that can be used to investigate specific food ingredients and demonstrate the authenticity of food. The characterization of the milk metabolome is a promising approach for the assessment of the overall quality and authenticity of milk. Recent advances in analytical technologies have created the conditions for metabolite profiling to help us understand complex molecular processes and physiology. There are numerous analytical platforms that have been used for metabolomic applications including:

- NMR,
- Fourier transform-infrared spectroscopy (FT-IR) and
- MS, coupled to separation techniques such as NMR, GC-MS, LC-MS, FT-MS and UPLC-MS.

NMR-based metabolomics: A commonly used technique is nuclear magnetic resonance (NMR) spectroscopy, which has several advantages. NMR can uniquely identify and simultaneously quantify a wide range of organic compounds in the micro-molar range. It has the ability to directly analyze different sample matrices with little pretreatment. In addition, this technique presents good repeatability and reproducibility of the results. Under certain circumstances, NMR-based metabolomics can provide a "holistic view" of metabolites, making it useful and helpful for metabolomic studies. Because NMR is simple, generally automated, and non-destructive, samples can be further analysed afterwards. The relatively low sensitivity and the high cost of the equipment required for NMR spectroscopy compared to other analytical techniques, such as mass spectrometry, are fundamental drawbacks for comprehensive metabolite profiling, making it unsuitable for the study of many low-abundance compounds. NMR has been used for analysis of metabolites, such as amino acids, nucleotides and nucleosides, vitamins, thiols, carbohydrates and peptides [61,64].

MS-based metabolomics: MS is becoming increasingly useful in high-throughput metabolomics, often in combination with other techniques like Gas Chromatography (GC) and High-Performance Liquid Chromatography (HPLC).

Using GC-MS, it is possible to profile hundreds of chemicals simultaneously, including organic acids, most amino acids, sugars, sugar alcohols, aromatic amines and fatty acids, in addition to the immediate separation and quantification of volatile metabolites. Although, the most effective technique for eliminating adulteration in milk is High-Resolution Mass Spectrometry coupled to Liquid Chromatography (LC-HRMS), as it provides numerous benefits throughout the entire sample analysis process:

- ✓ As an alternative to injection after extraction or microfiltration, liquid samples can be injected directly.
- ✓ During the separation phase, both reversed-phase chromatography (for the separation of non-polar and medium-polar metabolites) and HILIC chromatography (for the separation of polar metabolites) can be used. Fast analysis and high resolution are also possible at this stage.
- ✓ The technique can also provide high mass accuracy, selectivity, high sensitivity and a broad detection range at the metabolite detection stage [18].

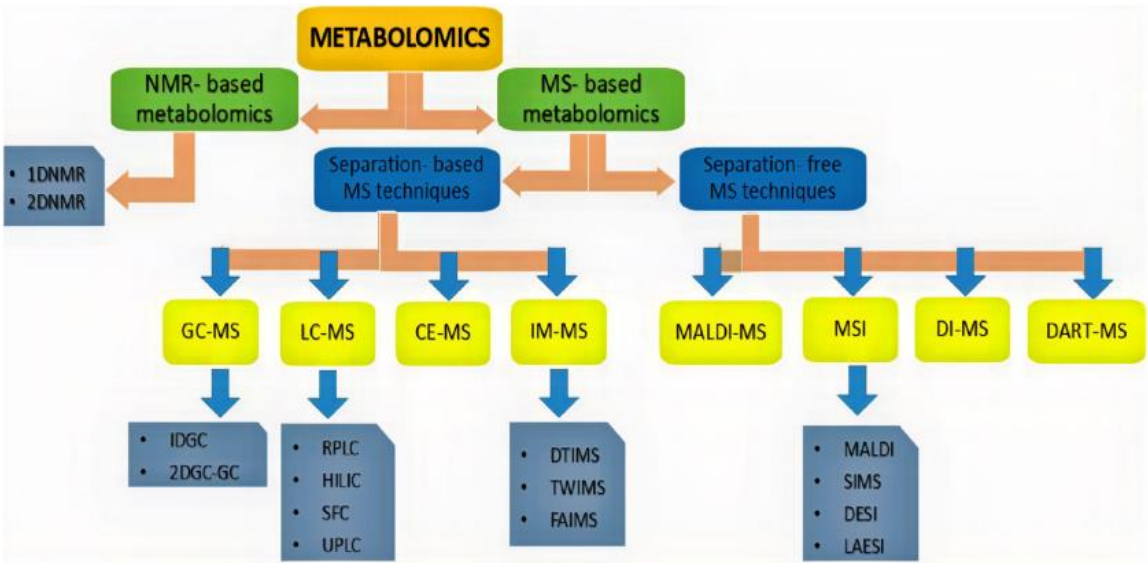


Figure 14: Analytical methodologies used to detect milk metabolites [18].

The studies reported in this literature review are described in detail below in Tables 1 and 2, where their separation is based on:

- the type of milk analysed,
- the sample preparation,
- the technique - instruments used,
- the metabolite markers used,
- the results obtained and
- the relevant literature reference.

3.3. Summary of literature on metabolomic analysis in milk authenticity

Table 1: Metabolites by animal species - Milk quality and authenticity

Scope	Factor	Sample	Metabolite	Metabolic Pathways	Instrument	Reference
To study the differences between cow's and goat's milk	Genetic traits	Cow's and goat's milk	Glucose, fructose, ribose, talose, glycine, valine	Carbohydrate and amino acid metabolism	GC–MS	**Scano et al., (2014)** [65]
To study metabolic differences between goat and sheep milk.		Goat and sheep's milk	Myo-inositol, citric acid, arabitol, a-ketoglutaric acid, glyceric acid, glycine, mannose 6-phosphate, valine, pyroglutamic acid, fucose, leucine	Protein and fat metabolism	GC–MS	**Caboni, Murgia, et al., (2019)** [66]
To study differentially regulated metabolites in milk from different species (e.g. markers to distinguish cow's milk from milk from other dairy animals).		cow's, buffalo, goat's, yak, horse and camel milk	Electric acid, choline, lactic acid, acetic acid, pyruvic acid, carnitine, uridine, pyroglutamic acid, hydroxyoctadecadienoic acid, linoleic acid, 2-(3-carboxy-3-aminopropyl)-L-histidine, capric acid, indoleacetic acid	Glycerophospholipid metabolism, valine, leucine, isoleucine and unsaturated fatty acid biosynthesis	LC-MS and NMR	**Yang et al., (2016)** [67]
To clarify the metabolic changes in milk produced under hot weather conditions.	Environment	Milk produced in hot weather (heat stressed cows)	lactic acid, pyruvic acid, citrate, galactose 1-phosphate, phosphatidylcholine, sphingomyelin, monoacylglycerol, diacylglycerol, triacylglycerol), isoleucine, glycine, proline, proline, serum acid, trimethylamine	Carbohydrate, lipid and amino acid metabolic pathways, gut microbiome metabolism	LC-MS and NMR	**Tian et al., (2016)**[68]
To identify seasonal variations.		Cow's milk from different seasons	Lipids (phosphatidylcholine , sphingomyelin, phosphatidylethanolamine)	Polar lipid pathway	LC-MS	**Liu et al., (2017)** [69]
To study different metabolic characteristics in milk from buffaloes reared at different altitudes.		Buffalo milk from different altitudes	Amino acids (pyroglutamic acid, glutamine), vitamin B6, phosphorylated sugars (Man-1P, Glc-1P, Gal-1P), fatty acids (dodecanoic acid, caprylic acid, capric acid, palmitic acid, myristic acid, etc.)	Metabolism of amino sugars, nucleotide sugars and galactose, fatty acid biosynthesis	LC-QTOF-MS	**Pu et al., (2021)** [70]
To evaluate the nutritional quality and authenticity of organic milk.	Nutrition	Buffalo milk from conventional and organic feeding	PC, mono- and polyunsaturated fatty acids	Polar lipid pathway, fatty acid biosynthesis	NMR	**Mazzei & Piccolo, (2018)** [71]
Assess the nutritional quality and authenticity of organic milk		Conventional and organic cow's milk	Unsaturated fatty acids (9-cis, 11-trans-linoleic acid, alpha-linolenic acid)	Fatty acid biosynthesis	NMR	**Tsiafoulis et al., (2019)** [72]

To investigate the differences in milk metabolites in cattle (cows) on a pasture-based feeding system	Nutrition	Pasture and non-pasture fed cows' milk	Carbohydrates, Nucleosides (1-methylene), Amino Acids (tyrosyl-threonine, l-cysteine, histidynyl cysteine, selenocysteine), vitamin derivatives	Carbohydrate, Amino Acid and Purine Metabolism	LC-MS	**Ashokan et al., 2021)** [73]
To study the metabolic differences in milk from different dietary combinations		Cow's milk from different feed combinations (maize, hay, mixture of hay and fresh feed)	Phenolic metabolites (3,3',4'5 tetrahydroxystilbene, schisandrin, etc.), lipids (PC, DG, SM, TG)	Polyphenol and lipid metabolism	LC-QTOF-MS	**Rocchetti et al., (2020)** [74]
To investigate the identity of metabolites in milk samples from maize, hay (cultivated) and pasture hay diets.		Cow's milk from different feeds	Lactic acid, hydroxycinnamic acid, glutamic acid, creatinine	Energy, lactate and lipid metabolism	DART-HRMS	**Riuzzi et al.,(2021)** [75]

Table 2: Reported analytical approaches for dairy products authenticity characterization.

References	Milk samples	Sample preparation	Analytical technique	Metabolite ("marker")	Type of study-Type of chemometrics	Results
Jia et al., (2020)[12]	Cow, buffalo, goat's milk.	A. 1g of each sample was weighed into polypropylene centrifuge tubes (50 mL). B. B. 10 mL ACN was added as extraction solvent to dissolve the samples. C. The tube was stoppered tightly and vortexed at 3,200 rpm for 1 minute. D. 1 g NaCl and 400 mg anhydrous MgSO4 were added. The tube was immediately shaken for 20 minutes using an SA31 oscillator and centrifuged for 20 minutes at 4 °C, 14,000 rpm^{-1}. E. The acetonitrile layer was evaporated to dryness under a gentle nitrogen stream in a water bath at 25 °C. F. The extracts were reconstituted with 1 mL methanol. G. Finally, 5 mL of the diluted extract was injected into the Q-Orbitrap UHPLC/ESI Q-Orbitrap system for analysis.	UHPLC-(H-ESI)/Q-Orbitrap-HRMS (Thermo Fisher 160 Scientific, Bremen, Germany) Ionisation: Positive Column: Hypersil Gold aQ C-18 (100 mm × 161 2.1 mm, 1.9 µm, Thermo Scientific, San Jose, CA, USA) Mobile phase Solvent A: 0.1% (v/v) formic acid in water Solvent B: 0.1% (v/v) formic acid in methanol	▪ **B-carotene** → cow's milk, ▪ **ergocalciferol** → buffalo milk**,** ▪ **nonanoic acid, decanoic acid, octanoic acid**→ goat's milk	Non-target screening, Principal Component Analysis (PCA) Orthogonal Projection to Latent Structure Discriminant Analysis (OPLS-DA)	The metabolites - markers with the highest discrimination were beta-carotene, which was only found in cow's milk, ergocalciferol, which was only found in buffalo milk, and nonanoic , decanoic and octanoic acids, which were higher in goat's milk than in cow's and buffalo milk.
Yang et al., (2016) [67]	Cow (Holstein and Jersey cows), goat, buffalo milk.	A. Raw milk samples were centrifuged at 3000 × g for 15 minutes at 4 °C to remove the fat layer. B. Samples of skimmed milk were collected and centrifuged at 100 000 × g for 1 hour at 4 °C in an ultracentrifuge to obtain the supernatant. The supernatant was centrifuged again to obtain the whey. C. Samples were stored at -80 °C until further analysis.	LC/Q-TOF-MS ((1290 Infinity LC, Agilent), ((6530 Accurate-Mass Q-TOF LC-MS, Agilent). Ionization: positive and negative Column: C18 reversed-phase column (100 mm × 2.1 mm, 1.8 µm; Agilent) Mobile phase Solvent A: 0.1% (v/v) formic acid in water (Milli-Q) Solvent B: 0.1% (v/v) formic acid in ACN	▪ **Choline,** ▪ **electric acid,** ▪ **lactic acid,** ▪ **acetic acid, pyruvic acid** →in higher concentrations in cow's milk (Holstein and Jersey). ▪ **Carnitine, uridine, pyroglutamic acid** → significantly lower in cow's milk (Holstein and Jersey). ▪ **Valine, capric acid** → at higher levels in goat's milk than in cow's milk	Non-targeted approach, Principal Component Analysis (PCA), Partial Least Squares Discriminant Analysis (PLS-DA) and Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA)	Different metabolites were identified and some were used to distinguish cow's milk (Holstein and Jersey breeds) from goat and buffalo milk.
Zhang et al., (2022) [63]	Goat, cow's milk.	A. Complete extraction of carotene and metabolites: by cold pressure induced acetonitrile double aqueous phase separation system (ATPS). B. 4 mL of the samples were mixed with 16 mL of 80% ACN. C. Vortexing for 5 min and freezing the mixture in a freezer at -80 °C for about 30 min (complete phase separation and minimisation of lipid extraction in milk). D. Defrost the mixture and centrifuge at 5 000 × g for 15 minutes at 4 °C. E. Phase separate the supernatant by centrifugation at 10 000 × g for 35 minutes at 4 °C. F. Add 1.0 g anhydrous MgSO4, 0.3 g PSA and 0.2 g C18 to	UHPLC-ESI/Q-Orbitrap-MS (Ultimate 3000 (Thermo Fisher Scientific) Ionisation: Positive, APCI ion source Column: Hypersil Gold aQ C18 reversed-phase capture column (100 mm × 2.1 mm, 1.9 µm, Thermo Scientific) Mobile phase Solvent A: 0.1% (v/v) formic acid in water Solvent B: 0.1% (v/v) formic acid in methanol	▪ **B-carotene** → Cow's milk, ▪ **Retinol, retinaldehyde, retinoic acid abscisic acid** → goat's milk	Non-targeted quantitative analysis, PCA, OPLS-DA	This study showed that beta-carotene acted as a discriminatory biomarker in cow's milk and retinol, retinaldehyde, retinoic acid and abscisic acid in goat's milk.

		the 5 mL supernatant. Stir the mixture. G. Centrifuge at 10 000 × g for 15 minutes at 4 °C. H. Evaporate the supernatant in 2 mL under a slow stream of nitrogen. I. Filter through a 0,22 µm membrane and analyzed by UHPLC-Q-Orbitrap system.				
Shetty et al.,(2020) [76]	Cow, goat, buffalo milk.	A. After thawing, samples were placed in an ultrasonic bath for 1 minute and 50 µL of milk was used for analysis. B. 250 µL ice-cold methanol (protein precipitation) was added to the milk samples. C. Centrifugation 5 min, 7379 x g. D. The supernatant was transferred to another tube and inoculation with ISTDs was performed (10 µL of 0.1 µg/mL except for nicotinamide which is 0.5 µg/mL). E. Dry under vacuum and redissolve in 100 µL water. F. Remove hydrophobic moieties (cholesterol) by washing with 100 µL diethyl ether. G. Stir and centrifuge for 5 minutes at 7379 × g. H. Discard the upper ether layer. I. Transfer 80 µL of the lower layer to HPLC vials. J. Inject 10 µL into the column for analysis of B vitamins in milk using a validated UHPLC-MS/SRM method.	UHPLC/QTRAP-MS (Agilent 1290 infinity II UHPLC (Agilent Technologies India Pvt. Ltd., Bangalore, India)) Ionization: Positive Column: Hypersil Gold aQ C-18 (100 mm × 161 2.1 mm, 1.9 µm, Thermo Scientific, San Jose, CA, USA) Mobile phase Solvent A: 0.3% FA in water Solvent B: acetonitrile	▪ Vitamins B	Targeted analysis with quantification.	Goat's milk had higher levels of vitamin B1 compared to cow's and buffalo's milk. Vitamin B2 was higher in buffalo milk. Buffalo milk had the lowest amount of B3. Pantothenic acid (B5) was higher in cow's milk samples. The same cow's milk samples that showed a higher amount of B1 also showed a higher amount of B6 compared to the other samples. Biotin was found at similar levels in cow's and goat's milk and at lower levels in buffalo milk. Folic acid was not found in the milk samples. B9 was higher in cow's milk. Total choline was higher in cow's milk and lower in buffalo milk. In general, B1, B3, B6 are more abundant in goat milk. B2 is more abundant in buffalo milk. B5, B9, TC (total choline) are higher in cow's milk. B8 i similar in cow's and goat's milk.
Yuan et al., (2022) [77]	Cow, buffalo milk.	A. Milk samples were thawed at room temperature and then 50 µL of each sample was mixed with 350 µL of cold extraction fluid, methanol:methyltri-butyl ether (MTBE) in a ratio of 1:1 (vol:vol), containing vitamin E acetate (25 ppm) as an internal standard. B. Vortex the mixture for 3 minutes. C. Then incubate for 10 minutes at 4 °C. D. Centrifuge at 14,000g for 15 minutes at 4°C. E. The supernatant was transferred to a 0,22 µm filter (Corning, SLC, USA) and centrifuged at 14 000 g for 5 minutes at 4 °C. F. The clear supernatant was then used for analysis.	LC-MS/MS (Dionex UltiMate 3000 Uhplc) equipped with a mass spectrometer (Q Exactive, Thermo Fisher Scientific, CA, USA). Ionization: positive and negative. Column: Hypersil GOLD HPLC (50×2.1 mm, 1.9 µm). Mobile phase Solvent A: 10 mM ammonium formate in water. Solvent B: 10 mM ammonium formate in methanol.	▪ **Choline, acetylcholine** → higher in buffalo milk than in cow's milk. ▪ **Nicotinamide** → cow's milk, ▪ **Uric acid** → cow's milk	Non-targeted metavolume analysis, PCA, OPLS-DA	Four metabolic pathways, including glycerophospholipid metabolism, nicotine and nicotinamide metabolism, glycine, serine and threonine metabolism, and purine metabolism, may influence milk yield. Four metabolites, including acetylcholine, choline, nicotinamide and uric acid, are associated with differences in milk content characteristics between Italian Mediterranean buffaloes and Chinese Holsteins.

CHAPTER 4

SCOPE

Due to the high commercial and nutritional value of milk and dairy products, it has been observed that several studies have been carried out focusing on the investigation of metabolites as authenticity markers in these food products. Food authenticity is often closely related to food safety when it comes to food adulteration. Therefore, the increasing interest in the quality and safety of milk and dairy products and the requirements for safety compliance have accelerated the need to apply analytical techniques for their metabolomic analysis. In recent years, the authenticity assessment of milk has been carried out using chromatographic separation methods (GC or LC) combined with sensitive and specific detection techniques such as MS. In particular, there has been an even greater application of metabolomics with high performance liquid chromatography (HPLC) coupled to high resolution mass spectrometry (HRMS).

In conclusion, the aim of this study was to develop a sophisticated analytical workflow for the detection and identification of metabolites in different types of milk, such as cow, goat, sheep, in order to discriminate them and thus evaluate the origin of the milk. This discrimination was achieved by performing both targeted and suspect screening approaches, using liquid chromatography combined with high resolution mass spectrometry (LC-QTOFMS). Reversed phase (RP) chromatography and two ionization modes (positive, negative) were used to cover as many compounds-metabolites as possible and to obtain the maximum analytical information. In reversed phase chromatography, the static phase is less polar (non-polar) than the mobile phase (polar) and is used to separate less polar analytes. PCA and PLS-DA chemometric tools are used.

CHAPTER 5

MATERIALS AND METHODS

5.1. Instrumentation: LC-QTOFMS system

An Ultra-High Performance Liquid Chromatography (UHPLC) system (UltiMate 3000 RSLC, Thermo Fisher Scientific, Germany) coupled to a QuadrupoleTime of Flight Mass Spectrometer (QTOF-MS) (Maxis Impact, Bruker Daltonics, Bremen, Germany) was used for the analysis of the samples.
The **UHPLC apparatus** consists of:

- ✓ a solvent rack degasser,
- ✓ a binary pump with solvent selection valve (HPG-3400),
- ✓ an auto-sampler and
- ✓ a column.

The **QTOF-MS apparatus** consists of an Electrospray Ionization (ESI) source operating in positive and negative mode and a QTOF Mass Analyzer (Ion funnel➔Quadrupole➔CID Cell➔Flight Tube➔Reflector➔Detector)

ESI: The operating parameters of the ESI interface were the following:

- ✓ capillary voltage 3000 V for negative mode,
- ✓ end plate offset 500 V,
- ✓ nebulizer pressure (N_2) 2.0 bar,
- ✓ drying gas (N_2) 8.0 L/min and
- ✓ drying temperature 200^{o}C.

Calibration: The external calibration of the QTOF mass spectrometer was performed with a sodium formate solution before the analysis, when a calibrant injection was automatically performed at the beginning of each run. A segment of 0.1- 0.25 min was used for internal calibration. The calibrant solution of sodium formate consisted of 10 mM sodium formate in a mixture of water:isopropanol 1:1. In particular in a 100 Ml volumetric flask were mixed:

- ✓ water:isopropanol 1:1 (50 ml 2-propanol, 50 ml H_2O ultrapure)
- ✓ 200 μL formic acid

- ✓ 1 ml NaOH 1M.

The flow rate used for the external calibration was initially set at 0.001 ml/min and, after some time, was adjusted to 0.0007 ml/min, which was the flow rate of the sequence.

Figure 15: UHPLC-QTOFMS, Maxis Impact, Bruker Daltonics [78]**.**

5.2. Reversed-phase Chromatography

For the reversed-phase liquid chromatography an Acclaim RSLC 120 C18 column (2.1 × 100 mm, 2.2 µm) (Dionex Bonded Silica Products, Thermo Scientific, Dreieich, Germany) was used, thermostated at 30˚C.

In the positive ESI mode, the aqueous mobile phase consisted of 90% H_2O, 10% MeOH with 5mM ammonium formate and 0.01% formic acid and the organic mobile phase of MeOH with 5mM ammonium formate and 0.01% formic acid.
In the negative ESI mode, the aqueous mobile phase consisted of 90% H2O, 10% MeOH with 5mM ammonium acetate and the organic mobile phase of MeOH with 5mM ammonium acetate.

The injection volume was set to 10 µL.

The gradient elution program for both positive and negative ESI, applied changes in mobile phase and in flow rate, is presented in the following table:

Table 3: Gradient elution program for reversed-phase chromatography.

Time (min)	Flow rate (mL/min)	Aqueous solvent %	Organic solvent %
0	0.2	99.0	1.0
1	0.2	99.0	1.0
3	0.2	61.0	39.0
14	0.4	0.1	99.9
16	0.48	0.1	99.9
16.1	0.48	99.0	1.0
19.1	0.2	99.0	1.0
20	0.2	99.0	1.0

The data were collected using the broad-band Collision Induced Dissociation (bbCID) scan mode of Data Independent Acquisition (DIA), which allows the simultaneous acquisition of accurate mass data at both low and high collision energies. There is no fragmentation at low energy in the collision cell, so the data is actually a full scan MS spectrum. At high collision energies the ions fragment and MS/MS spectra are obtained.

Information was also collected using the Data Dependent Acquisition Mode (AutoMS), which is based on the fragmentation of the five most abundant precursor ions per scan. If the m/z intensity was low for certain masses of interest, additional analysis was performed using the list of selected precursor ions in AutoMS mode.

5.3. Laboratory Equipment

In the laboratory equipment used were included:

- ✓ mobile phase solvent filtration apparatus (Millipore, XX15.04705),
- ✓ calibrated analytical balance with four decimal digits (Kern & Sohn GmbH),
- ✓ ultra-pure water apparatus 18.2MΩ / cm (Millipore Direct-Q UV),
- ✓ ultrasonic bath (Metason 60 Stuers),
- ✓ a Vortex spinner apparatus (Velp Scientifica),
- ✓ a sample shaker (Velp Scientifica) and
- ✓ a centrifugation apparatus (Remi 4500rpm, NEYA 8 basic and REMI NEYA 16R Refrigerated, 16000 Rpm).

They were also used, 5, 10, 20, 50 and 100 mL **volumetric flasks**, 100 and 250 mL **beakers**, 10, 25 and 100 mL **volumetric cylinders**, 15 mL **centrifuge tubes**, **glass tubes**, calibrated 0.1, 1 and 5 mL **pipettes**, glass and plastic **pasteur pipettes**, 1ml **syringes** and 1.5 mL autosampler **vials**.

5.4. Chemicals and Reagents

- ✓ Acetonitrile and methanol LC–MS grade were purchased from Merck (Darmstadt, Germany),
- ✓ formic acid 99% and ammonium formate from Fluka (Buchs, Switzerland).
- ✓ hexane and petroleum ether were purchased from Carlo Erba (Milan, Italy),
- ✓ distilled water was provided by a MilliQ purification apparatus (Millipore Direct-Q UV, Bedford, MA, USA) and
- ✓ syringe filters (15 mm diameter, 0.2 mm pore size) were provided from Phenomenex (Torrance, CA, USA).

The standards that were used for the development of the target databases are commercially available and can be found in the in-house laboratory database.

5.5. Preparation of Standard solutions

For quantification, several standard solutions of the targeted compounds were prepared.

Standard mixture solution of 6 compounds 100 ppm: Stock solutions Carnitine (1020ppm, H_2O), Glycine, Proline, Valine, Succinic Acid (980ppm, MeOH LC-MS), 2-cis,4-trans-Abscisic acid (1000ppm, MeOH LC-MS) were found in our laboratory. Standard mixture solution at concentration of 100 ppm of each compound were prepared by dilution of the stock solutions with methanol LC-MS in 1ml vial (100 µL of each compound were diluted in total 400 µL methanol LC-MS). The final solutions were prepared by diluting all the compounds in methanol LC-MS/ultrapure water 50:50 at final concentrations 5 and 0.5 ppm.

