NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

SCHOOL OF SCIENCES

DEPARTMENT OF CHEMISTRY

Master's Degree Program
«Analytical Chemistry – Quality Assurance»

Master Thesis

Chemical characterization of Cypriot carob syrup using NMR & High-Resolution Mass Spectrometry

Agapiou Constantinos
CHEMIST

ATHENS

JUNE 2023

Master Thesis

Chemical characterization of Cypriot carob syrup using NMR & High-Resolution Mass Spectrometry

CONSTANTINOS AGAPIOU

A.M.: 7111112100003

Supervising Professor:
Dr. Nikolaos S. Thomaidis**,** Professor, Department of Chemistry, NKUA

Consultative Committee

Dr. Nikolaos S. Thomaidis, Professor, Department of Chemistry, NKUA

Dr. Thomas Mavromoustakos, Professor, Department of Chemistry, NKUA

Dr. Evangelos Gikas, Professor, Department of Chemistry, NKUA

DEFENDING DATE 23/06/2023

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χημικός χαρακτηρισμός του Κυπριακού χαρουπόμελου
με χρήση Πυρηνικού Μαγνητικού Συντονισμού & Υψηλής Διακριτικής Ικανότητας Φασματομετρίας Μαζών

ΑΓΑΠΙΟΥ ΚΩΝΣΤΑΝΤΙΝΟΣ

Α.Μ.: 7111112100003

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:
Δρ. Νικόλαος Σ. Θωμαΐδης**,** Καθηγητής, Τμήμα Χημείας, ΕΚΠΑ

ΤΡΙΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Δρ. Νικόλαος Σ. Θωμαΐδης**,** Καθηγητής, Τμήμα Χημείας, ΕΚΠΑ

Δρ. Θωμάς Μαυρομούστακος, Καθηγητής, Τμήμα Χημείας, ΕΚΠΑ

Δρ. Ευάγγελος Γκίκας, Καθηγητής, Τμήμα Χημείας, ΕΚΠΑ

ΗΜΕΡΟΜΗΝΙΑ ΕΞΕΤΑΣΗΣ 23/06/2023

Scope

The primary aim of my MSc Thesis is to chemically characterize Cyprus carob syrup, a traditional candidate good for Protected Geographical Indication (PGI), through the detection of metabolites. Carob has been one of Cyprus' primary sources of income since ancient times, and as a result it has been referred to as the "black gold" of the island. As a traditional product with increased commercial potential, it appears to meet consumers' growing demand for food that has both nutritional and pharmaceutical benefits for their health, such as anticancer, antioxidant, and antiaging properties, while also being a distinctive component of the Cypriot flora and the island's history. This highlights both its unique nutritional worth and its medicinal potential, emphasizing the above-mentioned goal.

A traditional product also benefits from standardization in the manufacturing process (e.g., temperature, extraction time, geographic areas, country, etc.). Thus, a secondary goal of this research is to advance knowledge of the importance of carob to Cyprus so that more steps can be taken to restore and surpass pre-invasion production levels. The goal is to make use of the rising carob production and demand by using a cutting-edge scientific methodology to transform it into novel food and pharmaceutical goods that will be pushed on the global market.

There are not many studies on the metabolomics of carob syrup, and this emphasizes the novelty and importance of the present Thesis. Towards this, two state-of-the-art advanced instruments, High-Resolution Mass Spectrometry (HRMS) and NMR, were used to highlight the metabolites in carob syrup, further comprehend the crucial elements for the production process, and distinguish commercial from traditional products.