Standard mixture solution of 4 compounds 100 ppm: Decanoid acid, Pyruvic acid, Palmitic acid and Pyroglutamic acid (Fatty acids). About 10 mg of each individual standard was accurately weighed and placed in a 10-mL volumetric flask and were dissolved in methanol LC-MS (Stock Solutions 1000 ppm). Standard solutions at concentration of 100 ppm of each compound were prepared by dilution of the stock solutions with methanol LC-MS in 1ml vial (100 µL of each compound were diluted in total 600 µL methanol LC-MS). The final solutions were prepared by diluting all the compounds in methanol LC-MS/ultrapure water 50:50 at final concentrations 10 and 5 ppm.

Standard mixture solution of 9 compounds 100 ppm: Standard solutions of Taurine (10000 ppm), Lysine (10000 ppm), Threonine (1000 ppm), Zeaxanthine (500 ppm), Vitamin B_6 (1000 ppm) were also found in our laboratory. In a 5-mL volumetric flask were diluted in 1350 µL MeOH LC-MS:

- 500 µL Decanoid acid
- 500 µL Pyruvic acid
- 500 µL Succinic acid
- 500 µL Pyroglutamic acid
- 50 µL Taurine
- 50 µL Lysine

- 50 µL Threonine
- 1000 µL Zeaxanthine
- 500 µL Vitamin $B_{6.}$

Internal Standard (IS) mixture: Working standard solution of internal standards came by dilution of their stock solutions in MeOH LC-MS. Internal Standard (IS) mixture included:

- Syringaldehyde (for negative mode) and
- Lysine-d4 (for positive mode).

Matrix- matched standards: They were prepared with the final sample preparation protocol used for milk samples, but the addition of IS mixture and standard mixture was made at the end of the sample preparation.

Procedural blanks: Sample preparation procedure without any matrix.

Quality Control (QC) samples: They were prepared by mixing equal-volume aliquots of all milk sample extracts studied.

Stock and standard solutions of each compound were obtained and stored at -80 ºC in brown glass to prevent the photo degradation.

5.6. Samples and Storage

In our study, 35 milk samples belonging to three different sample groups according to animal species (cow, sheep and goat) were prepared and analysed. All samples were stored in the freezer at -20oC until analysis.
In the table below (Table 3), the samples that were studied are presented.

Table 4: Group and Coding of milk samples

Sample No.	Sample code	Group
1	FS 01.09	Goat
2	FS 01.10	Goat
3	FS 01.16	Goat
4	FS 01.21	Goat
5	FS 01.22	Goat
6	FS 01.28	Goat
7	FS 01.29	Goat
8	FS 01.30	Goat
9	FS 01.34	Goat
10	FS 01.35	Goat
11	FS 01.01	Sheep
12	FS 01.03	Sheep
13	FS 01.05	Sheep
14	FS 01.19	Sheep
15	FS 01.24	Sheep
16	FS 01.25	Sheep
17	FS 01.27	Sheep
18	FS 01.31	Sheep
19	FS 01.32	Sheep

20	FS 01.33	Sheep
21	1.MON 10/4	Cow
22	2.MON 10/4	Cow
23	3.MON 10/4	Cow
24	4.MON 10/4	Cow
25	5.MON 10/4	Cow
26	1.A. CARLOS 8/4	Cow
27	2.A. CARLOS 8/4	Cow
28	3.A. CARLOS 8/4	Cow
29	4.A. CARLOS 8/4	Cow
30	5.A. CARLOS 8/4	Cow
31	1.A. GIANNOUDAKIS 8/4	Cow
32	2.A. GIANNOUDAKIS 8/4	Cow
33	3.A. GIANNOUDAKIS 8/4	Cow
34	4.A. GIANNOUDAKIS 8/4	Cow
35	5.A. GIANNOUDAKIS 8/4	Cow

5.7. Data processing

Bruker software packages (Bruker Daltonics, Bremen, Germany) were used for target screening. These include TASQ Client 1.4 and Data Analysis 4.4.

TASQ Client 1.4: It was used to identify the target compounds, with the following parameters:

- ✓ mass accuracy threshold set to 5 mDa,
- ✓ satisfactory isotopic fit denoted only when mSigma (mSigma-Value is a measure of the goodness of fit between measured and theoretical isotopic pattern) was less than 100,
- ✓ signal-to-noise threshold set to 3 and
- ✓ relative tolerance of retention time window set from 0.2 to 0.4 minutes.

Furthermore, the identification of the target compounds required the use of more than two qualifier ions.

Data Analysis 4.4: It was used as a second step identification tool that provides:

- ✓ confirmation that the peaks represent the target compounds and
- ✓ the possibility to examine the MS/MS fragments of the compounds.

This identification process was applied to each compound detected in RPLC chromatography and in each ionization mode (positive, negative).

TASQ Client 1.4: It was used to identify the suspect compounds, with the following parameters:

- ✓ mass accuracy threshold set to 5 mDa,
- ✓ satisfactory isotopic fit denoted only when mSigma (mSigma-Value is a measure of the goodness of fit between measured and theoretical isotopic pattern) was less than 1000,
- ✓ signal-to-noise threshold set to 3 and
- ✓ relative tolerance of retention time window set from 0.2 to 0.4 minutes.

5.7.1. Chemometrics

The use of multivariate statistical techniques, established in analytical chemistry, has been widely used in food science and technology over the last decade. Chemometrics is typically used when the data set is large and complex in terms of number, type and reactions of samples. The data are used to verify geographical origin, farming practices and even adulteration of high value commodities [79]. An additional objective of chemometrics in foodomics studies is to find markers that address the difference between two different sets of samples. In this study, PCA was used as the chemometric tool as it is the most popular multivariate statistical technique and is widely used for handling LC-HRMS data.

5.8. Sample preparation – extraction procedures

In order to select the most appropriate sample preparation method and therefore the extraction procedure, six different sample preparation protocols with the most suitable extraction conditions of milk metabolites were collected from the literature review. LLE with ice-cold Methanol/ Defatting with Diethyl ether and Quechers method were only tested. A schematic representation of the final extraction procedure is presented in Figure 14.

5.8.1. Cold-induced acetonitrile dual aqueous-phase separation system (ATPS) - QuEChERS extraction

Each milk sample was accurately aspirated to 4 mL and then 16 mL of 80% acetonitrile was added. After five minutes of vortexing, the mixture was placed in a freezer set at -80°C for approximately 30 minutes to ensure complete phase separation and reduce co-extraction of milk lipids. The mixture was defrosted and centrifuged at 5,000 × g for 15 min at 4 ◦C. Phase separation of the supernatant was induced by centrifugation at 10,000 × g for 35 min at 4 ◦C. Then 1.0 g anhydrous magnesium sulphate ($MgSO_4$), 0.3 g PSA and 0.2 g C18 were added to 5 mL of supernatant, vortexed and centrifuged at 10,000 × g for 15 min at 4 ◦C. The supernatant was further evaporated under a slow stream of nitrogen to 2 mL, filtered through a 0.22 μm membrane and finally analysed on a UHPLC-Q-Orbitrap system. Six aliquot of the spiked cow milk sample (200

µg L-1) was added with 50%, 60%, 70%, 80%, 90% and 100% 16 mL acetonitrile/water mixture to induced phase separation by ATPS experiments. [63].

5.8.2. LLE with ice-cold Methanol/ Defatting with Diethyl ether

Samples were thawed on ice, thoroughly sonicated using a water bath sonicator for 1 minute and then 50 µL of milk was used for analysis. Milk samples were centrifuged at 7379 × g for five minutes after treatment with 250 ml ice-cold methanol to precipitate proteins. ISTDs (10 µL of 0.1 g/mL except nicotinamide which is 0.5 g/mL) were spiked after the supernatant was transferred to another tube. After vacuum drying, it was redissolved in 100 µL of water. A wash with 100 µL of diethyl ether (vortexed and centrifuged for five minutes at 7379 × g) was performed to remove hydrophobic molecules such as cholesterol. To assess the B-vitamins in the milk using the validated UHPLC-MS/SRM method, the top ether layer was discarded and 80 µL of the bottom layer was transferred to HPLC vials and 10 µL was injected on the column [76].

5.8.3. LLE with Methanol/ Methyl-tert-butyl ether (MTBE)

Milk samples were thawed at room temperature, then 50 µL of each sample was mixed with 350 µL of cooled extraction solvent, methanol: Methyl-tert-butyl ether (MTBE) in a 1:1 (vol:vol) ratio, with the addition of vitamin E acetate (25 ppm) as an internal standard. The mixture was vortexed for 3 minutes, incubated at 4 °C for 10 minutes and then centrifuged at 14,000 x g for 15 minutes at 4 °C. The supernatant was transferred to a 0.22 m filter (Corning, SLC, USA) and centrifuged at 14,000 x g for 5 minutes at 4 °C. Finally, the clean supernatant was discarded. The clean supernatant was then used for analysis [77].

5.8.4. LLE with ice-cold Methanol/Acetonitrile

After collection at 4 °C, fresh raw milk was centrifuged at 3000 x g for 15 minutes to remove the fat. Skim milk samples were centrifuged at 100,000 x g for 4 minutes at 4 °C to obtain the whey-containing supernatant, which was then

stored at 80 °C until further analysis. The extracted buffalo milk whey and the consistency of the whey samples were confirmed by SDS-PAGE gel with Coomassie blue staining. An aliquot of 100 µL serum (whey) was thawed at 4 °C and combined with 400 µL cold methanol/acetonitrile (1:1, v/v). The mixtures were kept at 20 °C for 60 min before centrifugation (Eppendorf Centrifuge 5430R, 14000 x g, 4 °C) for 15 min. The supernatants were dried by vacuum centrifugation and the dried samples were redissolved in 100 µL of acetonitrile/water (1:1, v/v) solvent mixture and centrifuged for 15 minutes (14 000 x g, 4 °C) [70].

5.8.5. LLE with Acetonitrile

1g of each sample (accurate to 0.1 mg in liquid form) was weighed into 50 mL polypropylene centrifuge tubes. Acetonitrile (10 mL) was added as an extraction solvent to help dissolve the samples. A Vortex-Genie 2T turbine mixer (Scientific Industries) was used to thoroughly mix the tube for 1 minute at 3200 rpm. 400 mg anhydrous magnesium sulphate and 1 g sodium chloride were then added. The tube was immediately shaken on a Yamata SA31 oscillator for 20 minutes and centrifuged on a Beckman Coulter Avnti J-30I PI for 20 minutes at 4 °C and 14,000 rpm. The acetonitrile layer was then gently evaporated to dryness in a water bath at 25 °C under a stream of nitrogen. The dried extracts were reconstituted in 1 mL of methanol. Finally, 5 µL of diluted extract was injected into a UHPLC/ESI Q-Orbitrap system for analysis. [12].

5.8.6. LLE with Methanol-H2O-ACN/Defatting with Hexane-Petroleum Ether

1 g of each sample was weighed into a 15 mL centrifuge tube. The target compounds of the study were spiked into the weighed samples at the appropriate concentrations. All spiked samples were allowed to rest for 10 minutes after spiking before further analysis. The samples were treated with 2 mL H2O, 2 mL MeOH and 2 mL ACN to precipitate the proteins and extract the metabolites. Each solvent was added and each tube was vortexed for 30 seconds. To increase the efficiency of the extraction, the samples were placed on an overhead shaker for 30 minutes. Ultrasonic extraction for 20 min at 40◦C

was performed for all samples followed by centrifugation for 5 min at 4000 rpm. The supernatants were transferred to new tubes and the remaining proteins and lipids were precipitated by keeping the mixture at -23°C for 12 hours (overnight). The samples were centrifuged again to remove the precipitate. The supernatant was transferred to another tube for defatting, 3 mL of hexane was added, the tube was vortexed for 1 minute and centrifuged as before. After removal of the hexane layer, 3 mL of petroleum ether was used for a second defatting step. Under the same conditions, each extract was vortexed again for 1 minute and then centrifuged. The final extracts were evaporated with a stream of nitrogen until completely dry. 40°C is the maximum temperature allowed. All samples were filtered through 0.22 mm RC filters before reconstitution in 0.5 mL methanol/water (80:20 v/v) to remove any residuals. Multi-analyte solutions were added to blank aliquots and vortexed for 10 seconds to prepare the matrix-matched standards required for analysis. The reconstituted extracts were transferred to the appropriate LC-HRMS acquisition vials and 5 µL was then injected into the instrument [80].

5.9. Method Applicability

Method validation is defined as the process of proving (through scientific studies) that an analytical method is acceptable for its intended use and is intended to demonstrate that a method is fit-for-purpose [81,82]. The validation should demonstrate the identity and concentration of the analyte, taking into account matrix effects, provide a statistical characterization of recovery results, and indicate whether the frequency of false positives and negatives is acceptable. In order to performance remains adequate over time, method validation should be continuously assessed (e.g. recovery peaks) [82].

The applicability of the selected method was tested according to the accepted criteria for the validation of analytical methods as laid down in Commission Decision 2002/657/EC for quantitative screening methods. Screening method means methods that are used to detect the presence of a substance or class of substances at the level of interest [83].

Performance parameters for analytical methods

Validation requirements depend upon the type of test method, including:

- ✓ *Specificity*: the ability to assess the exact components in a mixture.
- ✓ *Selectivity*: the ability to differentiate the components in a mixture from each other.
- ✓ *Precision*: agreement between a series of measurements.
- ✓ *Detection limit (LOD)*: lowest amount of analyte that can be detected.
- ✓ *Quantitation limit: (LOQ)*: lowest amount of analyte that can be measured [81].
- ✓ *Repeatability*: precision under repeatability conditions. Repeatability conditions means conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment.
- ✓ *Intermediate Precision*: it means that test results are obtained with the same experimental procedure, the same analyst but different days.
- ✓ *Reproducibility*: reproducibility means precision under reproducibility conditions. Reproducibility conditions means conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment, or in the same laboratory by the same operator using the same equipment but different days.
- ✓ *Linearity*: proportionality of measured value to concentration [81,82].
- ✓ *Matrix Effects*: Matrix effects derive from various physical and chemical processes and may be difficult or impossible to eliminate. They relate to the concentrations and protonation levels of co-extracted components. Co-extracted, undetected matrix components can reduce or increase the ion intensity of the analytes, depending on the analyte/matrix combination, and affect the reproducibility and accuracy of the assay. The degree of ion suppression or enhancement for an analyte and an internal standard may vary between different batches of the same biofluid [83,84].
- ✓ *Trueness and Recovery*: Trueness is the closeness of agreement between a test result and the accepted reference value of the property

being measured. Recovery refers to the proportion of analyte determined in the final result compared with the amount added (usually to a blank) sample prior to extraction, generally expressed as a percentage. Errors in measurement will lead to biased recovery figures that will deviate from the actual recovery in the final extract [82].

- ✓ *Range*: concentration interval where method is precise, accurate, and linear. The validated range is the interval of analyte concentration within which the method can be regarded as validated. The LVL is the lowest concentration assessed during validation that meets method performance criteria. It is important to realize that the validated range is not necessarily identical to the useful range of the instrumental calibration. While the calibration may cover a wide concentration range, the validated range (which is usually more important in terms of uncertainty) will typically cover a more restricted range.
- ✓ *Calibration*: Calibration is done to nullify or remove the deviation by comparing with a reference or know standard [81,82].

5.9.1. Linearity

A five-point calibration curve of standards (0.5, 1, 5, 10, 20 ppm) in pure solvent as well as in matrix-matched milk samples spiked at at five different concentrations (0.5, 1, 5, 10, 20 µg/g) were used to assess the linearity of the calibration curves. A 100 ppm mixed solution containing 9 compounds (Decanoid Acid, Pyruvic Acid, Succinic Acid, Pyroglutamic Acid, Vitamin B_6, Taurine, Lysine, Threonine, Zeaxanthine) was prepared. The milk samples were spiked with it. Syringaldehyde and Lysine-d4 were used as internal standards at a concentration of 5µg/g. The calibration curves were obtained by plotting the response of the respective analyte peak area versus its concentration. Acceptable linear regression R2 values were obtained for all compounds over the concentration ranges and calibration curves were generated by least squares linear regression analysis

5.9.2. Accuracy

The precision of a method is a part of the assessment of its accuracy.

5.9.2.1. Precision

In this study, precision was evaluated in terms of repeatability (intra-day precision) and within-laboratory reproducibility (inter-day precision). To assess repeatability, two sets (1µg/g and 10µg/g), each of them with six spiked milk samples, were prepared and analysed. The relative standard deviation (R.S.D.) was calculated. The procedure was repeated on 2 different days in order to determine inter-day precision (12 replicate analysis for reproducibility at two concentration levels).

CHAPTER 6

RESULTS AND DISCUSSION

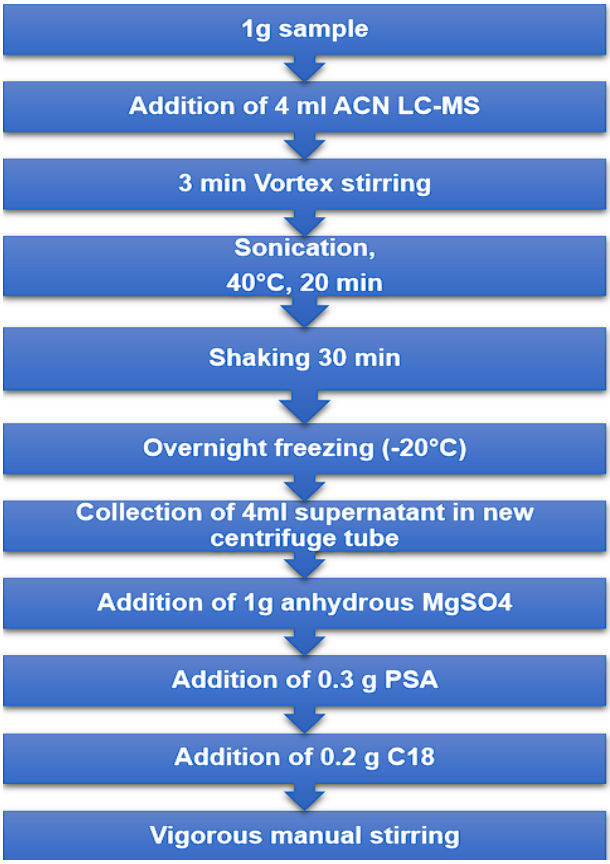
As mentioned in Chapter 4, the main goal of this study was to develop a methodology for a holistic metabolomic study and authenticity assessment of Greek cow, sheep and goat milk. The aim was to differentiate the milk of these species on the basis of its animal origin and thus to detect adulteration between them. Therefore, metabolites were investigated as 'markers' of authenticity using the LC-QTOFMS technique.

6.1. Selection of the extraction procedure

After a thorough study of the six sample preparation methods mentioned above, the two procedures that were selected are described in sections 5.7.1. and 5.7.6. with some modifications.

The following modified experimental procedures were subsequently carried out:

QuEChERS:



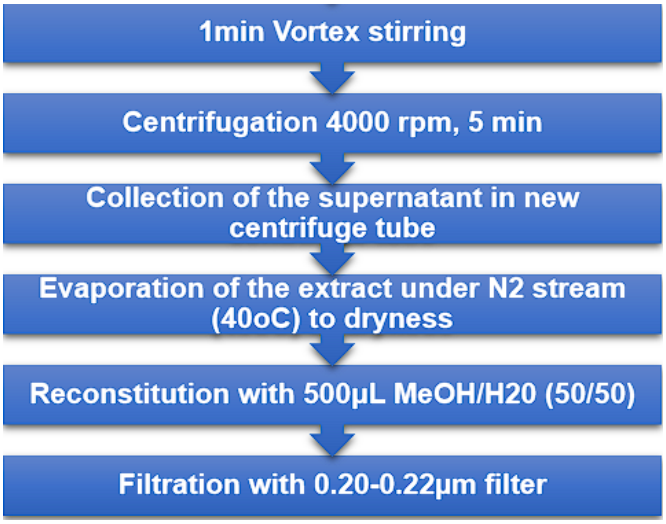


Figure 16: Schematic diagram of Quetchers sample preparation method.

LLE with Methanol-H2O-ACN/Defatting with Hexane-Petroleum Ether:

Figure 17: Schematic diagram of LLE sample preparation method.

Two different ultrasonic cleaners, 40 and 60 Hz, were used for these two sample preparation methods to investigate if the instrument frequency affects the results. Slightly better results were obtained with the 60Hz ultrasonic cleaner, so both methods were carried out with it.

Following the comparison of the two methods above, a qualitative and quantitative table and corresponding graph are presented to illustrate the superiority of results in the LLE sample preparation method.

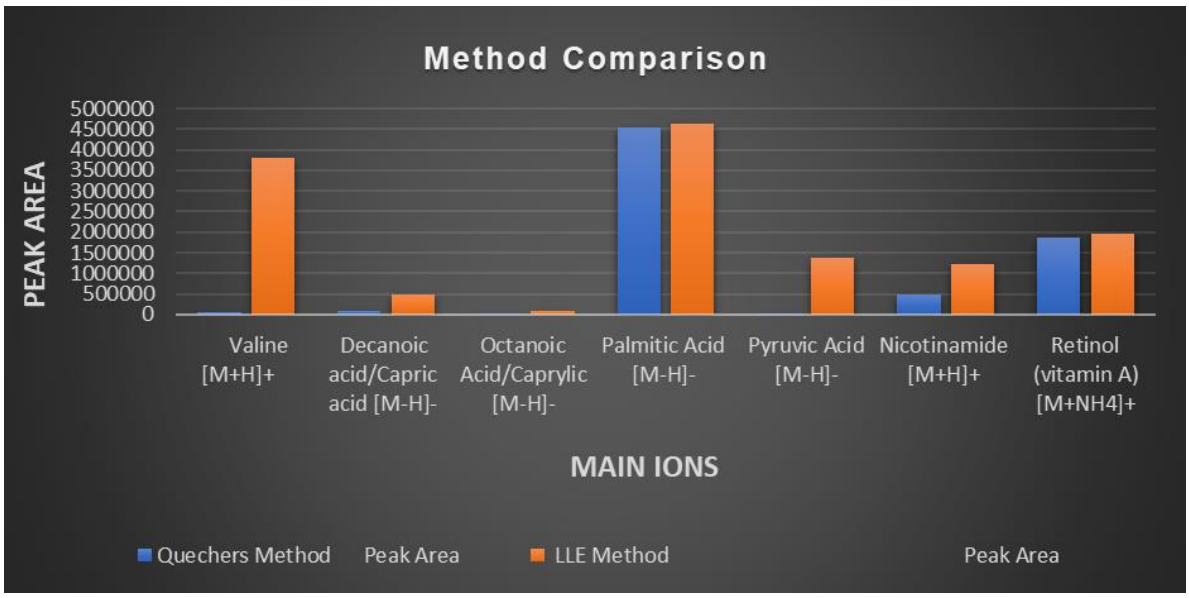
Table 5: Quality table of detected compounds

Compound	Ionization	Main Ion	Other Ions	RT
Carnitine	Positive	Exact Mass	[M+H]+	1,7 min
Proline	Positive Negative	[M+NH4]+	[M+H]+, [M+Na]+ [M-H]-, [M+HCOOH-H]-	6,5 min
Valine	Positive Negative	[M+H]+	[M+Na]+, [M-H]-	1,6 min
Glycine	Positive	[M+H]+	[M+NH4]+, [M+Na]+	1,7 min
Succinic acid	Positive Negative	[M-H]-	[M+NH4]+	1,3 min
Decanoic acid/Capric acid	Positive Negative	[M-H]-	[M+H]+	10,2 min
Octanoic Acid/Caprylic	Positive Negative	[M-H]-	[M+H]+, [M+Na]+	8,0 min
Palmitic Acid	Positive Negative	[M-H]-	[M+NH4]+	14,1 min
Pyruvic Acid	Positive Negative	[M-H]-	[M+Na]+ [M+HCOOH-H]-	1,3 min
L-pyroglutamic acid	Positive	[M+NH4]+	-	2,3 min
b-carotene	-	-	-	-
Nicotinamide	Positive	[M+H]+	[M+Na]+ [M+NH4]+	2,8 min
threonine	Negative	[M-H]-	[M+HCOOH-H]-	1,4 min
2 -cis, 4-trans- abscisic acid	-	-	-	-
Taurine	Positive Negative	[M-H]-	[M+H]+ [M+Na]+	1,4 min
Retinol (vitamin A)	Positive	[M+NH4]+	-	14,4 min

Table 6: Quantitive table showing the compounds detected with a significant difference in Peak Areas between the two methods.

Main Ions	Quechers Method Peak Area	LLE Method Peak Area
Valine [M+H]+	63140	3806896
Decanoic acid/Capric acid [M-H]-	75190	495645
Octanoic Acid/Caprylic [M-H]-	24022	100757
Palmitic Acid [M-H]-	4535846	4628363
Pyruvic Acid [M-H]-	13102	1385241
Nicotinamide [M+H]+	473382	1235546
Retinol (vitamin A) [M+NH4]+	1865881	1956050

Figure 18: Method Comparison Chart.



As can be seen, LLE method detected more compounds with higher intensities and peak areas.

6.2. Final extraction procedure

According to the above comparison of two methods, the LLE with Methanol-H2O-ACN/Defatting with Hexane-Petroleum Ether sample preparation method described in section 5.8 was selected. The preconcentration factor is important to achieve high recovery sensitivity at lower concentrations with as little matrix effect as possible. In the selected method a two times preconcentration was reached.

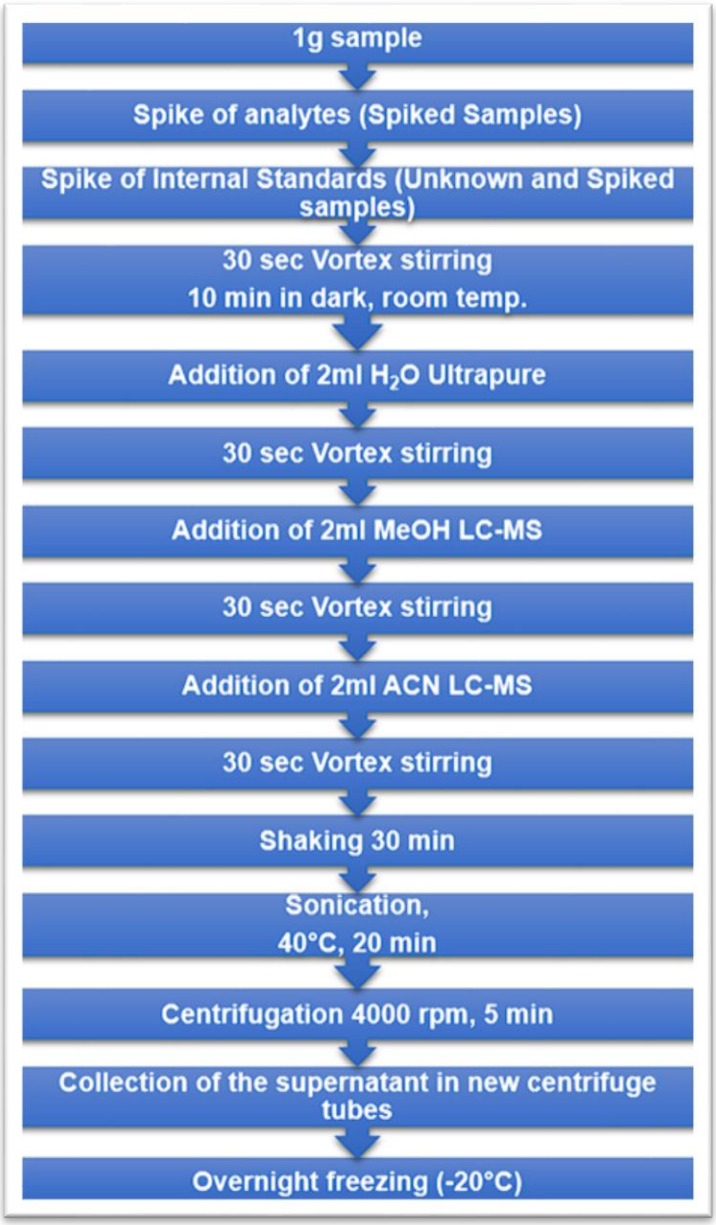


Figure 19: Schematic diagram of the final sample preparation method.

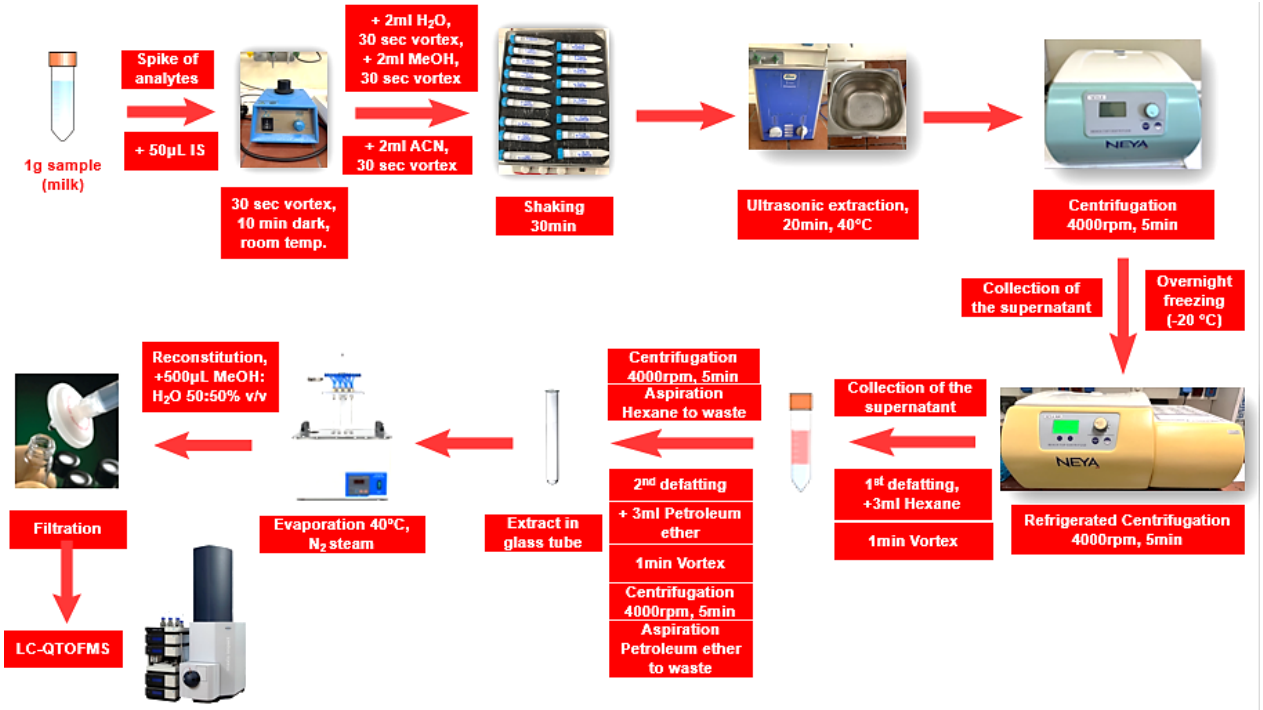


Figure 20: Schematic workflow of the sample preparation procedure.

6.3. Method Applicabilty

6.3.1. Linearity

A five-point calibration curve of standards (0.5, 1, 5, 10, 20 ppm) in pure solvent as well as in blank samples of milk at five different concentrations (0.5, 1, 5, 10, 20 µg/g) were used to assess the linearity of the calibration curves. Linear regression analysis was performed by plotting the peak area versus analyte concentrations.

For the compounds present at high levels in the milk samples, linearity is shown by a standard calibration curve, as the samples were spiked at much lower concentrations and therefore not successfully spiked. These compounds are shown below:

RP POSITIVE

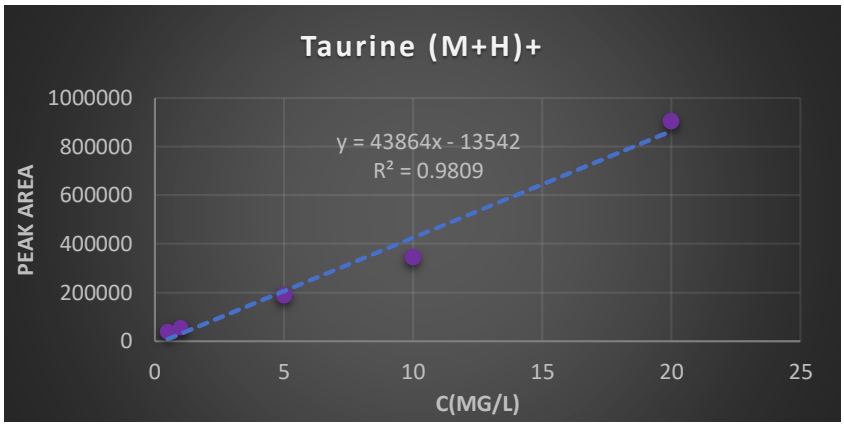


Figure 21: Calibration curve of Taurine $[M+H]^+$ standard solutions.

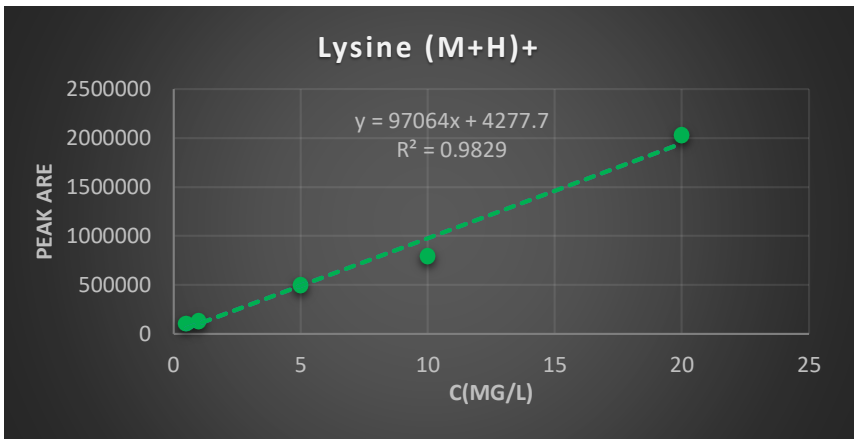


Figure 22: Calibration curve of Lysine $[M+H]^+$ standard solutions.

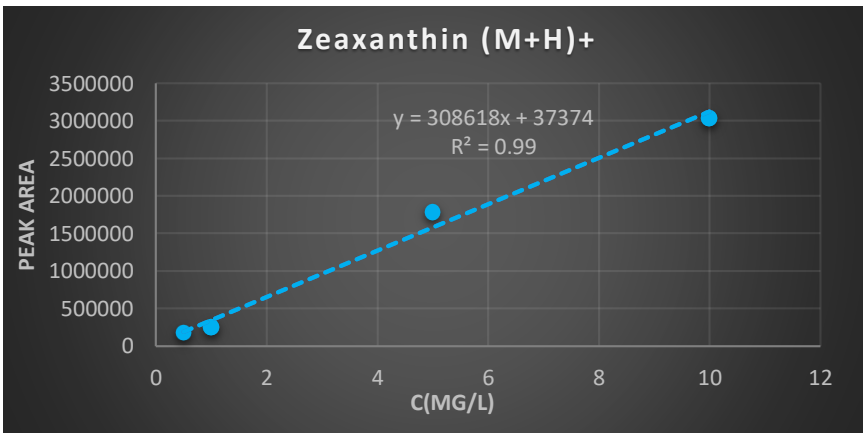


Figure 23: Calibration curve of Zeaxanthin $[M+H]^+$ standard solutions.

RP NEGATIVE

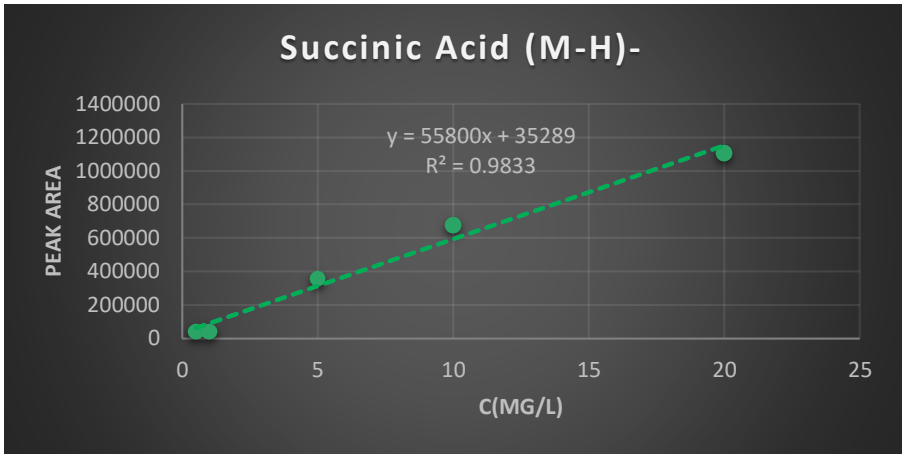


Figure 24: Calibration curve of Succinic Acid $[M-H]^-$ standard solutions.

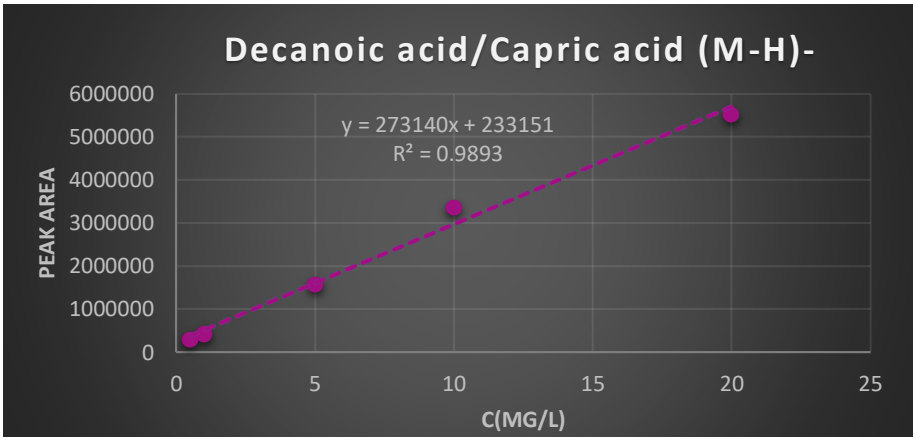


Figure 25: Calibration curve of Decanoid Acid $[M-H]^-$ standard solutions.

For the remaining compounds, linearity is illustrated by the spiked calibration curves after subtracting the peak area of the blank sample from the peak area of the analyte, as follows:

RP POSITIVE

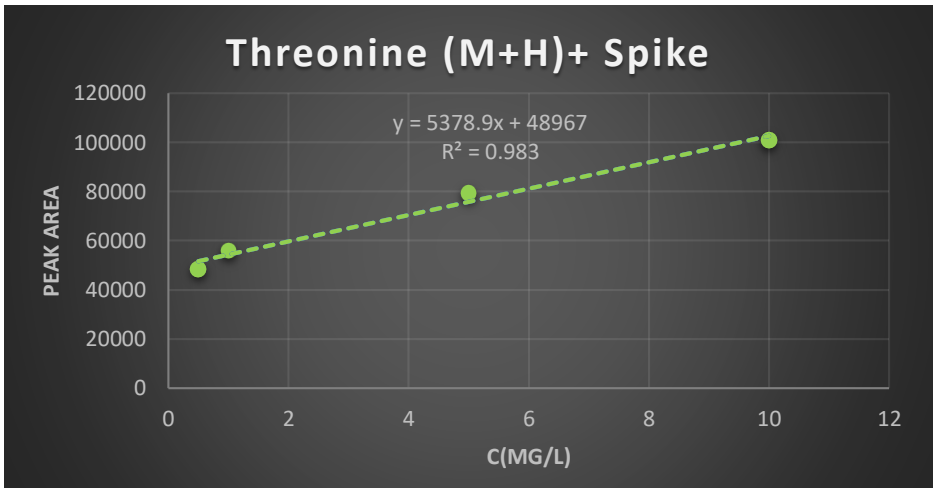


Figure 26: Calibration curve of Threonine $[M+H]^+$ in spiked milk sample.

RP NEGATIVE

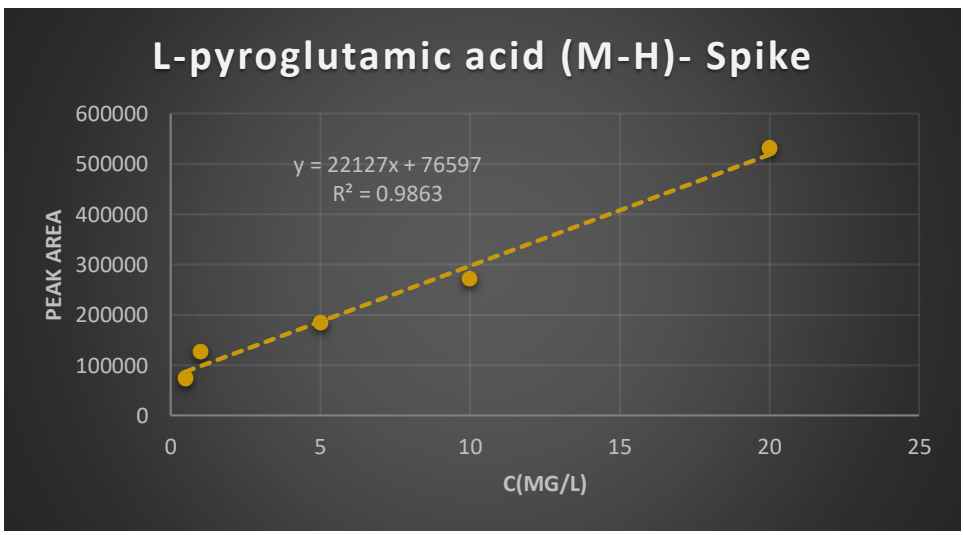


Figure 27: Calibration curve of L-Pyroglutamic acid $(M-H)^-$ in spiked milk sample.

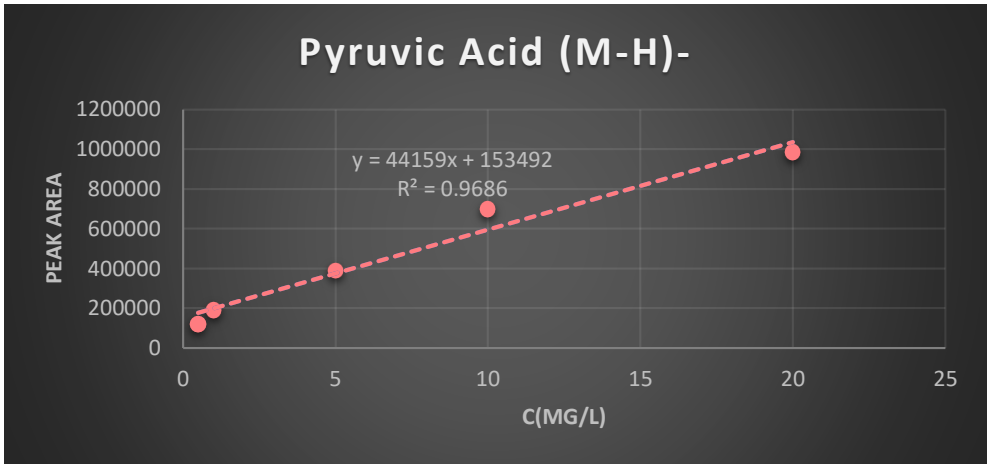


Figure 28: Calibration curve of Pyruvic Acid $(M-H)^-$ in spiked milk sample.

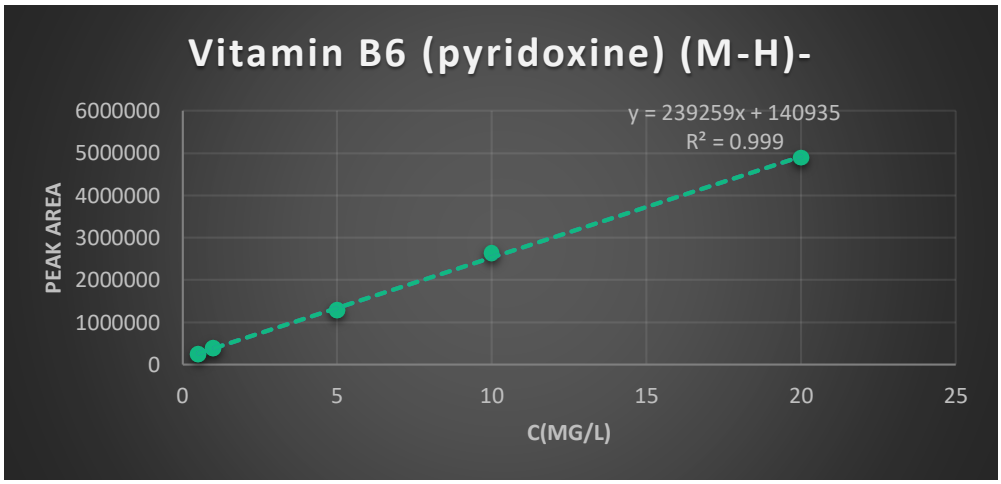


Figure 29: Calibration curve of Vitamin B6 (pyridoxine) $(M-H)^-$ in spiked milk sample.

As can be seen from the graphs above, R2 ranged from 0.9686 to 0.9990 for both the standards and the spiked samples, indicating good linearity.

6.3.2. Precision

The precision of this method was calculated as intra-day precision (repeatability) and inter-day precision (within-laboratory reproducibility). It can be observed that relative standard deviations were always lower than 20 for all analytes. These results indicate the good precision of the used method. Precision results for all compounds at concentrations of 1ppm and 10ppm are presented in Tables 7 and 8.

RP POSITIVE

PEAK AREA			
1ppm	**Threonine**	**Taurine**	**Lysine**
spike 1	-	301725	-
spike 2	9.98	315987	120.5252512
spike 3	-	345570	-
spike 4	10.83	509193	144.6433872
spike 5	10.10	354679	161.1418976
spike 6	-	336775	-
Average	10.30	360654.8333	142.103512
SD	0.46	75319.36905	20.42709511
%RSD	4.50	20.88405924	14.37479962
10ppm	**Threonine**	**Taurine**	**Lysine**
spike 1	7.707920576	392904	2374929
spike 2	9.086870681	386788	2282830
spike 3	10.95768498	289418	1677235
spike 4	-	254241	1977539
spike 5	9.892609065	377348	2577908
spike 6	9.153624181	329459	1937511
Average	9.359741896	338359.6667	2137992
SD	1.191933975	57255.45666	331324.0362
%RSD	12.7346885	16.92147803	15.49697268

Table 7: Intermediate Precision for RP Positive analytes.

RP NEGATIVE

PEAK AREA					
1ppm	**L-pyroglutamic acid**	**Pyruvic Acid**	**Vitamin B6 (pyridoxine)**	**Decanoic acid**	**Succinic acid**
spike 1	460846	388846	0.427779875	66.87713835	1544462
spike 2	440013	368393	0.385923916	54.90769637	1777589
spike 3	372582	319963	0.464008834	59.59941014	1937324
spike 4	251657	323569	0.466199368	55.04319467	1655078
spike 5	378525	369079	0.354614204	55.27271397	2060854
spike 6	375113	376374	0.420604176	57.51910806	2593200
Average	379789.3333	357704	0.419855062	58.20321026	1928084.5
SD	73062.13844	28815.34215	0.043745643	4.627922152	375350.3441
%RSD	19.23754356	8.055638782	10.41922484	7.951317688	19.46752562
10ppm	**L-pyroglutamic acid**	**Pyruvic Acid**	**Vitamin B6 (pyridoxine)**	**Decanoic acid**	**Succinic acid**
spike 1	493357	787120	2.80115724	35352808	3386605
spike 2	526028	784241	3.87204863	27103554	3476241
spike 3	404033	666880	4.745660566	40699860	2768272
spike 4	382833	514895	4.656041088	26647921	2130662
spike 5	615387	890563	3.512412711	32725892	3435257
spike 6	460589	696453	3.444435852	25212836	2407111
Average	480371.1667	723358.6667	3.838626014	31290478.5	2934024.667
SD	85105.81333	128918.8196	0.752348139	6051789.956	583192.0601
%RSD	17.71667811	17.8222541	19.5994123	19.34067565	19.87686289

Table 8: Intermediate Precision for RP Negative analytes.

Zeaxanthin was not detected in any sample. For some compounds an IS correction was applied by the ratio of the analyte peak area to the IS peak area (Lysine-d4 for positive and Syringaldehyde for RP negative mode).

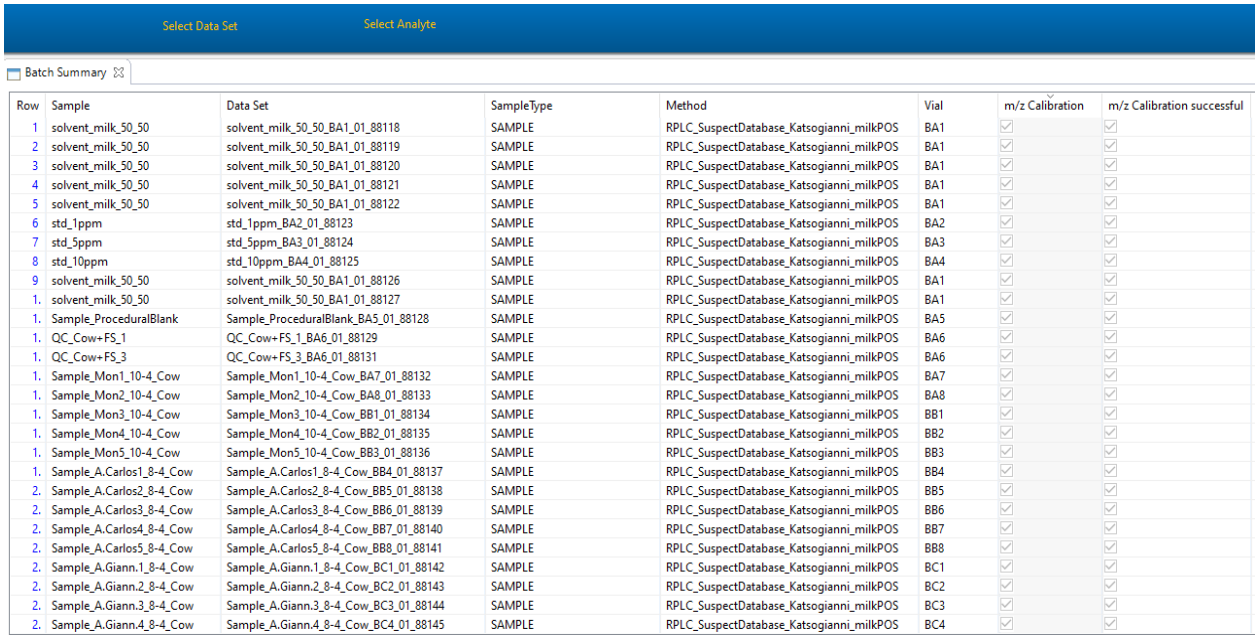
6.4. Target screening using RPLC chromatography

For metabolite target screening, 2 different lists of compounds were used, 1 for positive and 1 for negative ionization mode, as shown in the Appendix. The RPLC-positive list contained 171 metabolites, and the RPLC-negative list also contained 171 metabolites (Table 11).

The substances detected in the RPLC chromatography and in each ionization were then identified using the data processing described in section 5.7. First, the results of the LC-QTOFMS analysis were entered into TASQ Client 1.4 and the appropriate processing method was selected. Subsequently, TASQ Client calibrated all sample data. Next, screening was performed for each target compound in each sample to obtain the correct information from the chromatographic peaks. Finally, Data Analysis was used as a second stage identification method to ensure that the peaks represented the target compounds, as it allows the MS/MS fragments of the compounds to be examined.

TASQ Client 1.4 was used for the identification of the target compounds, using as parameters: mass accuracy, retention time, diagnostic ions detection and observed isotopic patterns, as it was mentioned in paragraph 5.7.

The schematic diagram of this data processing is shown in the following figures (30-31).



Row	Sample	Data Set	SampleType	Method	Vial	m/z Calibration	m/z Calibration successful
1	solvent_milk_50_50	solvent_milk_50_50_BA1_01_88118	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA1	☑	☑
2	solvent_milk_50_50	solvent_milk_50_50_BA1_01_88119	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA1	☑	☑
3	solvent_milk_50_50	solvent_milk_50_50_BA1_01_88120	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA1	☑	☑
4	solvent_milk_50_50	solvent_milk_50_50_BA1_01_88121	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA1	☑	☑
5	solvent_milk_50_50	solvent_milk_50_50_BA1_01_88122	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA1	☑	☑
6	std_1ppm	std_1ppm_BA2_01_88123	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA2	☑	☑
7	std_5ppm	std_5ppm_BA3_01_88124	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA3	☑	☑
8	std_10ppm	std_10ppm_BA4_01_88125	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA4	☑	☑
9	solvent_milk_50_50	solvent_milk_50_50_BA1_01_88126	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA1	☑	☑
1.	solvent_milk_50_50	solvent_milk_50_50_BA1_01_88127	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA1	☑	☑
1.	Sample_ProceduralBlank	Sample_ProceduralBlank_BA5_01_88128	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA5	☑	☑
1.	QC_Cow+FS_1	QC_Cow+FS_1_BA6_01_88129	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA6	☑	☑
1.	QC_Cow+FS_3	QC_Cow+FS_3_BA6_01_88131	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA6	☑	☑
1.	Sample_Mon1_10-4_Cow	Sample_Mon1_10-4_Cow_BA7_01_88132	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA7	☑	☑
1.	Sample_Mon2_10-4_Cow	Sample_Mon2_10-4_Cow_BA8_01_88133	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BA8	☑	☑
1.	Sample_Mon3_10-4_Cow	Sample_Mon3_10-4_Cow_BB1_01_88134	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BB1	☑	☑
1.	Sample_Mon4_10-4_Cow	Sample_Mon4_10-4_Cow_BB2_01_88135	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BB2	☑	☑
1.	Sample_Mon5_10-4_Cow	Sample_Mon5_10-4_Cow_BB3_01_88136	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BB3	☑	☑
1.	Sample_A.Carlos1_8-4_Cow	Sample_A.Carlos1_8-4_Cow_BB4_01_88137	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BB4	☑	☑
2.	Sample_A.Carlos2_8-4_Cow	Sample_A.Carlos2_8-4_Cow_BB5_01_88138	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BB5	☑	☑
2.	Sample_A.Carlos3_8-4_Cow	Sample_A.Carlos3_8-4_Cow_BB6_01_88139	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BB6	☑	☑
2.	Sample_A.Carlos4_8-4_Cow	Sample_A.Carlos4_8-4_Cow_BB7_01_88140	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BB7	☑	☑
2.	Sample_A.Carlos5_8-4_Cow	Sample_A.Carlos5_8-4_Cow_BB8_01_88141	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BB8	☑	☑
2.	Sample_A.Giann.1_8-4_Cow	Sample_A.Giann.1_8-4_Cow_BC1_01_88142	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BC1	☑	☑
2.	Sample_A.Giann.2_8-4_Cow	Sample_A.Giann.2_8-4_Cow_BC2_01_88143	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BC2	☑	☑
2.	Sample_A.Giann.3_8-4_Cow	Sample_A.Giann.3_8-4_Cow_BC3_01_88144	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BC3	☑	☑
2.	Sample_A.Giann.4_8-4_Cow	Sample_A.Giann.4_8-4_Cow_BC4_01_88145	SAMPLE	RPLC_SuspectDatabase_Katsogianni_milkPOS	BC4	☑	☑

Figure 30: Calibration of all sample data, performed by the TASQ Client.

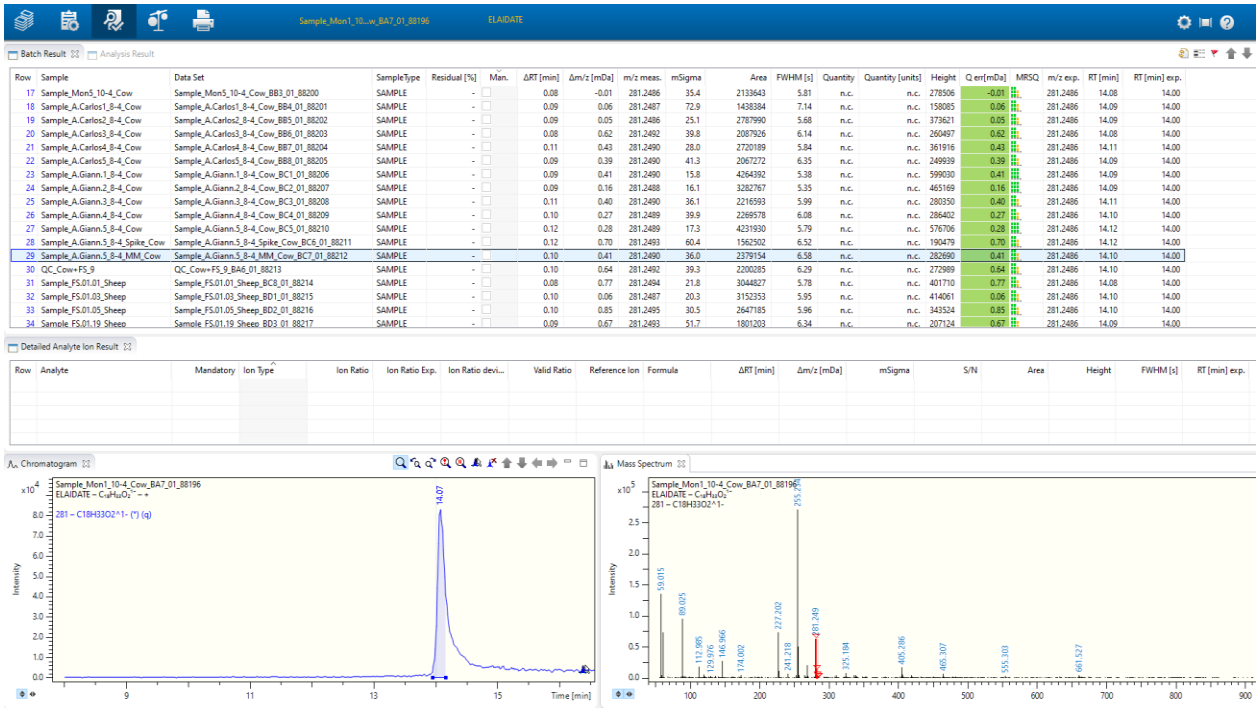


Figure 31: TASQ Client performing the screening for each compound in each sample.

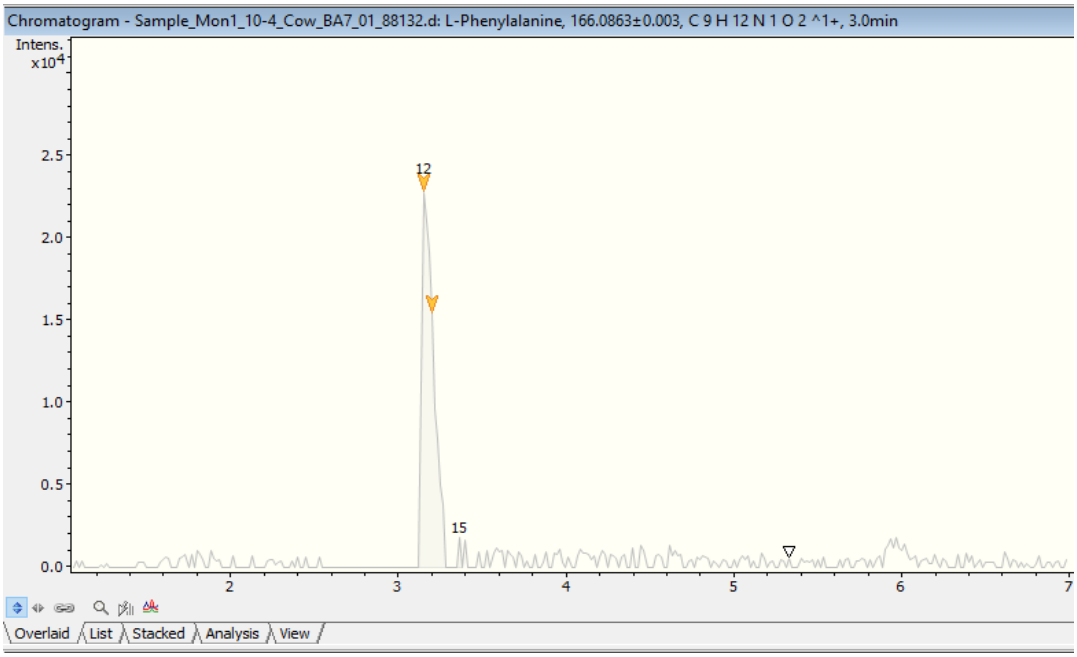


Figure 32: Chromatogram of L-Phenylalanine in Data Analysis 4.4.

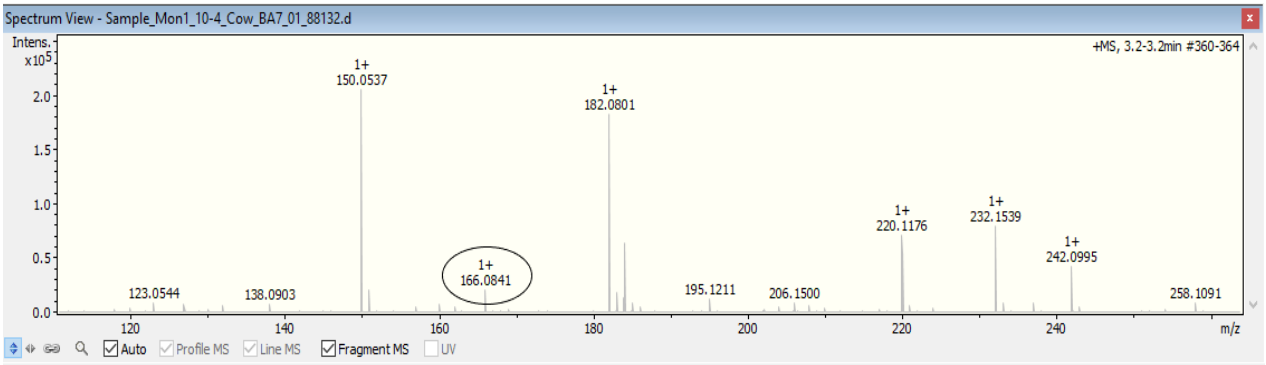


Figure 33: Mass Spectrum of L-Phenylalanine.

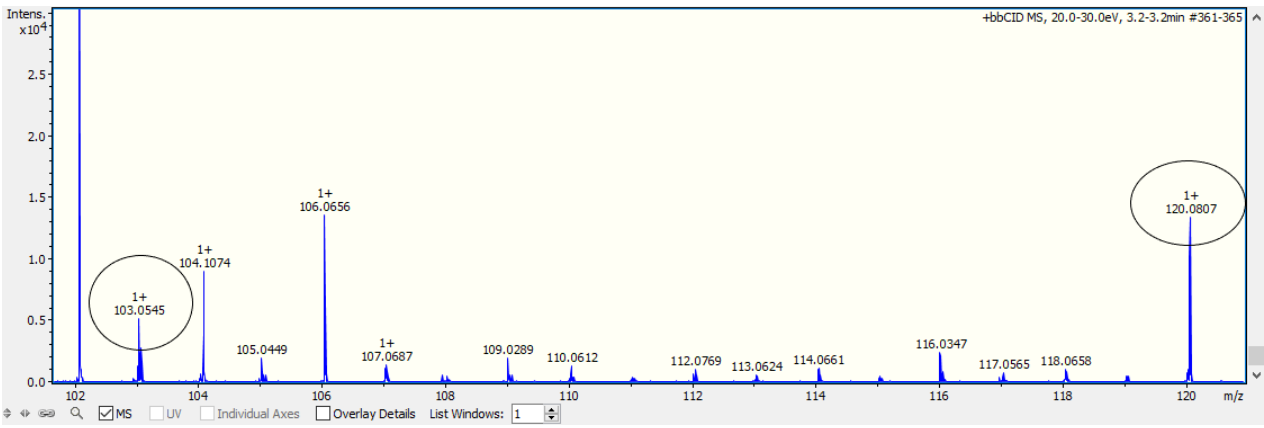


Figure 34: MS/MS spectrum of L-Phenylalanine.

In the Target screening, 52 metabolites were identified, of which 21 metabolites were identified in RPLC negative and 31 metabolites were identified in RPLC positive ionization.

Table 12 (Appendix) shows how the identified compounds were classified based on their detection in RPLC chromatography and positive or negative ionization. It can be observed that a greater number of compounds were detected and identified in positive ionization. In the Target screening, 21 compounds were identified in the negative ionization and 31 compounds in the positive ionization. To confirm the presence or absence of a compound in milk samples, we set the criterion that it should be detected in 90% of the samples. This means that 14 out of 15 cow's milk samples should contain the compound, and similarly 9 out of 10 sheep and goat milk samples.

The following tables present the compounds identified in cow's milk (15 samples), sheep's milk (10 samples) and goat's milk (10 samples):

Table 9: Identified metabolites in cow's milk.

Compound Name	Ionization
10-Hydroxydecanoic acid	Negative
2-Hydroxybutyric acid sodium salt	Negative
2-Methylglutaric acid	Negative
3-Methylglutaric acid	Negative
4-Hydroxybenzoic acid	Negative
4-Imidazoleacrylic acid	Negative
Azelaic acid	Negative
D-(+)-Galactose	Negative
Dulcitol	Negative
Elaidate	Negative
L-Norvaline	Negative
L-Pyroglutamic acid	Negative
meso-Tartaric acid monohydrate	Negative
myo-Inositol	Negative
Myristate	Negative
Nonanoate	Negative
Oleate	Negative
Petroselinate	Negative
Succinic acid	Negative

Compound Name	Ionization
Sucrose	Negative
Tagatose	Negative
4-Hydroxybenzoic acid	Negative
1-Aminocyclopropanecarboxylic acid	Positive
4-Imidazoleacrylic acid	Positive
5-Aminovaleric acid	Positive
Betaine	Positive
Dihydrouracil	Positive
Creatine	Positive
L-Glutamine	Positive
D-Mannosamine hydrochloride	Positive
DL-Normetanephrine hydrochloride	Positive
L-Isoleucine	Positive
Dopamine hydrochloride	Positive
L-Ornithine monohydrochloride	Positive
L-Phenylalanine	Positive
L-Valine	Positive
Histamine dihydrochloride	Positive
Methyl 4-aminobutyrate hydrochloride	Positive
Nα-Acetyl-L-lysine	Positive
Sarcosine	Positive
Trigonelline hydrochloride	Positive
L-Alanine	Positive
L-Arginine monohydrochloride	Positive
L-Glutamic acid	Positive
L-Homoserine	Positive
L-Methionine	Positive
L-Serine	Positive
L-Threonine	Positive
N-Acetylglycine	Positive
Xanthine	Positive
Taurine	Positive
N-Methyl-L-glutamic acid	Positive

Table 10: Identified metabolites in sheep's milk.

Compound Name	Ionization
10-Hydroxydecanoic acid	Negative
2-Hydroxybutyric acid sodium salt	Negative
2-Methylglutaric acid	Negative
3-Methylglutaric acid	Negative
4-Hydroxybenzoic acid	Negative
4-Imidazoleacrylic acid	Negative
Azelaic acid	Negative
D-(+)-Galactose	Negative
Dulcitol	Negative
Elaidate	Negative
L-Norvaline	Negative
L-Pyroglutamic acid	Negative
meso-Tartaric acid monohydrate	Negative
myo-Inositol	Negative
Myristate	Negative
Nonanoate	Negative
Oleate	Negative
Petroselinate	Negative
Succinic acid	Negative
Sucrose	Negative
Tagatose	Negative
4-Hydroxybenzoic acid	Negative
L-Isoleucine	Positive
L-Phenylalanine	Positive

Table 11: Identified metabolites in goat's milk

Compound Name	Ionization
10-Hydroxydecanoic acid	Negative
2-Hydroxybutyric acid sodium salt	Negative
2-Methylglutaric acid	Negative
3-Methylglutaric acid	Negative
4-Hydroxybenzoic acid	Negative
4-Imidazoleacrylic acid	Negative
Azelaic acid	Negative
D-(+)-Galactose	Negative
Dulcitol	Negative
Elaidate	Negative
L-Norvaline	Negative

Compound Name	Ionization
L-Pyroglutamic acid	Negative
meso-Tartaric acid monohydrate	Negative
myo-Inositol	Negative
Myristate	Negative
Nonanoate	Negative
Oleate	Negative
Petroselinate	Negative
Succinic acid	Negative
Sucrose	Negative
Tagatose	Negative
4-Hydroxybenzoic acid	Negative
1-Aminocyclopropanecarboxylic acid	Positive
4-Imidazoleacrylic acid	Positive
5-Aminovaleric acid	Positive
Betaine	Positive
Dihydrouracil	Positive
Creatine	Positive
L-Glutamine	Positive
D-Mannosamine hydrochloride	Positive
DL-Normetanephrine hydrochloride	Positive
L-Isoleucine	Positive
Dopamine hydrochloride	Positive
L-Ornithine monohydrochloride	Positive
L-Phenylalanine	Positive
L-Valine	Positive
Histamine dihydrochloride	Positive
Methyl 4-aminobutyrate hydrochloride	Positive
Nα-Acetyl-L-lysine	Positive
Sarcosine	Positive
Trigonelline hydrochloride	Positive
L-Alanine	Positive
L-Arginine monohydrochloride	Positive
L-Carnitine hydrochloride	Positive
L-Homoserine	Positive
L-Methionine	Positive
L-Serine	Positive
L-Threonine	Positive
N-Acetylglycine	Positive
Xanthine	Positive
Taurine	Positive
N-Methyl-L-glutamic acid	Positive

From the above tables, according to the criterion for confirming the presence of a compound in milk, the compound L-glutamic acid was identified only in cow's milk, whereas the compound L-carnitine hydrochloride was identified only in goat's milk.

Table 12 summarizes the identified metabolites, the categorization, the ionization, the Precursor Ion, the Adduct of the precursor ion and the fragments of the identified metabolites in Target screening.

Table 12: Ionization (positive, negative), precursor ion, adduct of the precursor ion, categorization and fragments of the identified compounds in the Target Screening.

Compound name	Category	Ionization	Precursor Ion	Adduct of precursor ion	Fragments
10-Hydroxydecanoic acid	omega-hydroxy fatty acid	Negative	187.1340	[M-H]-	-
2-Hydroxybutyric acid sodium salt	Carboxylic Acid	Negative	103.0401	[M-H]-	-
2-Methylglutaric acid	Carboxylic Acid	Negative	145.0506	[M-H]-	-
4-Hydroxybenzoic acid	Benzoic acid, paraben	Negative	137.0244	[M-H]-	93.035
4-Imidazoleacrylic acid	Carboxylic Acid	Negative	137.0357	[M-H]-	93.046
Azelaic acid	Carboxylic Acid	Negative	187.0976	[M-H]-	-
D-(+)-Galactose	Carbohydrate	Negative	179.0561	[M-H]-	-
Dulcitol	Carbohydrate	Negative	181.0718	[M-H]-	-
Elaidate	Fatty Acid	Negative	281.2486	[M-H]-	-
L-Norvaline	Amino Acid	Negative	116.0717	[M-H]-	-
L-Pyroglutamic acid	Carboxylic Acid, amino acid derivative	Negative	128.0353	[M-H]-	-
meso-Tartaric acid monohydrate		Negative	149.0092	[M-H]-	-
myo-Inositol	Carbohydrate	Negative	179.0561	[M-H]-	-
Myristate	Fatty Acid anion	Negative	227.2017	[M-H]-	-

Compound name	Category	Ionization	Precursor Ion	Adduct of precursor ion	Fragments
Nonanoate	Fatty acid anion	Negative	157.1234	[M-H]-	-
Oleate	Fatty acid anion	Negative	281.2486	[M-H]-	
Petroselinate	fatty acid anion	Negative	281.2486	[M-H]-	-
Succinic acid	Carboxylic Acid	Negative	117.0193	[M-H]-	73.030
Sucrose	Carbohydrate	Negative	341.1089	[M-H]-	-
Tagatose	carbohydrate	Negative	179.0561	[M-H]-	-
1-Aminocyclopropanecarboxylic acid	non-proteinogenic alpha-amino acid	Positive	102.0550	[M+H]+	-
4-Imidazoleacrylic acid	carboxylic acid	Positive	139.0502	$[M+H]^+$	93.045/121.040
5-Aminovaleric acid	Amino Acid	Positive	118.0863	[M-H]-	55.054
Betaine	Amine	Positive	118.0863	$[M+H]^+$	58.065
Dihydrouracil	Pyrimidine	Positive	115.0502	$[M+H]^+$	-
Creatine	Amino acid derivative	Positive	132.0768	$[M+H]^+$, $[M+K]^+$, $[M+Na]^+$	170.033/ 154.059/ 90.055/44.049
L-Glutamine	Amino Acid	Positive	147.0764	$[M+H]^+$	84.044/ 56.049
D-Mannosamine hydrochloride	Carbohydrate	Positive	180.0866	$[M+H]^+$	-
DL-Normetanephrine hydrochloride	O-methylated metabolite of norepinephrine (Biogenic Amine)	Positive	166.0863	$[M+H]^+$	134.060/ 121.065

Compound name	Category	Ionization	Precursor Ion	Adduct of precursor ion	Fragments
L-Isoleucine	Amino Acid	Positive	132.1019	$[M+H]^+$	86.096/ 69.070
Dopamine hydrochloride	Biogenic Amine	Positive	154.0863	$[M+H]^+$	137.060
L-Ornithine monohydrochloride	Intermediate, ammonia metabolism	Positive	133.0972	$[M+H]^+$	70.065
L-Phenylalanine	Amino Acid	Positive	166.0863	[M-H]-	120.081/ 103.054
L-Valine	Amino Acid	Positive	72.0808	$[M+H]^+$	55.054/ 53.039
Histamine dihydrochloride	Biogenic Amine	Positive	112.0869	[M-H]-	68.049/ 95.060
Methyl 4-aminobutyrate hydrochloride	Carboxylic Acid	Positive	118.0863	$[M+H]^+$	59.049
Nα-Acetyl-L-lysine	Peptide	Positive	189.1234	$[M+H]^+$	84.081/130.086/129.102
Sarcosine	Peptide	Positive	90.0550	$[M+H]^+$	-
Trigonelline hydrochloride	Alkaloid	Positive	138.0550	$[M+H]^+$	94.065/ 65.039/ 110.061
L-Alanine	Amino Acid	Positive	90.0550	[M+H]+	-
L-Arginine monohydrochloride	hydrochloride salt of arginine	Positive	175.1190	[M+H]+	70.065/ 157.085

Compound name	Category	Ionization	Precursor Ion	Adduct of precursor ion	Fragments
L-Carnitine hydrochloride	Vitamin B_T	Positive	162.1125	$[M+H]^+$	60.081/85.028/103.039
L-Glutamic acid	Amino Acid	Positive	148.0604	$[M+H]^+$	-
L-Homoserine	Amino Acid	Positive	120.0655	$[M+H]^+$	-
L-Methionine	Amino Acid	Positive	150.0583	$[M+H]^+$	61.011/ 56.049
L-Serine	Amino Acid	Positive	106.0499	$[M+H]^+$	-
L-Threonine	Amino Acid	Positive	120.0655	$[M+H]^+$	84.044/ 56.049/ 74.060
N-Acetylglycine	Amino Acid	Positive	118.0499	$[M+H]^+$	-
Xanthine	oxopurine	Positive	153.0407	$[M+H]^+$	-
Taurine	amino sulfonic acid	Positive	126.0219	$[M+H]^+$	-
N-Methyl-L-glutamic acid	amino acid derivative	Positive	162.0761	$[M+H]^+$	-

Figure 32 compares the 2 different ionizations (positive, negative) that were used in the Target Screening.

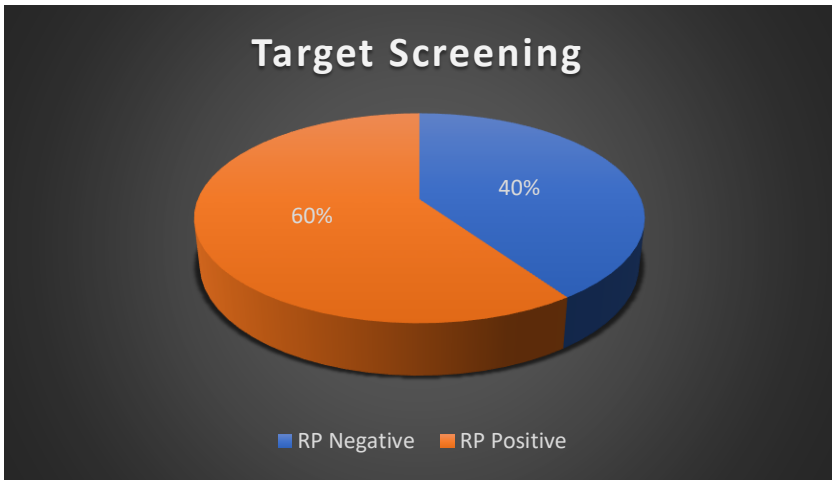


Figure 35: Comparison of the percentage (%) of the identified compounds between the 2 different ionizations (positive, negative).

Figure 33, summarizes the percentages (%) of the identified compounds in each category, based on the Table 9.

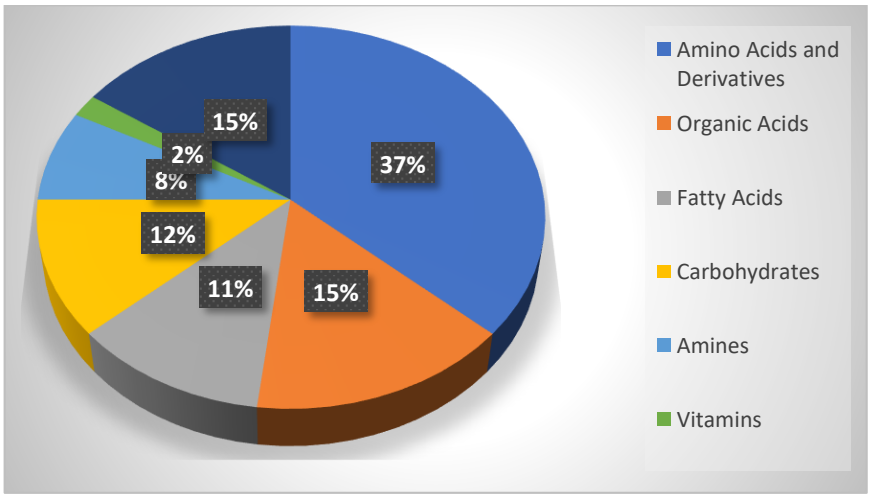


Figure 36: The percentages (%) of the identified compounds in each category, in Positive and Negative Ionization.

6.5. Suspect screening using RPLC chromatography

For the metabolite suspect screening, 1 list of compounds was generated from a literature review and tested in both ionization modes. This list contained only the molecular formula and the compound name of the suspect compounds. The RPLC list for the suspect screening contained 111 metabolites (Table 16).

The substances detected in the RPLC chromatography and in each ionization were then identified using the data processing described in section 5.7. First, the results of the LC-QTOFMS analysis were entered into TASQ Client 1.4 and the appropriate processing method was selected. Subsequently, TASQ Client calibrated all sample data. Next, screening was performed for each suspect compound in each sample to obtain the correct information from the chromatographic peaks.

A total of **58** metabolites were identified in Suspect screening in which 31 metabolites were identified in RPLC negative and 27 metabolites were identified in RPLC positive. To confirm the presence or absence of a compound in milk samples, we set the criterion that it should be detected in 90% of the samples. This means that 14 out of 15 cow's milk samples should contain the compound, and similarly 9 out of 10 sheep and goat milk samples.

The following tables present the compounds identified in cow's milk (15 samples), sheep's milk (10 samples) and goat's milk (10 samples):

Table 13: Detected metabolites in cow's milk.

Compound Name	Ionization
3-Hydroxyisovaleric acid	Negative
Alpha-D-glucose	Negative
Arachidonic acid	Negative
Aspartic acid	Negative
Butyrate	Negative
Caproic acid	Negative
cis-Aconitic acid	Negative

Compound Name	Ionization
citrate ion	Negative
D-Glucuronic Acid	Negative
DL-Lactate	Negative
Galactonic acid	Negative
glyceric acid	Negative
Homovanillic acid	Negative
Isocitric acid	Negative
L-Glutamate	Negative
linoleic acid	Negative
Malic Acid	Negative
Muricholic acid	Negative
orotate	Negative
P-Cresol	Negative
Pyridoxal	Negative
pyruvate	Negative
retinoic acid	Negative
Ricinoleic acid	Negative
α-ketoglutaric acid	Negative
5-Hydroxymethylfurfural	Positive
7-Methylguanine	Positive
Choline	Positive
Cis-5-Tetradecenoylcarnitine	Positive
Creatinine	Positive
L-Tyrosine	Positive
Monobutyl phthalate	Positive
N-Formylpiperidine	Positive
Leucylproline	Positive
Pantothenic acid	Positive
Phenylacetylglycine	Positive
N-Acetylmannosamine	Positive
Pinoresinol	Positive
Pyridoxal	Positive
Riboflavin	Positive
Ribonic acid	Positive
Salinosporamide B	Positive
Trigonelline	Positive
Trimethylamine	Positive
Xylobiose	Positive
Hydroxyoctadecadienoic acid	Positive
L-arabinose	Positive

Table 14: Detected metabolites in sheep's milk.

Compound Name	Ionization
3-Hydroxyisovaleric acid	Negative
Alpha-D-glucose	Negative
Arachidonic acid	Negative
Aspartic acid	Negative
Butyrate	Negative
Caproic acid	Negative
cis-Aconitic acid	Negative
citrate ion	Negative
D-Glucuronic Acid	Negative
DL-Lactate	Negative
Galactonic acid	Negative
glyceric acid	Negative
Homovanillic acid	Negative
Isocitric acid	Negative
L-Glutamate	Negative
linoleic acid	Negative
Malic Acid	Negative
Muricholic acid	Negative
orotate	Negative
P-Cresol	Negative
pyruvate	Negative
retinoic acid	Negative
Ricinoleic acid	Negative
6-Hydroxyhexanoic acid	Negative
Arabitol	Negative
mannose-6-phosphate	Negative
α-ketoglutaric acid	Negative
Cinnamic acid	Negative
Clupanodonic acid	Negative
5-Hydroxymethylfurfural	Positive
7-Methylguanine	Positive
Choline	Positive
Creatinine	Positive
L-Tyrosine	Positive
Monobutyl phthalate	Positive
N-Formylpiperidine	Positive
Pantothenic acid	Positive
Phenylacetylglycine	Positive
Phenylalanylproline	Positive
N-Acetylmannosamine	Positive
Riboflavin	Positive

Compound Name	Ionization
Ribonic acid	Positive
Salicyluric acid	Positive
Salinosporamide B	Positive
Uridine	Positive
indolelactic acid	Positive
L-arabinose	Positive
Glucosamine	Positive
Isoleucyl-Hydroxyproline	Positive

Table 15: Detected metabolites in goat's milk.

Compound Name	Ionization
3-Hydroxyisovaleric acid	Negative
Alpha-D-glucose	Negative
Arachidonic acid	Negative
Aspartic acid	Negative
Butyrate	Negative
Caproic acid	Negative
cis-Aconitic acid	Negative
citrate ion	Negative
D-Glucuronic Acid	Negative
DL-Lactate	Negative
Galactonic acid	Negative
glyceric acid	Negative
Homovanillic acid	Negative
Isocitric acid	Negative
L-Glutamate	Negative
linoleic acid	Negative
Malic Acid	Negative
Muricholic acid	Negative
orotate	Negative
P-Cresol	Negative
pyruvate	Negative
retinoic acid	Negative
Ricinoleic acid	Negative
Arabitol	Negative
mannose-6-phosphate	Negative
α-ketoglutaric acid	Negative
Cinnamic acid	Negative
Clupanodonic acid	Negative
Diethyl hydrogen phosphate	Negative
5-Hydroxymethylfurfural	Positive

Compound Name	Ionization
7-Methylguanine	Positive
Choline	Positive
Creatinine	Positive
L-Tyrosine	Positive
Monobutyl phthalate	Positive
N-Formylpiperidine	Positive
Pantothenic acid	Positive
Phenylacetylglycine	Positive
Phenylalanylproline	Positive
Pinoresinol	Positive
Ribonic acid	Positive
Salicyluric acid	Positive
Uridine	Positive
indolelactic acid	Positive

From the above tables, according to the criterion for confirming the presence of a compound in milk, the compounds Pyridoxal, Cis-5-Tetradecenoylcarnitine, Leucylproline, Trigonelline, Trimethylamine, Xylobiose and Hydroxyoctadecadienoic acid were detected only in cow's milk, the compounds 6-Hydroxyhexanoic acid, Glucosamine and Isoleucyl-Hydroxyproline were detected only in sheep's milk and Diethyl hydrogen phosphate was detected only in goat's milk.

Table 16 summarizes the detected metabolites, the categorization and the ionization of the detected metabolites in Suspect screening.

Table 16: Ionization (positive, negative), precursor ion, adduct of the precursor ion, categorization and fragments of the detected compounds in the Suspect Screening.

Compound name	Category	Ionization
3-Hydroxyisovaleric acid	Carboxylic Acid	Negative
Alpha-D-glucose	monosaccharide	Negative
Arachidonic acid	Dietary Fatty Acid	Negative
Aspartic acid	Amino Acids	Negative
Butyrate	Carboxylic Acid	Negative
Caproic acid	Carboxylic Acid	Negative
cis-Aconitic acid	Carboxylic Acid	Negative
citrate ion	Carboxylic Acid	Negative
D-Glucuronic Acid	Carbohydrates	Negative
DL-Lactate	Carboxylic Acid	Negative
Galactonic acid	Carbohydrates	Negative
glyceric acid	Carbohydrate	Negative
Homovanillic acid	Carboxylic Acid	Negative
Isocitric acid	Carboxylic Acid	Negative
L-Glutamate	Amino Acids	Negative
linoleic acid	Fatty Acid	Negative
Malic Acid	Carboxylic Acid	Negative
Muricholic acid	Cholic Acids	Negative
orotate	carboxylic acid anion	Negative
P-Cresol	Phenols	Negative
Pyridoxal	Vitamin B Complex	Negative
pyruvate	Carboxylic Acid	Negative
retinoic acid	Vitamin	Negative
Ricinoleic acid	hydroxy fatty acid	Negative
6-Hydroxyhexanoic acid	Carboxylic Acid	Negative
Arabitol	Carbohydrate	Negative
mannose-6-phosphate	Carbohydrates	Negative
Cinnamic acid	Carboxylic Acid	Negative
α-ketoglutaric acid	keto acid	Negative

Compound name	Category	Ionization
Clupanodonic acid	Fatty Acid	Negative
Diethyl hydrogen phosphate	Organophosphates	Negative
5-Hydroxymethylfurfural	Aldehyde	Positive
7-Methylguanine	Purine	Positive
Choline	Amine, Vitamin	Positive
Cis-5-Tetradecenoylcarnitine	O-acylcarnitine	Positive
Creatinine	lactam	Positive
L-Tyrosine	Amino Acid	Positive
Monobutyl phthalate	Carboxylic Acid	Positive
N-Formylpiperidine		Positive
Leucylproline	Peptide	Positive
Pantothenic acid	Vitamin	Positive
Phenylacetylglycine	Amino Acid	Positive
Phenylalanylproline	Peptide	Positive
N-Acetylmannosamine	Carbohydrate	Positive
Pinoresinol	Lignan	Positive
Riboflavin	Vitamin	Positive
Ribonic acid	carboxylic acid	Positive
Salicyluric acid	Carboxylic Acid	Positive
Salinosporamide B	Lactam	Positive
Trigonelline	Alkaloid	Positive
Trimethylamine	Amine	Positive
Uridine	Nucleoside	Positive
Xylobiose	Disaccharide	Positive
Hydroxyoctadecadienoic acid	Fatty Acid	Positive
indolelactic acid	Indole	Positive
L-arabinose	Monosaccharides	Positive
Glucosamine	Carbohydrate	Positive
Isoleucyl-Hydroxyproline	peptide	Positive

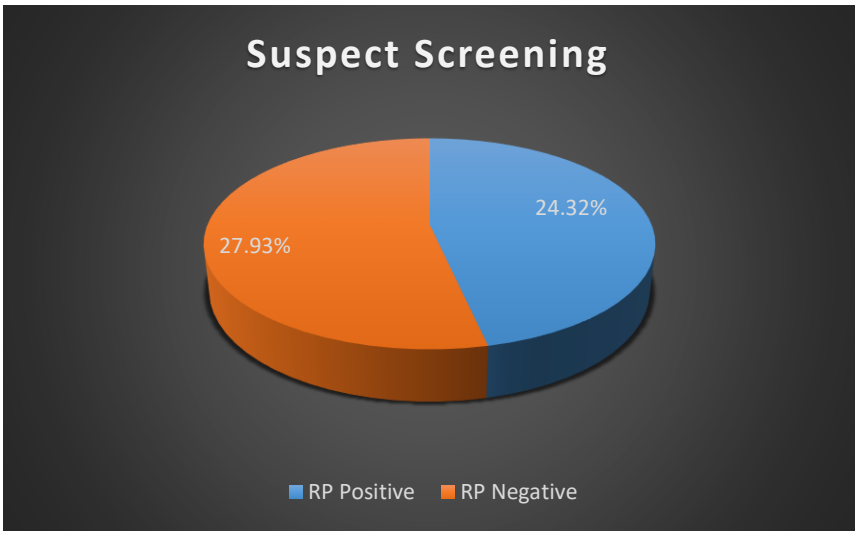


Figure 37: Comparison of the percentage (%) of the detected compounds between the 2 different ionizations (positive, negative).

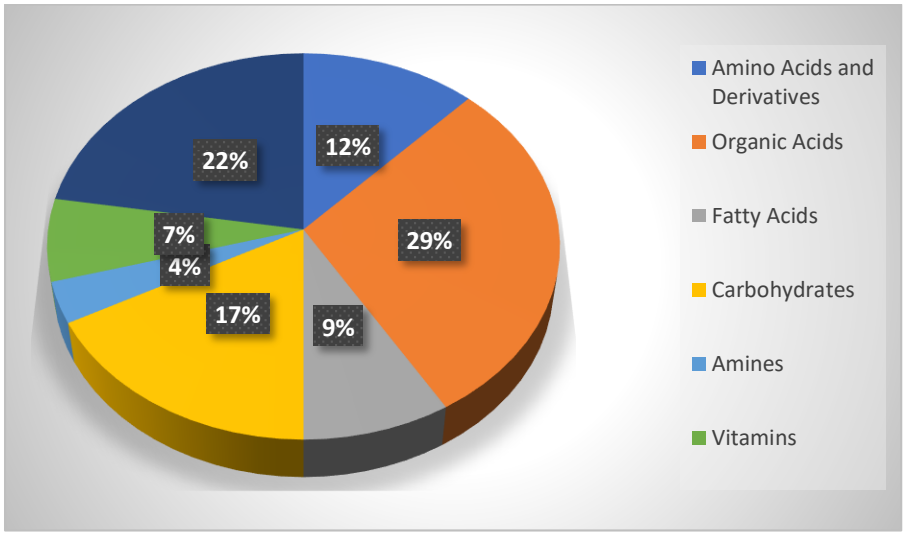


Figure 38: The percentages (%) of the detected compounds in each category, in Positive and Negative Ionization.

Out of the 110 metabolites that were identified in both target and suspect screening, the 58 metabolites of the suspect screening showed higher ion intensities than the 52 metabolites of the target screening. In particular, almost all of the detected metabolites in the suspect screening showed ion intensities above 10.000 and peak areas more than 2.000. In the target screening the majority of metabolites showed low ion intensities (< 6.000), although also their fragments were detected. The results showed high mass accuracy (<5 ppm), acceptable isotopic fit values (< 100 mSigma for the Target Screening and 100-200 mSigma for the Suspect Screening) and an acceptable retention time window (0.2-0.4 minutes).

6.6. Milk sample discrimination study: Animal Origin

In order to determine whether the samples of each milk type belonged to the same origin (cow, sheep, goat), the peak data from the different milk samples were used to create a PCA model.

6.6.1. PCA in Target Screening

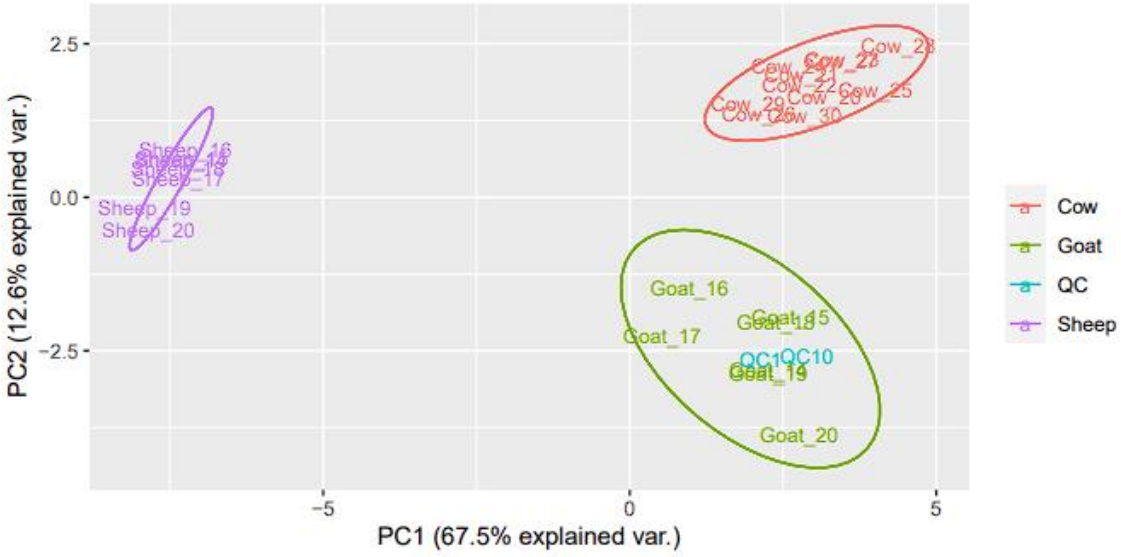


Figure 39: PCA plot for cow, sheep and goat milk in RP Positive mode.

The three types of milk are completely separated in this chemometric model, as shown in the PCA diagram (Figure 36). That means that the model works successfully for the classification in positive ionization.

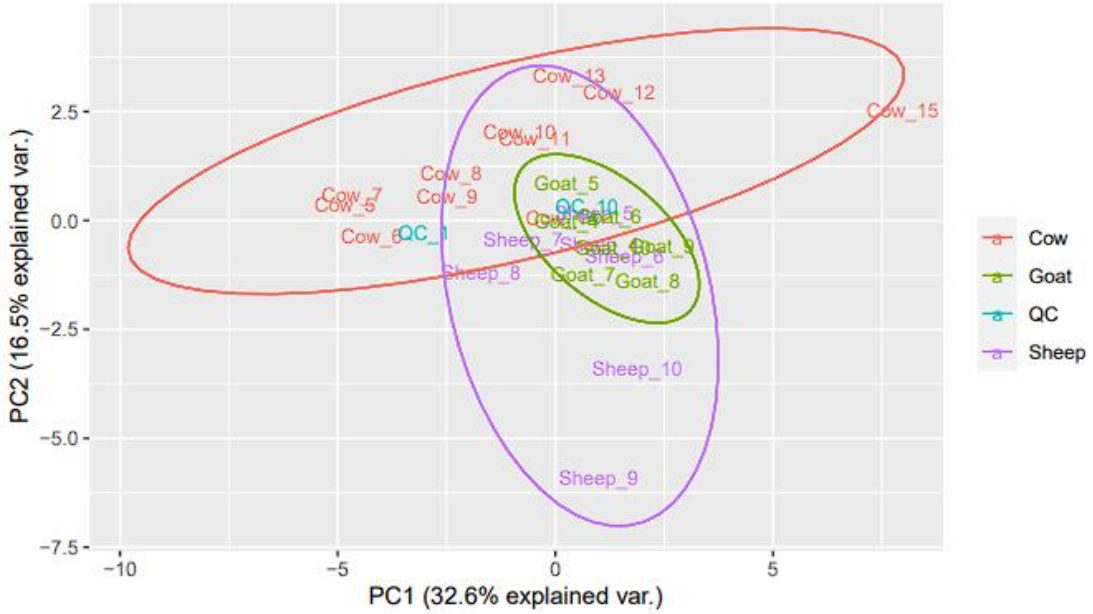


Figure 40: PCA plot for cow, sheep and goat milk in RP Negative mode.

As shown in the PCA diagram (Figure 37), the three types of milk samples were indistinguishable from each other, so the model didn't work well in Negative ionization.

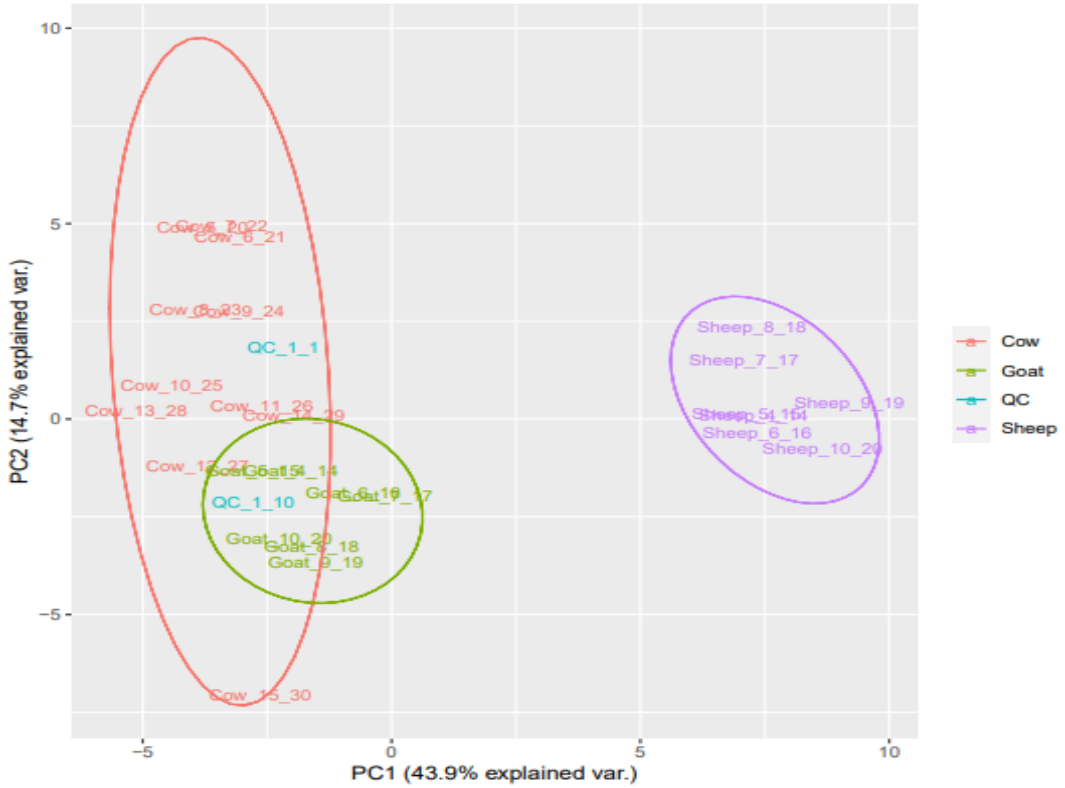


Figure 41: PCA plot for cow, sheep and goat milk in RP Negative and Positive mode.

As shown in the Merged (Positive and Negative) PCA plot, sheep's milk was completely separated from the other two types of milk, but this model couldn't distinguish cow's and goat's milk.

6.6.2. PLS-DA model in Target Screening

For the development of the PLS-DA model, 70% of the data was set as the training set and 30% of the final used data was set as the test set to evaluate the accuracy of the models. The choice of which samples would be the training data and which would be the test data was random, but the 70:30 ratio was maintained across groups.

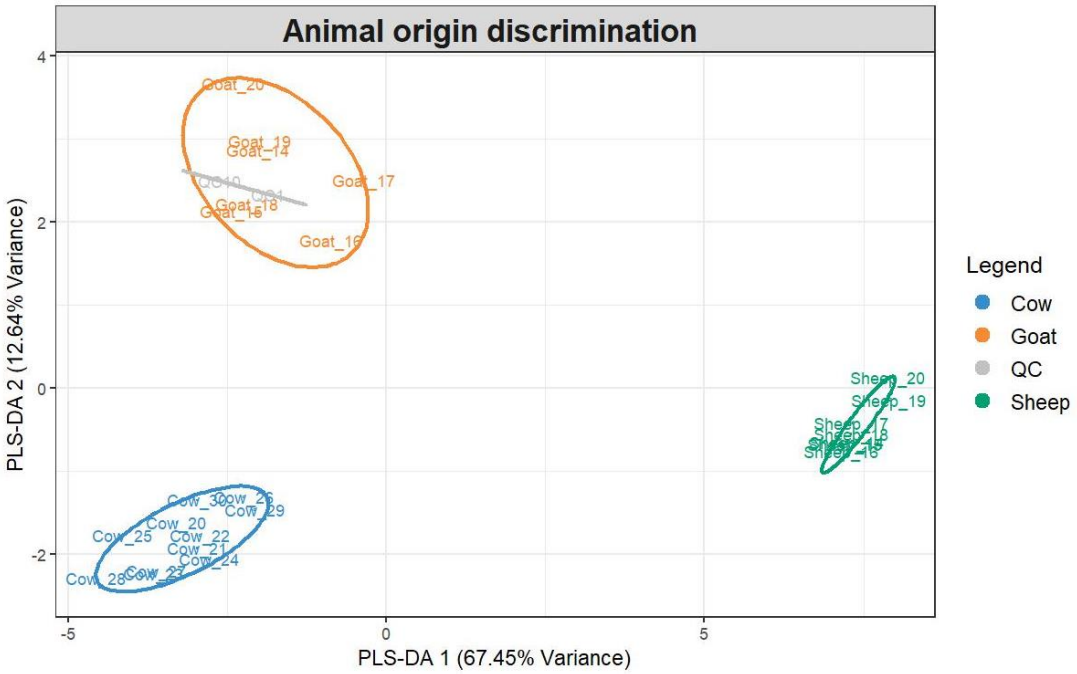


Figure 42: PLS-DA chart for cow, sheep and goat milk in RP Positive mode.

For the discrimination of authentic cow, sheep and goat milk, PLS-DA provided similar results to the PCA diagram (Figure 36). It is obvious that the milk samples were completely separated from each other, regarding to animal origin, in positive ionization.

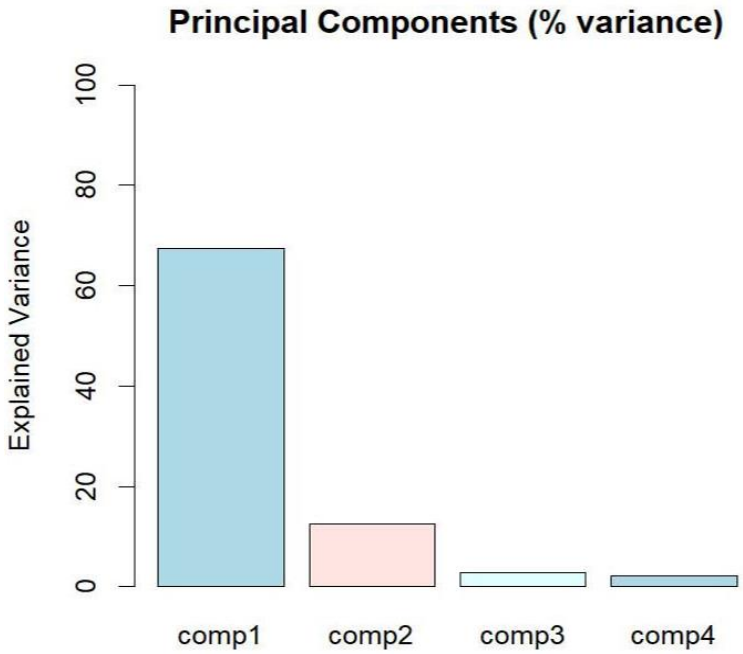


Figure 43: Evaluation of Principal Components (PCs) importance in RP Positive mode.

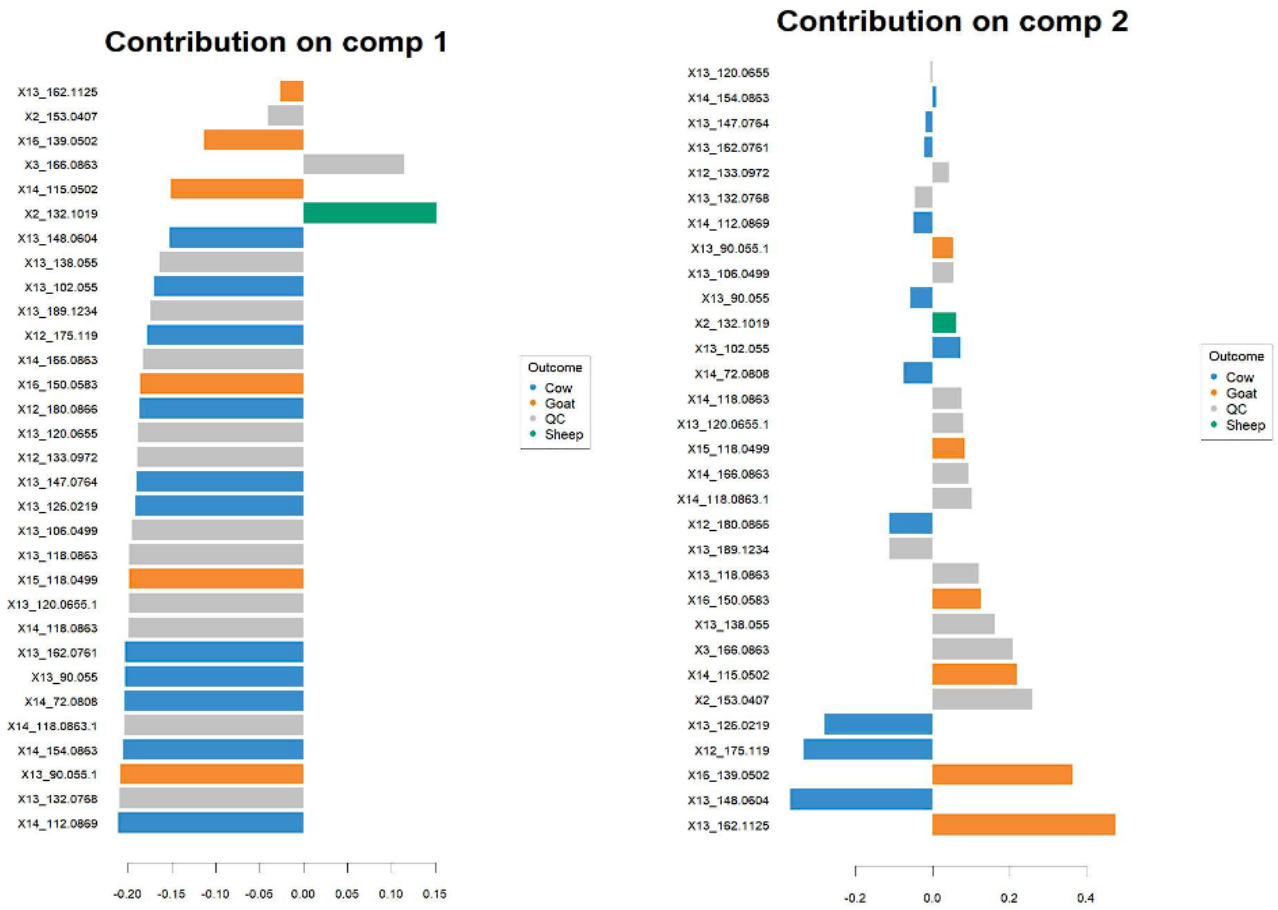


Figure 44: VIP score of the first two PCs of the discrimination study, in RP Positive mode..

In the figures are recorded the most important m/z values for each milk type depending on their contribution to PLS-DA diagram.

Table 17: VIP values for each milk type

COW	GOAT	SHEEP
1-Aminocyclopropanecarboxylic acid	4-Imidazoleacrylic acid	L-Isoleucine
L-Glutamine	Dihydrouracil	
D-Mannosamine hydrochloride	L-Alanine	
Dopamine hydrochloride	L-Carnitine hydrochloride	
L-Valine	L-Methionine	
Histamine dihydrochloride	N-Acetylglycine	
L-Arginine monohydrochloride		
L-Glutamic acid		
Taurine		
N-Methyl-L-glutamic acid		
Sarcosine		

.

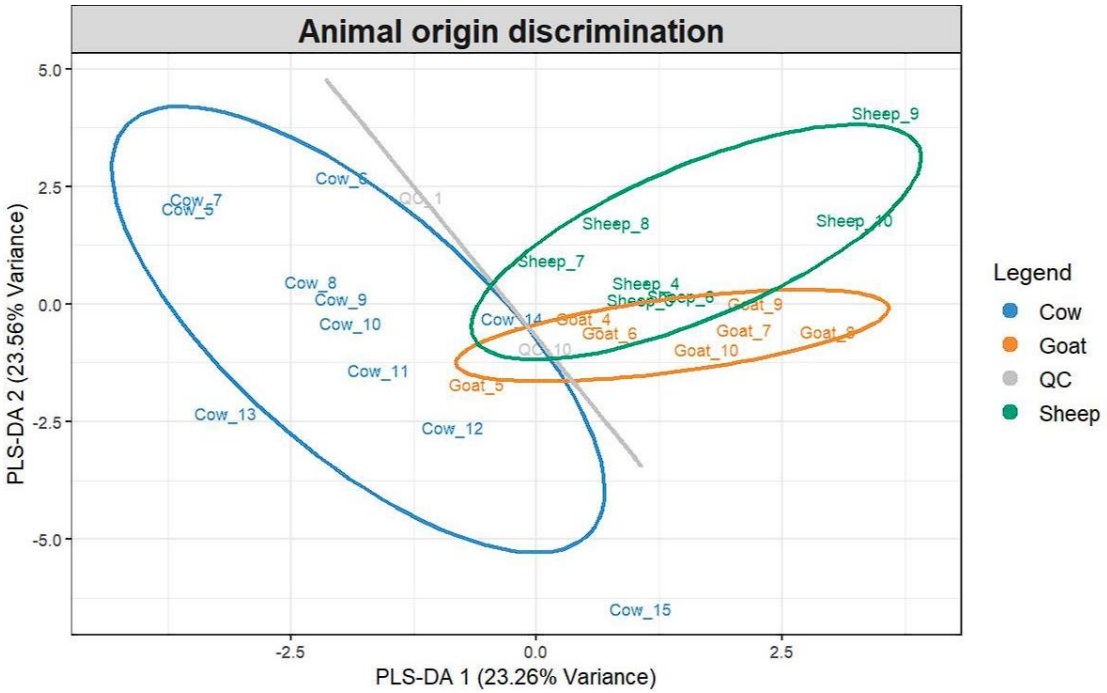


Figure 45: PLS-DA chart for cow, sheep and goat milk in RP Negative mode.

As can be seen in Figure 41, the application of the PLS-DA model didn't completely separate the milk samples in the Negative ionization, although a better classification can be observed than with the PCA model. The cow's milk samples were a bit more separated from the other groups.

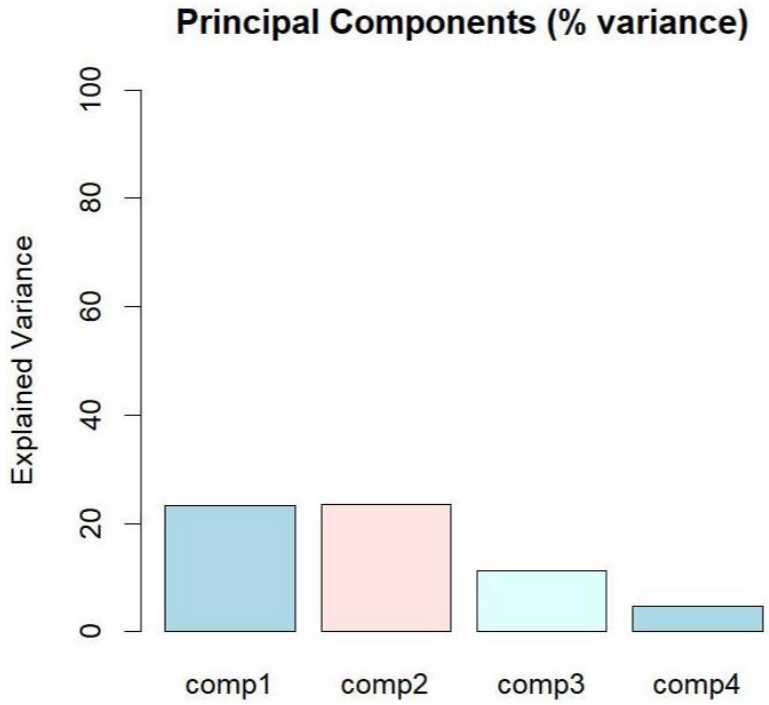


Figure 46: Evaluation of Principal Components (PCs) importance in RP Negative mode.

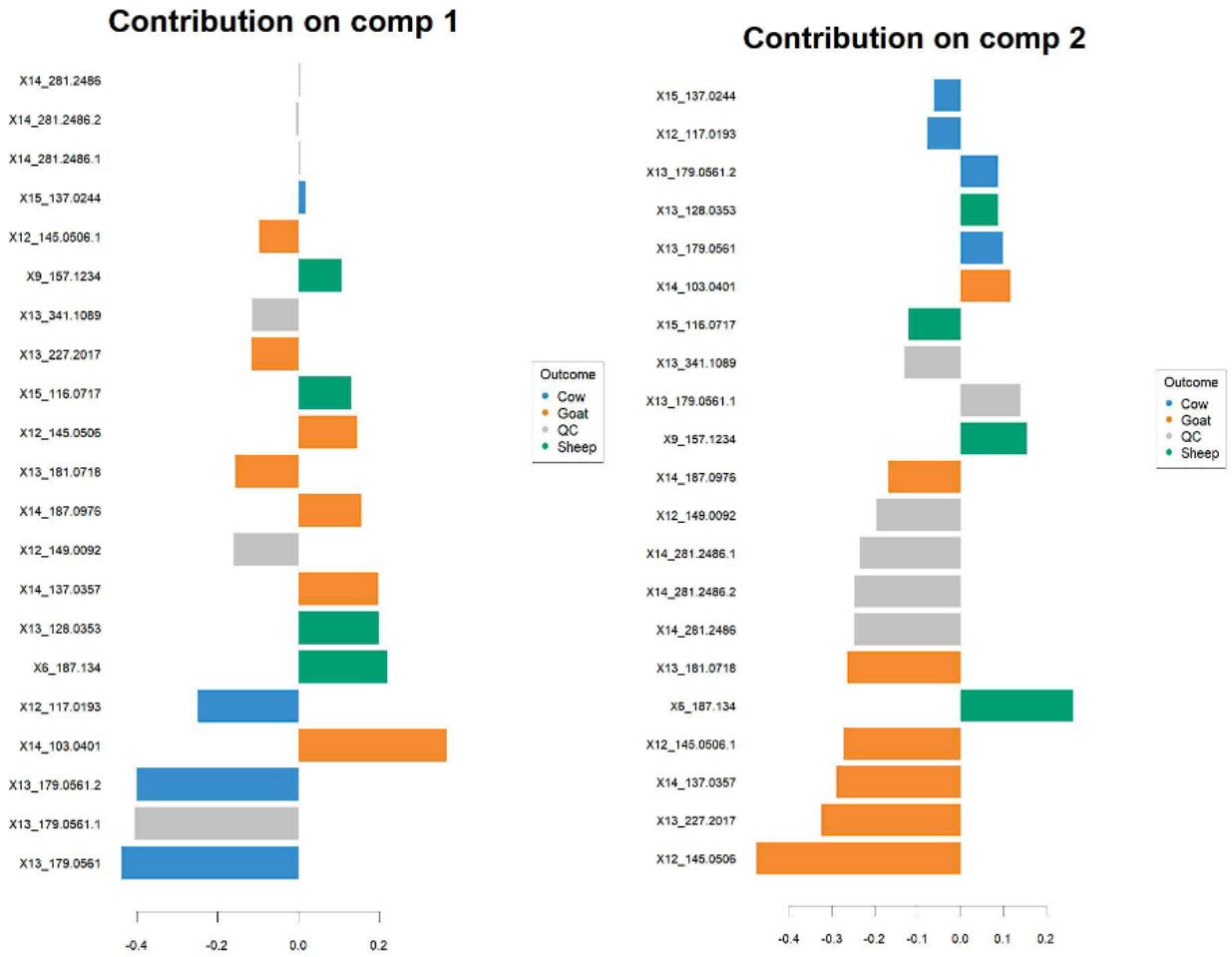


Figure 47: VIP score of the first two PCs of the discrimination study, in RP Negative mode.

Table 18: VIP values for each milk type.

COW	GOAT	SHEEP
4-Hydroxybenzoic acid	2-Hydroxybutyric acid sodium salt	10-Hydroxydecanoic acid
D-(+)-Galactose	2-Methylglutaric acid	L-Norvaline
Succinic acid	3-Methylglutaric acid	L-Pyroglutamic acid
Tagatose	4-Imidazoleacrylic acid	Nonanoate
	Azelaic acid	
	Dulcitol	
	Myristate	

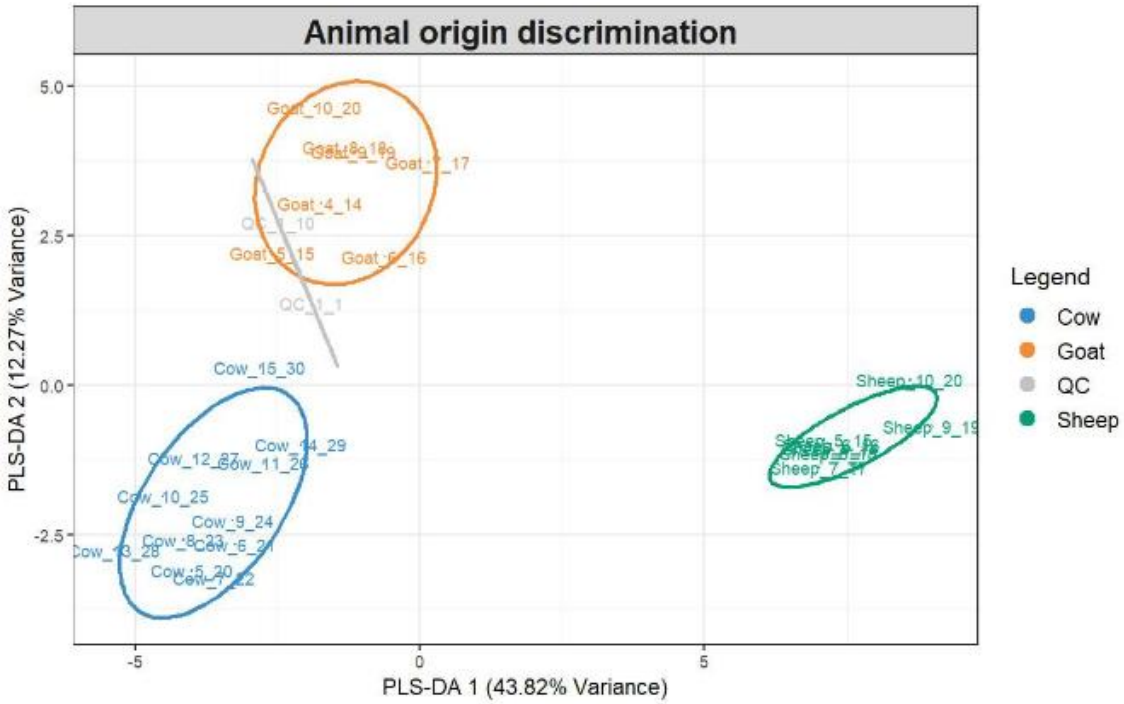


Figure 48: PLS-DA chart for cow, sheep and goat milk in RP Negative and Positive mode.

Unlike the PCA model, the PLS-DA model (figure 45) was able to fully discriminate the three types of milk in both ionizations (positive and negative) simultaneously.

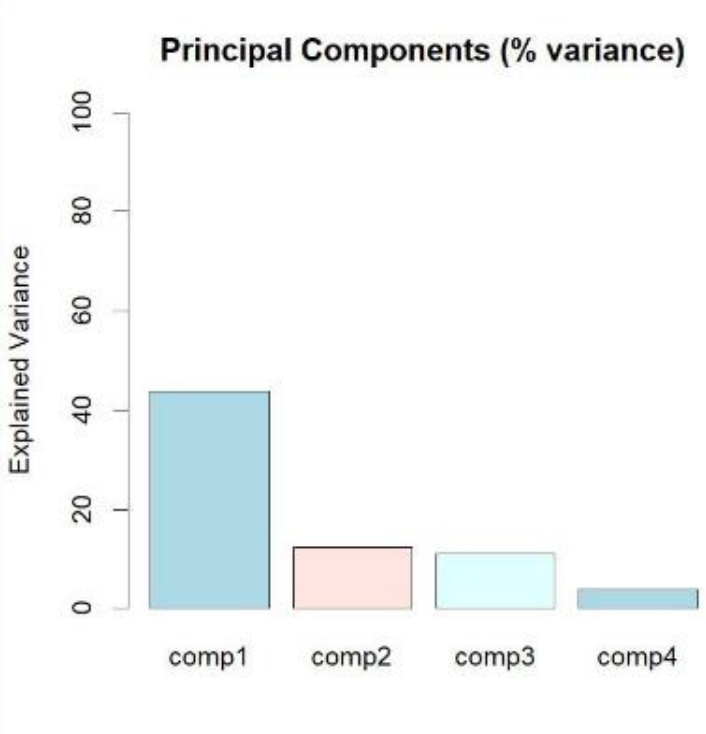


Figure 49: Evaluation of Principal Components (PCs) importance in RP Negative and Positive mode.

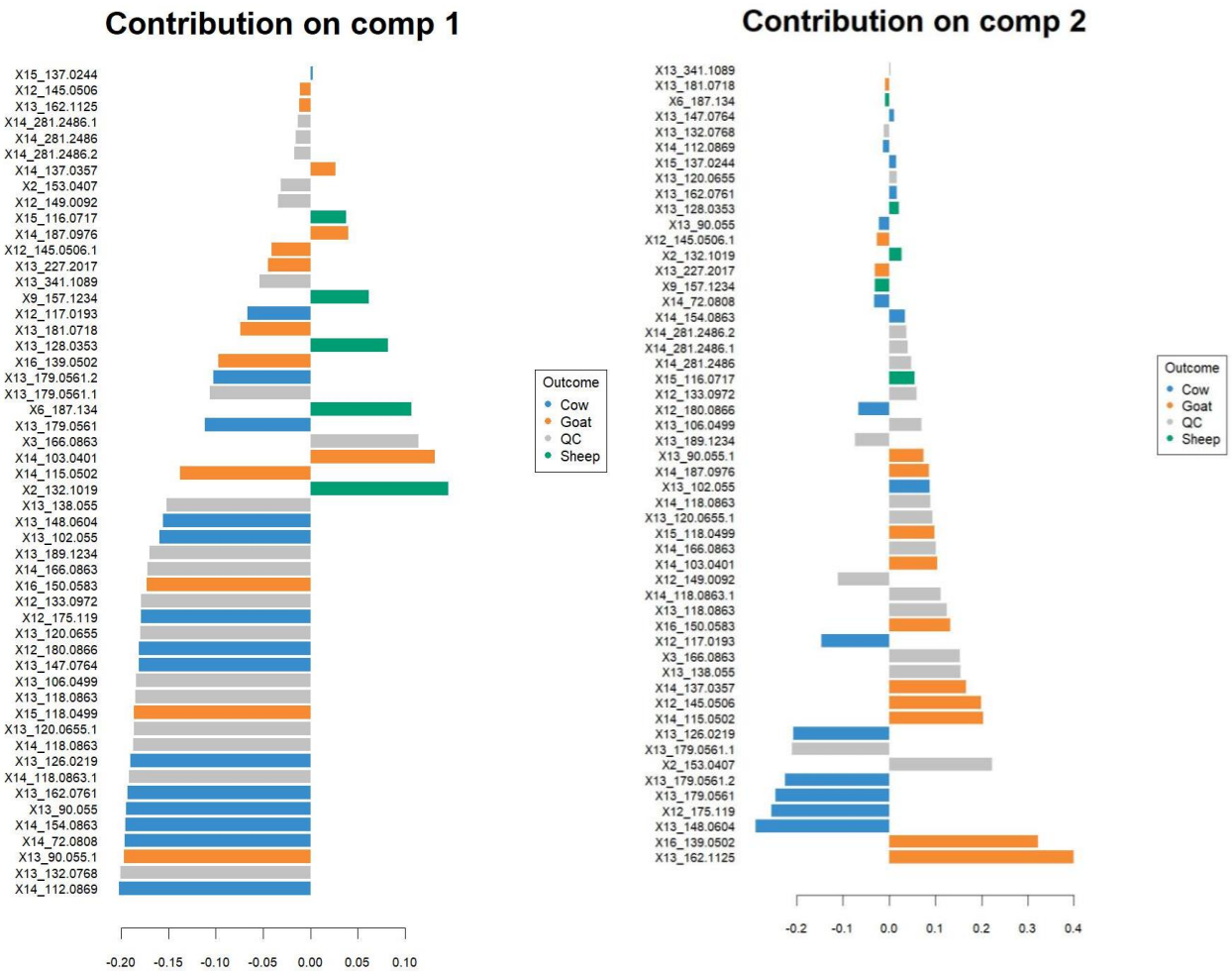


Figure 50: VIP score of the first two PCs of the discrimination study.

6.6.3. PCA in Suspect Screening

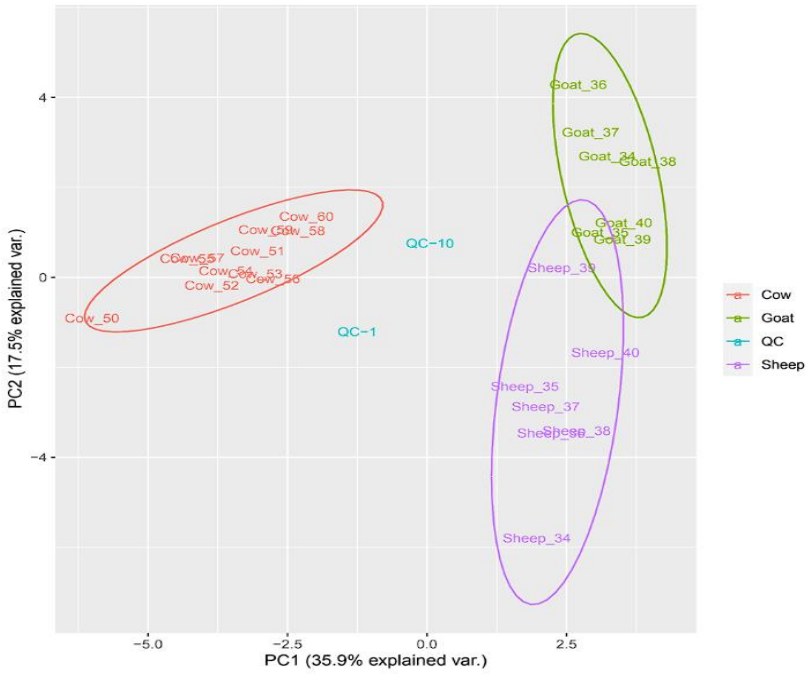


Figure 51: PCA plot for cow, sheep and goat milk in RP Positive mode.

The three types of milk are sufficiently separated in this chemometric model, as shown in the PCA diagram (Figure 45), with cow's milk completely separated from goat and sheep's milk. Goat's and sheep's milk samples were too close to each other, but the discrimination was also good.

6.6.4. PLS-DA model in Suspect Screening

For the development of the PLS-DA model, 70% of the data was set as the training set and 30% of the final used data was set as the test set to evaluate the accuracy of the models. The choice of which samples would be the training data and which would be the test data was random, but the 70:30 ratio was maintained across groups.

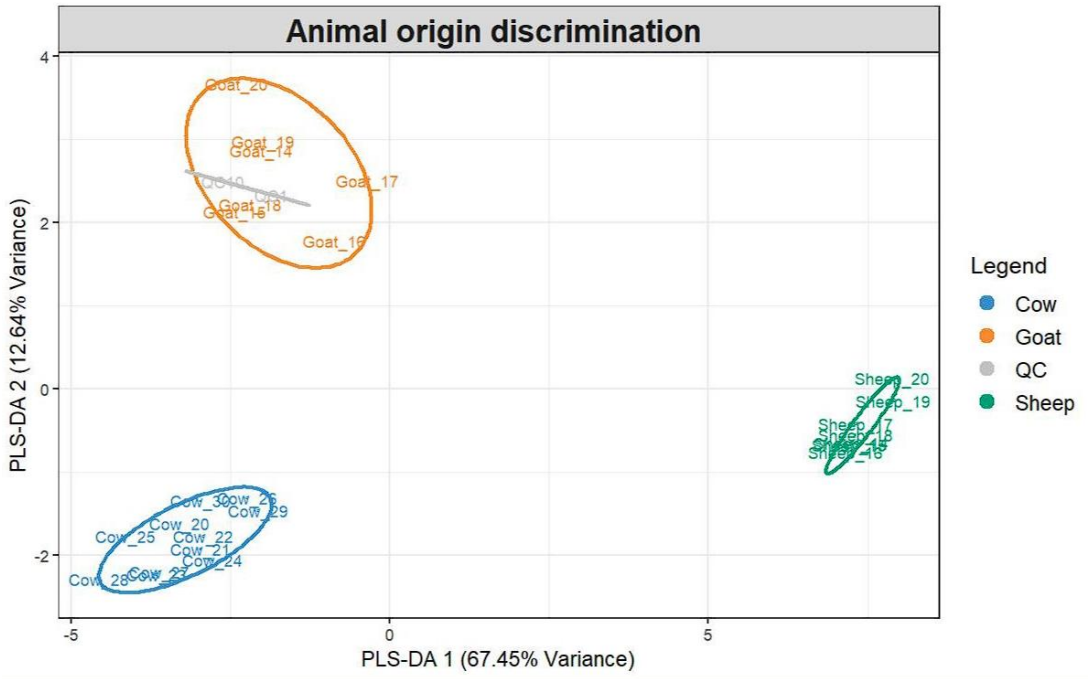


Figure 52: PLS-DA chart for cow, sheep and goat milk in RP Positive mode.

For the discrimination of authentic cow, sheep and goat milk, PLS-DA provided similar results to the PCA diagram (Figure 44). It is obvious that the milk samples were completely separated from each other, regarding to animal origin, in positive ionization.

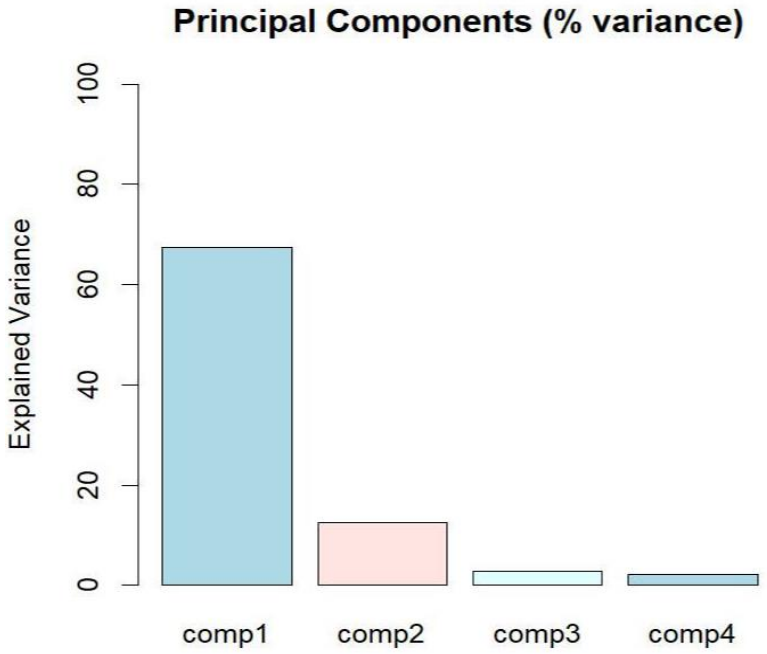
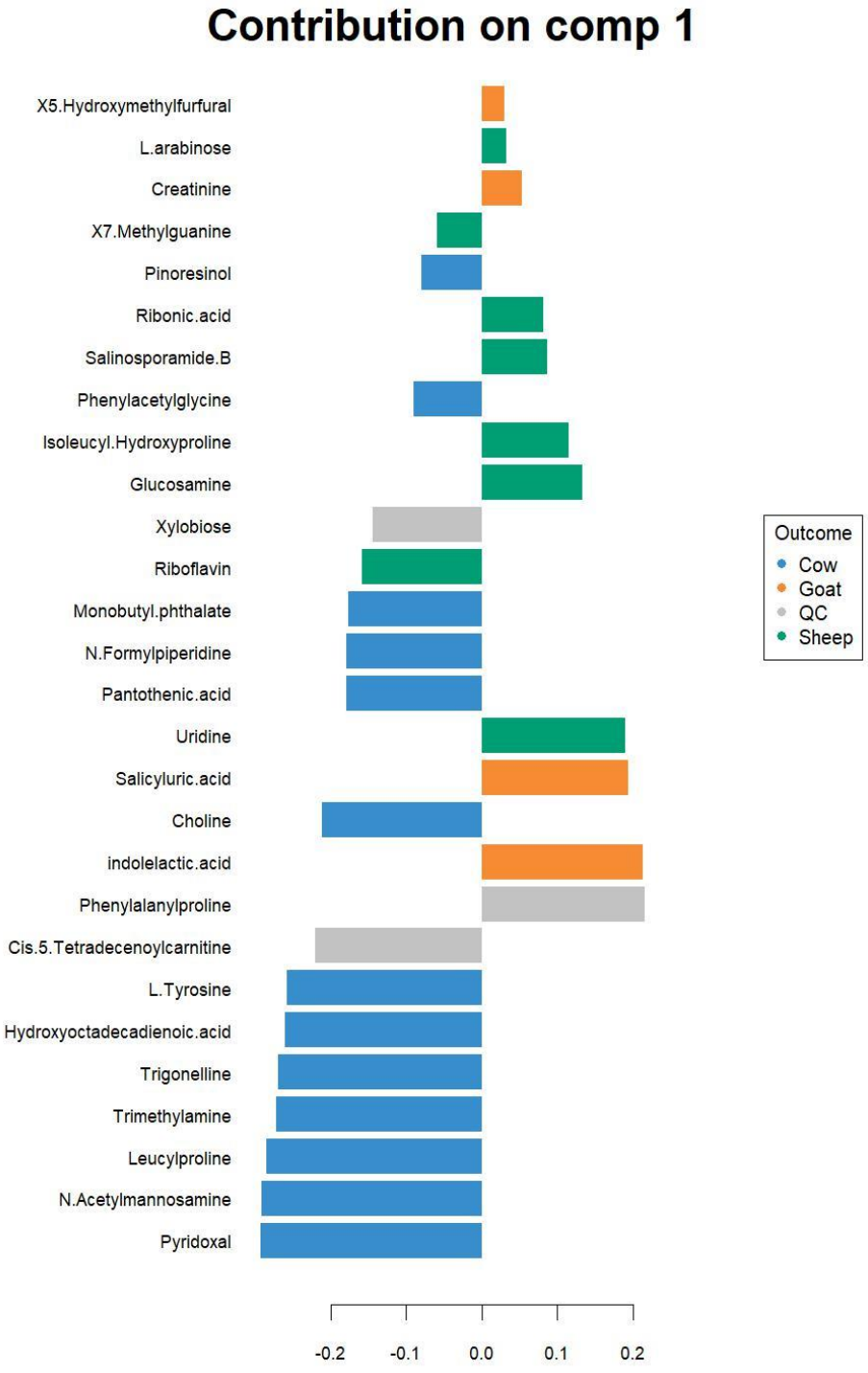


Figure 53: Evaluation of Principal Components (PCs) importance in RP Positive mode.


Contribution on comp 1
X5.Hydroxymethylfurfural
L.arabinose
Creatinine
X7.Methylguanine
Pinoresinol
Ribonic.acid
Salinosporamide.B
Phenylacetylglycine
Isoleucyl.Hydroxyproline
Glucosamine
Xylobiose
Riboflavin
Monobutyl.phthalate
N.Formylpiperidine
Pantothenic.acid
Uridine
Salicyluric.acid
Choline
indolelactic.acid
Phenylalanylproline
Cis.5.Tetradecenoylcarnitine
L.Tyrosine
Hydroxyoctadecadienoic.acid
Trigonelline
Trimethylamine
Leucylproline
N.Acetylmannosamine
Pyridoxal
-0.2
-0.1
0.0
0.1
0.2
Outcome
Cow
Goat
QC
Sheep

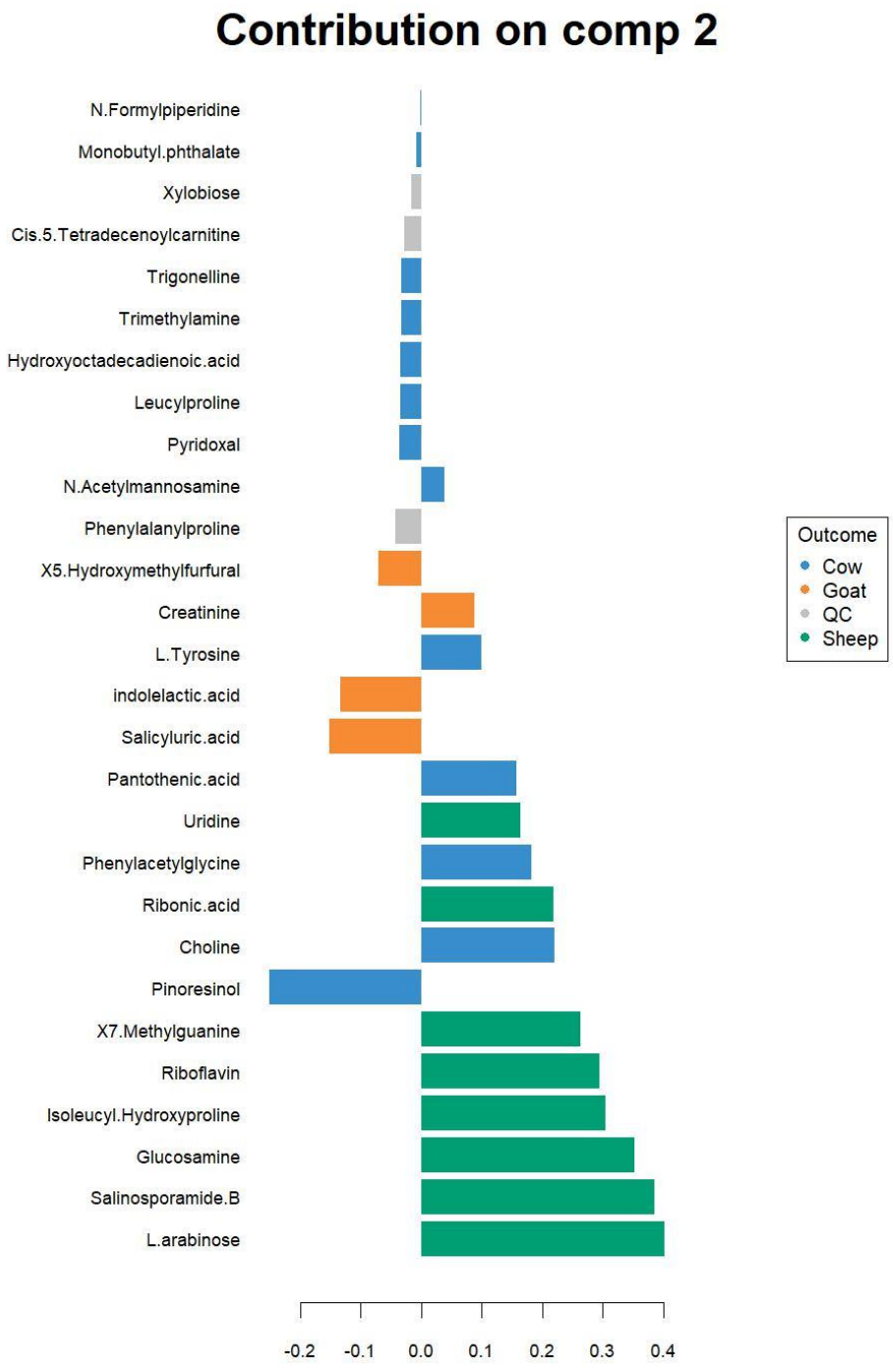


Figure 54: VIP score of the first two PCs of the discrimination study.

CHARTER 7

CONCLUSIONS

In the field of food analysis, there is a constant need for the development of more robust, efficient and cost-effective analytical methods and workflows for food authentication in order to ensure the safety, quality and traceability of food products in accordance with legislation and consumer requirements. Metabolomics is a promising tool in dairy science that has emerged as a valuable tool for screening and profiling metabolites in complex matrices such as milk. Given their considerable economic and nutritional value, milk and dairy products are often targeted for adulteration. The most common form of adulteration is the deliberate substitution or addition of a substance to a dairy product or the mixing of milk from different animal origins in order to increase its perceived value or reduce its production costs.

The aim of the specific thesis was to detect discriminating biomarkers between milk from different animal origins, such as cow's, sheep's and goat's milk. This was achieved by using PCA discriminant analysis and PLS-DA models, which were successfully developed to discriminate between authentic cow's, sheep's and goat's milk. In Target screening, level 1 of identification was achieved, because we had the necessary information such as retention time, m/z values, fragments, molecular formula etc. In Suspect screening, the Investigation of discriminating biomarkers has also been achieved, but the only available information was the compound name and its molecular formula. So, level 4 of identification was achieved. The future goals of this study are:

- the quantification of all the metabolites, found in target screening,
- the enrichment of the in-laboratory database with the above suspect metabolites after further confirmation. We want to examine all the fragments and fragmentation pattern (by using AutoMs samples), so we can reach a higher level of identification,
- the extension of the applicability of the method to other foods of either animal or plant origin.

ABBREVIATIONS – ACRONYMS

bbCID	broad-band Collision Induced Dissociation
BRC	British Retail Consortium
DNA	Deoxyribonucleic acid
EC	European Council
ECR	Europe Council Regulation
ESI	Electrospray Ionisation
EU	European Union
FT-ICR	Fourier Transform Ion Cyclotron Resonance
GC-MS	Gas Chromatography-Mass Spectrometry
HILIC	Hydrophilic Interaction Liquid Chromatography
HPLC	High-Performance Liquid Chromatography
HPLC/MS	Liquid Chromatography coupled with Mass Spectrometry
HRMS	High-Resolution Mass Spectrometry
ISO	International Organization for Standardization
IT	Quadrupole ion trap

LC-HRMS	Liquid Chromatography – High-Resolution Mass Spectrometry
LC-MS	Liquid Chromatography – Mass Spectrometry
LC-MS/MS	Liquid Chromatography-tandem MS
LE	Liquid Extraction
MS	Mass Spectrometry
NMR	Nuclear Magnetic Resonance
PCA	Principal Component Analysis
PDO	Protected Designation of Origin
PGI	Protected Geographical Indication
PLS-DA	Partial Least Square Discriminant Analysis
Q	Quadrupole
QC	Quality Control
QTOF	Quadrupole-Time-of-flight
RPLC	Reversed Phase Liquid Chromatography
RP-UHPLC	Reversed Phase Ultra High-Performance Liquid Chromatography
TOF	Time-of-flight
TSG	Traditional Specialties Guaranteed

UHPLC	Ultra High Performance Liquid Chromatography

APPENDIX

Table 19: Metabolites Compound list for RPLC Target Screening in both ionization modes.

RPLC Positive	RPLC Negative
b-Nicotinamide adenine dinucleotide hydrate	L-Glutamine
L-Glutamine	Citric acid
Hypotaurine	L-Threonine
Inosine 5'-monophosphate from Saccharomyces cerevisiae	Purine
L-Threonine	Uric acid
Purine	Cytidine
N-Acetylneuraminic acid	L-Serine
Cytidine	L-Cysteine
L-Serine	L-Citrulline
L-Citrulline	Taurine
Taurine	Nicotinic acid
Nicotinic acid	Inosine
γ-Aminobutyric acid	γ-Aminobutyric acid
Cytosine	Cytosine
L-Isoleucine	L-Isoleucine
L-Glutamic acid	L-Glutamic acid
L-Ascorbic acid	4-Hydroxyphenylacetic acid
N-Acetyl-D-glucosamine	N-Acetyl-D-glucosamine
Sarcosine	Sarcosine
(-) Quinic acid	(-) Quinic acid
L-Methionine	L-Dihydroorotic acid
Adenine	Malonic acid
5'-Deoxy-5'-(methylthio)adenosine	L-Methionine
Thymidine	Adenine
O-Phosphorylethanolamine	5'-Deoxy-5'-(methylthio)adenosine
Xanthine	Thymidine
L-Alanine	Xanthine
Uridine 5'-monophosphate	β-Nicotinamide adenine dinucleotide hydrate
Thymine	L-Glutamine
L-Carnosine	Hypotaurine
Nicotinamide	Inosine 5'-monophosphate from Saccharomyces cerevisiae
L-Phenylalanine	Citric acid
Uracil	L-Threonine
2'-Deoxycytidine 5'-monophosphate	Purine
Hypoxanthine	N-Acetylneuraminic acid

Creatine	Uric acid
3,4-Dihydroxy-L-phenylalanine	Succinic acid
Dihydrouracil	L-Phenylalanine
L-Asparagine	Uracil
L-Valine	2'-Deoxycytidine 5'-monophosphate
Guanine	Hypoxanthine
L-Homoserine	L-Asparagine
L-(+)-Tartaric acid	L-Valine
β-Nicotinamide mononucleotide	L-Homoserine
Folic acid	Pyridoxine
2,6-Diaminopimelic acid	2,6-Diaminopimelic acid
DL-2-Aminoadipic acid	3-Dehydroshikimic acid
2'-Deoxycytidine	Theophylline
Theophylline	Allantoin
Betaine	D-(+)-Glucuronic acid γ-lactone
O-Succinyl-L-homoserine	2-Aminoethylphosphonic acid
Allantoin	NG,NG'-Dimethyl-L-arginine di(p-hydroxyazobenzene-p'-sulfonate) salt
2-Aminoethylphosphonic acid	1,7-Dimethylxanthine
NG,NG'-Dimethyl-L-arginine di(p-hydroxyazobenzene-p'-sulfonate) salt	Dulcitol
D-(+)-Glucosamine hydrochloride	(R)-(+)-2-Pyrrolidone-5-carboxylic acid
1-Methyladenosine	4-Pyridoxic acid
2'-Deoxyuridine	2,3-Pyridinedicarboxylic acid
4-Imidazoleacrylic acid	D-Glucuronic acid sodium salt monohydrate
L-Pyroglutamic acid	1-Methyladenosine
4-ACETAMIDOBUTYRIC ACID	2'-Deoxyuridine
Dopamine hydrochloride	4-Imidazoleacrylic acid
Arabinose	L-Pyroglutamic acid
Cytidine 5'-diphosphocholine sodium salt dihydrate	4-ACETAMIDOBUTYRIC ACID
1,3-Diaminopropane	L-Norvaline
1-Aminocyclopropanecarboxylic acid	3-Hydroxy-3-methylglutaric acid
Glutaryl-L-carnitine lithium salt	myo-Inositol
L-Cystathionine	L-Homocystine
3-Hydroxy-3-methylglutaric acid	N-Methyl-L-glutamic acid
L-Arginine monohydrochloride	Thyrotropin releasing hormone
L-Homocystine	N-Methyl-D-aspartic acid
N-Methyl-L-glutamic acid	Nα-Acetyl-L-asparagine
Thyrotropin releasing hormone	D-(+)-Galactose
N-Methyl-D-aspartic acid	N-Acetylglycine
Nα-Acetyl-L-asparagine	N-Acetyl-L-aspartic acid

Agmatine sulfate salt	Adenosine 2':3'-cyclic monophosphate sodium salt
N-Acetylglycine	Sucrose
N-Acetyl-L-aspartic acid	O-Phosphorylethanolamine
Palmitoyl-DL-carnitine chloride	Adenosine-5'-diphosphoglucose disodium salt
Nicotinamide hypoxanthine dinucleotide sodium salt	Nα-Acetyl-L-lysine
S-(5'-Adenosyl)-L-methionine p-toluenesulfonate salt	5-Aminolevulinic acid hydrochloride
(3-Carboxypropyl)trimethylammonium chloride	2'-Deoxyuridine 5'-monophosphate disodium salt
Adenosine 2':3'-cyclic monophosphate sodium salt	2'-Deoxyguanosine monohydrate
Phosphocholine chloride calcium salt tetrahydrate	Indoxyl sulfate potassium salt
Thiamine hydrochloride	2,4-Dihydroxypyrimidine-5-carboxylic acid
5-Methylcytosine hydrochloride	N-Acetyl-L-alanine
Cytidine 2':3'-cyclic monophosphate monosodium salt	4-Guanidinobutyric acid
Nε,Nε,Nε-Trimethyllysine hydrochloride	N-Acetyl-DL-glutamic acid
Pyridoxal 5'-phosphate hydrate	Inosine 5'-monophosphate disodium salt hydrate
L-Carnitine hydrochloride	Aniline-2-sulfonic acid
O-Phosphorylethanolamine	L-Norleucine
3-Nitro-L-tyrosine	Adenosine
(+/-)-Octopamine hydrochloride	meso-Tartaric acid monohydrate
Nα-Acetyl-L-lysine	D-Saccharic acid potassium salt
Pyridoxamine dihydrochloride	Tagatose
5-Aminolevulinic acid hydrochloride	D-Lactose monohydrate
Flavin adenine dinucleotide disodium salt hydrate	D-(+)-Galacturonic acid monohydrate
2'-Deoxyguanosine monohydrate	cis-4-Hydroxy-D-proline
Lauroyl-L-carnitine	N-Formyl-L-methionine
N-Acetyl-DL-methionine	5-Hydroxy-L-tryptophan
Guanosine 3',5'-cyclic monophosphate sodium salt	(+/-)-Potassium citramalate monohydrate
L-Homocysteine thiolactone hydrochloride	Ophthalmate
L-Ornithine monohydrochloride	2-Hydroxybutyric acid sodium salt
DL-Normetanephrine hydrochloride	N-Acetyl-DL-serine
Uridine 5'-diphospho-N-acetylglucosamine sodium salt	3-Methylglutaric acid
NG,NG-Dimethylarginine dihydrochloride	4-Hydroxybenzoic acid
Selenocystamine dihydrochloride	Melatonin
Histamine dihydrochloride	Maleic acid
Thiamine monophosphate chloride dihydrate	Pregnenolone sulfate sodium salt

N-Acetyl-L-alanine	3-Hydroxybenzyl alcohol
4-Guanidinobutyric acid	3,5-Diiodo-L-tyrosine dihydrate
D-Mannosamine hydrochloride	Mandelic acid
2'-Deoxyadenosine monohydrate	Caffeic acid
N-Acetylputrescine hydrochloride	N-Acetyl-L-phenylalanine
N-Acetyl-D-galactosamine	N-Acetyl-L-proline
N-Acetyl-DL-glutamic acid	2,3-Dihydroxybenzoic acid
Lumazine	3-(2-Hydroxyethyl)indole
D-Pantothenic acid hemicalcium salt	4-Hydroxy-3-methoxycinnamic acid
2-Aminoisobutyric acid	6-(METHYLTHIO)PURINE
Aniline-2-sulfonic acid	2-Hydroxy-4-(methylthio)butyric acid calcium salt
L-Histidinol dihydrochloride	3,4-Dihydroxybenzoic acid
Sorbitol	(R)-(-)-Mandelic acid
3-Ureidopropionic acid	10-Hydroxydecanoic acid
5-Aminovaleric acid	Methylmalonic acid
L-Norleucine	Indole-3-acetamide
Adenosine	Hippuric acid
D-Lactose monohydrate	Ethylmalonic acid
4-Imidazoleacetic acid hydrochloride	3,5-Diiodo-L-thyronine
Methyl 4-aminobutyrate hydrochloride	4-Hydroxybenzaldehyde
N-Formyl-L-methionine	Pimelic acid
5-Hydroxy-L-tryptophan	3-Amino-4-hydroxybenzoic acid
Theobromine	2,5-Dihydroxybenzoic acid
Ophthalmate	4-Quinolinecarboxylic acid
Trigonelline hydrochloride	d-Desthiobiotin
(-)-Epinephrine	N-Acetyl-5-hydroxytryptamine
CDP-ethanolamine sodium salt hydrate	Itaconic acid
2,5-Dimethylpyrazine	Azelaic acid
L-Anserine nitrate salt	2-Oxoadipic acid
Biliverdin hydrochloride	2-Methylglutaric acid
sn-Glycerol 3-phosphate bis(cyclohexylammonium) salt	N-Acetyl-L-leucine
4-Hydroxy-L-phenylglycine	2',4'-Dihydroxyacetophenone
N-Acetyl-DL-serine	Methyl indole-3-acetate
3-Methylglutaric acid	(RS)-Mevalonic acid lithium salt
Tyramine	DL-4-Hydroxy-3-methoxymandelic acid
Hydrocortisone	Salicylamide
Melatonin	Pyridoxal hydrochloride
Kynurenic acid	Sebacic acid
3,5-Diiodo-L-tyrosine dihydrate	Pyrrole-2-carboxylic acid
Tryptamine	4-Coumarate
Lumichrome	Nonanoate
N-Acetyl-L-phenylalanine	Estradiol-17alpha

N-Acetyl-L-proline	Ursodeoxycholate
L-Tryptophanamide hydrochloride	Petroselinate
N-Methyltryptamine	Lithocholate
3-(2-Hydroxyethyl)indole	Elaidate
4-Hydroxy-3-methoxycinnamic acid	Myristate
Glycocholic acid hydrate	Rosmarinate
6-(METHYLTHIO)PURINE	Cortexolone
Sodium glycochenodeoxycholate	Palmitoleate
10-Hydroxydecanoic acid	Stearate
1,2-Didecanoyl-sn-glycero-3-phosphocholine	Palmitate
Dl-5-(1,2-dithiolan-3-yl)valeramide	Laurate
Indole-3-acetamide	Arachidate
Hippuric acid	Cortisone
3,5-Diiodo-L-thyronine	Heptadecanoate
3-Methoxytyramine hydrochloride	Omega-hydroxydodecanoate
Quinaldic acid	5,6 Dimethylbenzimidazole
2-Aminophenol	Cholate
Pimelic acid	-
3-Amino-4-hydroxybenzoic acid	-
Indole-3-acetic acid sodium salt	-
b-Nicotinamide adenine dinucleotide hydrate	-
L-Glutamine	-
Hypotaurine	-
Inosine 5'-monophosphate from Saccharomyces cerevisiae	-
L-Threonine	-
Purine	-
N-Acetylneuraminic acid	-
Cytidine	-
L-Serine	-
L-Citrulline	-
Taurine	-
Nicotinic acid	-
γ-Aminobutyric acid	-

Cytosine	-
L-Isoleucine	-
L-Glutamic acid	-
L-Ascorbic acid	-
N-Acetyl-D-glucosamine	-
Sarcosine	-
(-) Quinic acid	-
L-Methionine	-

Table 20: RPLC Target Screening identified compounds in both ionization modes

RPLC Positive	RPLC Negative
1-Aminocyclopropanecarboxylic acid	3-Hydroxyisovaleric acid
4-Imidazoleacrylic acid	Alpha-D-glucose
5-Aminovaleric acid	Arachidonic acid
Betaine	Aspartic acid
Dihydrouracil	Butyrate
Creatine	Caproic acid
L-Glutamine	cis-Aconitic acid
D-Mannosamine hydrochloride	citrate ion
DL-Normetanephrine hydrochloride	D-Glucuronic Acid
L-Isoleucine	DL-Lactate
Dopamine hydrochloride	Galactonic acid
L-Ornithine monohydrochloride	glyceric acid
L-Phenylalanine	Homovanillic acid
L-Valine	Isocitric acid

Histamine dihydrochloride	L-Glutamate
Methyl 4-aminobutyrate hydrochloride	Nonanoate
Nα-Acetyl-L-lysine	Oleate
Sarcosine	Petroselinate
Trigonelline hydrochloride	Succinic acid
L-Alanine	Sucrose
L-Arginine monohydrochloride	Tagatose
L-Carnitine hydrochloride	-
L-Glutamic acid	-
L-Homoserine	-
L-Methionine	-
L-Serine	-
L-Threonine	-
N-Acetylglycine	-
Xanthine	-
Taurine	-
N-Methyl-L-glutamic acid	-

Table 21: Metabolites Compound list for RPLC Suspect Screening.

Choline	Phenylacetylglycine
Pyruvate	Salicyluric acid
Hydroxyoctadecadienoic acid	1,2-Didecanoylglycerol
Linoleic acid	L-Glutamate
Indolelactic acid	3-Hydroxyisovaleric acid
Acetate	L-Tyrosine
Citrate ion	alpha-D-glucose
Urea	Selenocystine
Pantothenic acid	Maltotriose
Formic acid	Caproic acid
Orotate	DL-Lactate
Phosphocreatine	N-Acetylmannosamine
Trigonelline	1-Methyladenine
Fumarate	β-Carotene
Butyrate	Thyronine
Trimethylamine	Maltohexaose
Creatinine	Cyanocobalamin
Ornithine	Thioguanosine
Arginine	Threoninyl-Aspartate
Pyridoxal	Retinaldehyde
Fucose 1-phosphate	Retinoic acid
Isoleucyl-Hydroxyproline	2-cis,4-trans-Abscisic acid
Aspartic acid	α-Ketoglutaric acid
L-arabinose	Arabitol
Serylmethionine	Cholesterol
Biotin	Isomaltulose
Zeaxanthin	Mannose-6-phosphate
Karpoxanthin	Fucose
Mactraxanthin	Ergocalciferol
Secoisolariciresinol	Uridine
Pinoresinol	Pyrimidine
Lariciresinol	2-Dehydro-O-desmethylangolensin
Enterodiol	cis-5-Tetradecenoylcarnitine
N-oleoyl glutamine	Isocitric acid
Schisandrin	Riboflavin
Salinosporamide B	Muricholic acid
alpha-Galactosylceramide	1,2-Dipalmitoyl-sn-glycerol
Acetylcholine	Glucosamine
Glyceric acid	Histidinyl-Cysteine
Retinol	Isoleucylphenylalanine
Selenomethionine	N-Formylpiperidine
Succinoadenosine	Phenylalanylproline

3,3?,4?5-Tetrahydroxystilbene	3-Indoleacrylic acid
N-Methyl-3-aminomethylindole	5-Hydroxymethylfurfural
Ribonic acid	7-Methylguanine
Galactonic acid	Dexpanthenol
Clupanodonic acid	Diethyl hydrogen phosphate
Arachidonic acid	Leucylproline
Lactoferrin	Linoleoyl ethanolamide
Monobutyl phthalate	P-Cresol
Homovanillic acid	6-Hydroxyhexanoic acid
Cinnamic acid	Acrylic acid
Xylobiose	Benzoic Acid
Ricinoleic acid	Malic Acid
D-Glucuronic Acid	
cis-Aconitic acid	

Table 22: RPLC Suspect Screening identified compounds in both ionization modes

RPLC Positive	RPLC Negative
5-Hydroxymethylfurfural	3-Hydroxyisovaleric acid
7-Methylguanine	Alpha-D-glucose
Choline	Arachidonic acid
Cis-5-Tetradecenoylcarnitine	Aspartic acid
Creatinine	Butyrate
L-Tyrosine	Caproic acid
Monobutyl phthalate	cis-Aconitic acid
N-Formylpiperidine	citrate ion
Leucylproline	D-Glucuronic Acid
Pantothenic acid	DL-Lactate
Phenylacetylglycine	Galactonic acid
Phenylalanylproline	glyceric acid
N-Acetylmannosamine	Homovanillic acid
Pinoresinol	Isocitric acid

Riboflavin	L-Glutamate
Ribonic acid	linoleic acid
Salicyluric acid	Malic Acid
Salinosporamide B	Muricholic acid
Trigonelline	orotate
Trimethylamine	P-Cresol
Uridine	Pyridoxal
Xylobiose	pyruvate
Hydroxyoctadecadienoic acid	retinoic acid
indolelactic acid	Ricinoleic acid
L-arabinose	6-Hydroxyhexanoic acid
Glucosamine	Arabitol
Isoleucyl-Hydroxyproline	mannose-6-phosphate
-	α-ketoglutaric acid
-	Cinnamic acid
-	Clupanodonic acid
-	Diethyl hydrogen phosphate

REFERENCES

[1] Danezis G. P., Tsagkaris A. S., Camin F., Brusic V. and Georgiou C. A. 2016, Food authentication: Techniques, trends & emerging approaches *TrAC - Trends Anal. Chem.* **85** 123–32.

[2] A. Varnam J. P. S. 2001, *Milk and Milk Products: Technology, Chemistry and Microbiology* (New York: Aspen Publishers).

[3] Csapb-kiss Z., Stefler J., Martin T. G., Makray S. and Csap J. 1995 Kiss_1995.pdf **5** 403–15.

[4] Barron L. J. R., Aldai N., Virto M. and de Renobales M. 2017, Cheeses with protected land- and tradition-related labels: Traceability and authentication *Glob. Cheesemaking Technol. Cheese Qual. Charact.* 100–19.

[5] Dias C. and Mendes L. 2018, Protected Designation of Origin (PDO), Protected Geographical Indication (PGI) and Traditional Speciality Guaranteed (TSG): A bibiliometric analysis *Food Res. Int.* **103** 492–508.

[6] Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January2002. 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.

[7] Commission Implementing Regulation (EU) 2018/150 of 30 January 2018, amending Implementing Regulation (EU) 2016/1240 as regards methods for the analysis and quality evaluation of milk and milk products eligible for public intervention and aid for private storage.

[8] Regulation (EC) No 853/2004 of the European Parliament and of the Council of 29 April 2004 laying down specific hygiene rules for food of animal origin.

[9] Regulation (EU) 2017/625 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.

[10] Charlebois S., Schwab A., Henn R. and Huck C. W. 2016, Food fraud: An exploratory study for measuring consumer perception towards mislabeled food products and influence on self-authentication intentions *Trends Food Sci. Technol.* **50** 211–8.

[11] Qin C., Liu L., Wang Y., Leng T., Zhu M., Gan B., Xie J., Yu Q. and Chen Y. 2022, Advancement of omics techniques for chemical profile analysis and authentication of milk *Trends Food Sci. Technol.* **127** 114–28.

[12] Jia W., Dong X., Shi L., Chu X. and Chu X. 2020, Discrimination of Milk from Different Animal Species by a Foodomics Approach Based on High-Resolution Mass Spectrometry *J. Agric. Food Chem.* **68** 6638–45.

[13] Davies H. 2010, A role for “omics” technologies in food safety assessment *Food Control* **21** 1601–10.

[14] Valdés A., Álvarez-Rivera G., Socas-Rodríguez B., Herrero M., Ibáñez E. and Cifuentes A. 2022, Foodomics: Analytical Opportunities and Challenges *Anal. Chem.* **94** 366–81.

[15] Yun Z. and Geng F. 2022, Editorial overview: Foodomics focusing on food nutrition and health *Curr. Opin. Food Sci.* **48** 100950.

[16] Jariyasopit N., Khamsaeng S., Panya A., Vinaisuratern P., Metem P., Asawalertpanich W., Visessanguan W., Sirivatanauksorn V. and Khoomrung S. 2021, Quantitative analysis of nutrient metabolite compositions of retail cow’s milk and milk alternatives in Thailand using GC-MS *J. Food Compos. Anal.* **97** 103785.

[17] Suh J. H. 2022, Critical review: Metabolomics in dairy science – Evaluation of milk and milk product quality *Food Res. Int.* **154** 110984.

[18] Sen C., Ray P. R. and Bhattacharyya M. 2021, A critical review on metabolomic analysis of milk and milk products *Int. J. Dairy Technol.* **74** 17–31.

[19] Bruschetta G., Notti A., Lando G. and Ferlazzo A., 2021 A promising 31P NMR-multivariate analysis approach for the identification of milk

phosphorylated metabolites and for rapid authentication of milk samples *Biochem. Biophys. Reports* **27** 101087.

[20] Rocchetti G. and O'Callaghan T. F. 2021, Application of metabolomics to assess milk quality and traceability *Curr. Opin. Food Sci.* **40** 168–78.

[21] Plozza T., Craige Trenerry V. and Caridi D. 2012, The simultaneous determination of vitamins A, e and β-carotene in bovine milk by high performance liquid chromatography-ion trap mass spectrometry (HPLC-MS n) *Food Chem.* **134** 559–63.

[22] Lozano-Sánchez J., Borrás-Linares I., Sass-Kiss A. and Segura-Carretero A. 2018, Chromatographic Technique: High-Performance Liquid Chromatography (HPLC).

[23] Lee T. D. 2011, Introduction to Modern Liquid Chromatography, Third Edition vol 22.

[24] Luykx D. M. A. M. and van Ruth S. M. 2008, An overview of analytical methods for determining the geographical origin of food products *Food Chem.* **107** 897–911.

[25] Sanjay K. D. and Kumar H. D. 2012, Importance of RP-HPLC in analytical method development: A review *Int. J. Pharm. Sci. Res.* **3** 4626–33.

[26] Colin H. and Guiochon G. 1977, Introduction to reversed-phase high-performance liquid chromatography *J. Chromatogr. A* **141** 289–312.

[27] Trivedi D. K. and Iles R. K. 2014, Do not just do it, do it right: Urinary metabolomics -establishing clinically relevant baselines *Biomed. Chromatogr.* **28** 1491–501.

[28] D.A. Skoog, F.J. Holler S R C 2007, Principles of Instrumental Analysis.

[29] Kaklamanos G., Aprea E. and Theodoridis G. 2020, Mass spectrometry: principles and instrumentation (Elsevier Inc.).

[30] Hossain M. 2020, The Mass Spectrometer and Its Components.

[31] Kaklamanos G., Aprea E. and Theodoridis G. 2020 Mass spectrometry: principles and instrumentation (Elsevier Inc.).

[32] Ho C. S., Lam C. W. K., Chan M. H. M., Cheung R. C. K., Law L. K., Lit L. C. W., Ng K. F., Suen M. W. M. and Tai H. L. 2003, Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin. Biochem. Rev.* **24** 3–12.

[33] Banerjee S. and Mazumdar S. 2012, Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte *Int. J. Anal. Chem.* **2012** 1–40.

[34] Joyce J. R. and Richards D. S. 2011, Kinetic control of protonation in electrospray ionization *J. Am. Soc. Mass Spectrom.* **22** 360–8.

[35] Chen G., Fan M., Liu Y., Sun B., Liu M., Wu J., Li N. and Guo M. 2019, Advances in MS Based Strategies for Probing Ligand-Target Interactions: Focus on Soft Ionization Mass Spectrometric Techniques *Front. Chem.* **7** 1–17.

[36] G. Marshall A., T. Blakney G., Chen T., K. Kaiser N., M. McKenna A.l, P. Rodgers R., M. Ruddy B. and Xian F. 2013, Mass Resolution and Mass Accuracy: How Much Is Enough? *Mass Spectrom.* **2** S0009–S0009.

[37] Murray K. K. 2022, Resolution and Resolving Power in Mass Spectrometry *J. Am. Soc. Mass Spectrom.* **33** 2342–7.

[38] Brenton A. G. and Godfrey A. R. 2010, Accurate mass measurement: Terminology and treatment of data *J. Am. Soc. Mass Spectrom.* **21** 1821–35.

[39] Blichert-Toft J. 2018, *Hafnium.*

[40] Geer Wallace M. A. and McCord J. P. 2020, High-resolution mass spectrometry (Elsevier B.V.).

[41] Pitt J. J. 2009, Principles and Applications of Liquid ChromatographyMass Spectrometry in Clinical Biochemistry **30** 19–34.

[42] Tilvi S., Majik M. S. and Singh K. S. 2014, Mass spectrometry for determination of bioactive compounds vol. **65** (Elsevier B.V.).

[43] Rockwood A. L., Kushnir M. M. and Clarke N. J. 2018, Mass spectrometry

(Elsevier Inc.).

[44] Lacorte S. and Fernandez-Alba A. R. 2006, Time of flight mass spectrometry applied to the liquid chromatographic analysis of pesticides in water and food *Mass Spectrom. Rev.* **25** 866–80.

[45] Allen D. R. and McWhinney B. C. 2019, Quadrupole Time-of-Flight Mass Spectrometry: A Paradigm Shift in Toxicology Screening Applications *Clin. Biochem. Rev.* **40** 135–46.

[46] Utpott M., Rodrigues E., Rios A. de O., Mercali G. D. and Flôres S. H. 2022, Metabolomics: An analytical technique for food processing evaluation *Food Chem.* **366.**

[47] Bletsou A. A., Jeon J., Hollender J., Archontaki E. and Thomaidis N. S. 2015, Targeted and non-targeted liquid chromatography-mass spectrometric workflows for identification of transformation products of emerging pollutants in the aquatic environment *TrAC - Trends Anal. Chem.* **66** 32–44.

[48] Agüera A., Martínez-Piernas A. B. and Campos-Mañas M. C. 2017, Analytical Strategies Used in HRMS *Appl. High Resolut. Mass Spectrom. Food Saf. Pestic. Residue Anal.* 59–82.

[49] Schymanski E. L., Jeon J., Gulde R., Fenner K., Ruff M., Singer H. P. and Hollender J. 2014, Identifying small molecules via high resolution mass spectrometry: Communicating confidence *Environ. Sci. Technol.* **48** 2097–8.

[50] Milman B. L. 2015, General principles of identification by mass spectrometry *TrAC - Trends Anal. Chem.* **69** 24–33.

[51] Andre C. M. and Soukoulis C. 2020, Food quality assessed by chemometrics *Foods* **9** 2–5.

[52] Esteki M., Simal-Gandara J., Shahsavari Z., Zandbaaf S., Dashtaki E. and Vander Heyden Y. 2018, A review on the application of chromatographic methods, coupled to chemometrics, for food authentication *Food Control* **93** 165–82.

[53] Hohrenk L. L., Vosough M. and Schmidt T. C. 2019, Implementation of Chemometric Tools to Improve Data Mining and Prioritization in LC-HRMS for Nontarget Screening of Organic Micropollutants in Complex Water Matrixes *Anal. Chem.* **91** 9213–20.

[54] Ballabio D. and Consonni V. 2013, Classification tools in chemistry. Part 1: Linear models. PLS-DA *Anal. Methods* **5** 3790–8.

[55] Barker M. and Rayens W. 2003, Partial least squares for discrimination *J. Chemom.* **17** 166–73.

[56] Lasalvia M., Capozzi V. and Perna G. 2022, A Comparison of PCA-LDA and PLS-DA Techniques for Classification of Vibrational Spectra *Appl. Sci.* **12**.

[57] Mialon N., Roig B., Capodanno E. and Cadiere A. 2023, Untargeted metabolomic approaches in food authenticity: A review that showcases biomarkers *Food Chem.* **398** 133856.

[58] Valdés A., Álvarez-Rivera G., Sánchez-Martínez J. D. and Cifuentes A. 2023, Foodomics: LC and LC–MS-based omics strategies in food science and nutrition *Liquid Chromatography* (Elsevier) pp 647–81.

[59] Sadeghi A., Ebrahimi M., Hajinia F., Kharazmi M. S. and Jafari S. M. 2023, FoodOmics as a promising strategy to study the effects of sourdough on human health and nutrition, as well as product quality and safety; back to the future *Trends Food Sci. Technol.* **136** 24–47.

[60] Zhang A., Sun H., Wang P., Han Y. and Wang X. 2012, Modern analytical techniques in metabolomics analysis *Analyst* **137** 293–300.

[61] Rubert J., Zachariasova M. and Hajslova J. 2015, Advances in high-resolution mass spectrometry based on metabolomics studies for food – a review *Food Addit. Contam. - Part A Chem. Anal. Control. Expo. Risk Assess.* **32** 1685–708.

[62] Zhang R. and Jia W. 2022, Authenticity and traceability of goat milk: Molecular mechanism of β-carotene biotransformation and accessibility *Food Chem.* **388** 133073.

[63] Esslinger S., Riedl J. and Fauhl-Hassek C. 2014, Potential and limitations of non-targeted fingerprinting for authentication of food in official control *Food Res. Int.* **60** 189–204.

[64] Scano P., Murgia A., Pirisi F. M. and Caboni P. 2014, A gas chromatography-mass spectrometry-based metabolomic approach for the characterization of goat milk compared with cow milk *J. Dairy Sci.* **97** 6057–66.

[65] Caboni P., Murgia A., Porcu A., Manis C., Ibba I., Contu M. and Scano P. 2019, A metabolomics comparison between sheep's and goat's milk *Food Res. Int.* **119** 869–75.

[66] Yang Y., Zheng N., Zhao X., Zhang Y., Han R., Yang J., Zhao S., Li S., Guo T., Zang C. and Wang J. 2016, Metabolomic biomarkers identify differences in milk produced by Holstein cows and other minor dairy animals *J. Proteomics* **136** 174–82.

[67] Tian H., Zheng N., Wang W., Cheng J., Li S., Zhang Y. and Wang J. 2016, Integrated metabolomics study of the milk of heat-stressed lactating dairy cows *Sci. Rep.* **6** 1–10.

[68] Liu Z., Logan A., Cocks B. G. and Rochfort S. 2017, Seasonal variation of polar lipid content in bovine milk *Food Chem.* **237** 865–9.

[69] Pu J., Vinitchaikul P., Gu Z., Mao H. and Zhang F. 2021, The use of metabolomics to reveal differences in functional substances of milk whey of dairy buffaloes raised at different altitudes *Food Funct.* **12** 5440–50.

[70] Mazzei P. and Piccolo A. 2018, NMR-based metabolomics of water-buffalo milk after conventional or biological feeding *Chem. Biol. Technol. Agric.* **5** 1–10.

[71] Tsiafoulis C. G., Papaemmanouil C., Alivertis D., Tzamaloukas O., Miltiadou D., Balayssac S., Malet-Martino M. and Gerothanassis I. P. 2019, NMR-based metabolomics of the lipid fraction of organic and conventional bovine milk *Molecules* **24** 1–18.

[72] Ashokan M., Ramesha K. P., Hallur S., Karthikkeyan G., Rana E.,

Azharuddin N., Raj S. R., Jeyakumar S., Kumaresan A., Kataktalware M. A., Das D. N. and Keshava Prasad T. S. 2021, Differences in milk metabolites in Malnad Gidda (Bos indicus) cows reared under pasture-based feeding system *Sci. Rep.* **11** 1–13.

[73] Rocchetti G, Gallo A, Nocetti M, Lucini L and Masoero F 2020 Milk metabolomics based on ultra-high-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry to discriminate different cows feeding regimens *Food Res. Int.* **134** 109279.

[74] Riuzzi G, Tata A, Massaro A, Bisutti V, Lanza I, Contiero B, Bragolusi M, Miano B, Negro A, Gottardo F, Piro R and Segato S 2021 Authentication of forage-based milk by mid-level data fusion of (+/−) DART-HRMS signatures *Int. Dairy J.* **112** 104859.

[75] Shetty S A, Young M F, Taneja S and Rangiah K 2020 Quantification of B-vitamins from different fresh milk samples using ultra-high performance liquid chromatography mass spectrometry/selected reaction monitoring methods *J. Chromatogr. A* **1609** 460452.

[76] Yuan X, Shi W, Jiang J, Li Z, Fu P, Yang C, Rehman S ur, Pauciullo A, Liu Q and Shi D 2022 Comparative metabolomics analysis of milk components between Italian Mediterranean buffaloes and Chinese Holstein cows based on LC-MS/MS technology *PLoS One* **17** 1–13.

[77] Web page: https://www.directindustry.com/prod/bruker-daltonics/product-30029-991983.html.

[78] Granato D., Putnik P., Kovačević D. B., Santos J. S., Calado V., Rocha R. S., Cruz A. G. Da., Jarvis B., Rodionova O. Y. and Pomerantsev A. 2018, Trends in Chemometrics: Food Authentication, Microbiology, and Effects of Processing *Compr. Rev. Food Sci. Food Saf.* **17** 663–77.

[79] Tzora A., Nelli A., Kritikou A. S., Katsarou D., Giannenas I., Lagkouvardos I., Thomaidis N. S. and Skoufos I. 2022, The “Crosstalk” between Microbiota and Metabolomic Profile of Kefalograviera Cheese after the Innovative Feeding Strategy of Dairy Sheep by Omega-3 Fatty Acids *Foods* **11.**

[80] Breaux J., Jones K. and Boulas P. 2003, Understanding and Implementing Efficient Analytical Methods Development and Validation *Pharm. Technol.* 6–13.

[81] Fao 2017, Guidelines on performance criteria for methods of analysis for the determination of pesticide residues in food and feed *Codex Aliment.* 1–13.

[82] European Parliament and the Council of the European Union 2002 96/23/EC COMMISSION DECISION of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results (notified under document number C(2002) 3044)(Text withEEA relevance) (2002/657/EC) *Off. J. Eur. communities* 8–36.

[83] Matuszewski B. K., Constanzer M. L. and Chavez-Eng C. M. 2003, Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS *Anal. Chem.* **75** 3019–30.