ΠΕΡΙΛΗΨΗ

Σε μια εποχή που το αυθεντικό τρόφιμο αναζητείται και συνεχώς εμφανίζονται συμπληρώματα διατροφής, τα χαρούπια και κατ'επέκταση το χαρουπόμελο παρουσιάζουν μεγάλη εμπορική δυνατότητα. Τα τελευταία χρόνια παρατηρείται μια συνεχής αυξανόμενη ζήτηση σχετικών προϊόντων χαρουπιού από το αγοραστικό κοινό τόσο στην Κύπρο, όσο και στο εξωτερικό. Το χαρουπόμελο είναι άρρηκτα συνδεδεμένο με την παράδοσή της Κύπρου, παρέχοντας μία μοναδική διατροφική αξία και παράλληλη φαρμακολογική δράση. Στη συγκεκριμένη διπλωματική εργασία, μελετήθηκαν διάφορα χαρουπόμελα για την ανάδειξη των ιδιαίτερων χημικών χαρακτηριστικών τους. Συγκεκριμένα, εμπορικά (n=12) και παραδοσιακά (n=23) δείγματα χαρουπόμελων από την Κύπρο, Αυστραλία (n=1), Σλοβενία (n=1) και Κροατία (n=1) αναλύθηκαν με τη χρήση των αναλυτικών συστημάτων Υγρής χρωματογαφίας υπέρ-υψηλής απόδοσης συζευγμένη με φασματομετρία μάζας με αναλυτή μάζας τετράπολο - χρόνου πτήσης (Ultra-High Performance Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry, UHPLC-QTOF-ESI-MS) και υψηλής ακρίβειας Πυρηνικό Μαγνητικό Συντονισμό Nuclear Magnetic Resonance (NMR), για την εύρεση των χημικών μεταβολιτών τους. Συνολικά ανιχνεύτηκαν 48 ενώσεις, εκ των οποίων ποσοτικοποιήθηκαν οι 21, για τις οποίες υπήρχαν διαθέσιμα πρότυπα στο εργαστήριο. Οι ενώσεις-μεταβολίτες που ταυτοποιήθηκαν και εμφανίστηκαν σε όλα τα δείγματα (n=38) ανήκουν σε 7 μεγάλες κατηγορίες: α) αμινοξέα (n=13), β) λιπαρά ακύλια/λιπαρά οξέα και συζυγή (n=6), γ) οργανοοξυγονούχες ενώσεις/υδατάνθρακες και συζυγείς υδατάνθρακες (n=9), δ) φλαβονόλες και φλαβονοειδή (n=7), ε) οργανικά οξέα (n=9), στ) νουκλεοτίδια και παράγωγα νουκλεοτιδίων (n=3) και ζ) αλκαλοειδή (n=1). Η πιο πλούσια χημική κατηγορία είναι αυτή των αμινοξέων με ποσοστό 27%, ακολουθούμενη από τις οργανοοξυγονούχες ενώσεις/υδατάνθρακες και συζυγείς υδατάνθρακες και των οργανικών οξέων με ποσοστό 19%. Επιπλέον, το εμπορικό χαρουπόμελο είναι πλούσιο σε αλανίνη (μέση τιμή συγκέντρωσης=16,2 mg kg^{-1}) και σε D-φρουκτόζη (μέση τιμή συγκέντρωσης=12, 4 mg kg^{-1}), ενώ τα αντίστοιχα παραδοσιακά σε αλανίνη (μέση τιμή συγκέντρωσης=14,7 mg kg^{-1}) και σε L-σερίνη (μέση τιμή συγκέντρωσης=10,8 mg kg^{-1}). Το NMR, αλληλοσυμπλήρωσε τα αποτελέσματα της Φασματομετρίας Μάζας και κυρίως ανέδειξε την ιδιαιτερότητα της αρωματικής και αλειφατικής περιοχής και τη μεγάλη αφθονία των σακχάρων (σορβιτόλη, D-ξυλόζη, D-γλυκόζη, D-λακτόζη, σουκρόζη). Η αύξηση της θερμοκρασίας (20-100°C) κατά την παρασκευή των παραδοσιακών χαρουπόμελων οδήγησε στην αύξηση

συγκεγκριμένων μεταβολιτών. Η περαιτέρω χημειομετρική επεξεργασία των αποτελεσμάτων του NMR, μετά την αφαίρεση των σακχάρων, επέτρεψε το διαχωρισμό των εμπορικών από τα παραδοσιακά χαρουπόμελα. Αυτό οδήγησε στο συμπέρασμα ότι η αρωματική και αλειφατική περιοχή μπορούν να δώσουν σημαντικές πληροφορίες ως προς την αυθεντικότητα του προϊόντος. Συνοψίζοντας, ο χημικός χαρακτηρισμός του Κυπριακού χαρουπόμελου, μέσω των αναλυτικών συστημάτων UHPLC-QTOF-ESI/MS, σε συνδυασμό με το NMR, συμβάλλει στην άμεση καταγραφή και ποσοτικοποίηση των μεταβολιτών του, καθώς και στην περαιτέρω κατανόηση της ιδιαίτερης διατροφικής του αξίας.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Ενόργανη χημική ανάλυση-Τρόφιμα

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: χαρούπι, χαρουπόμελο, παραδοσιακό προϊόν, μεταβολίτες, Υγρο-χρωματογαφία, Φασματομετρία Μάζας υψηλής διακριτικής ικανότητας, πυρηνικός μαγνητικός συντονισμός, μεταβολομική τροφίμων

ABSTRACT

In an era where authentic food is sought after and dietary supplements and numerous chemical drugs constantly appear, carob and, by extension, carob syrup present significant commercial potential. In recent years, there has been a continuous increasing demand for carob-based products from the consumer market in Cyprus and abroad. Carob syrup is inextricably linked with the tradition of Cyprus, providing unique nutritional value and concurrent pharmacological action. In this thesis, various carob syrups were studied to highlight their distinctive chemical characteristics. Specifically, commercial (n=12) and traditional (n=23) carob syrup samples from Cyprus, Australia (n=1), Slovenia (n=1), and Croatia (n=1) were analyzed using the Ultra-High Performance Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry (UHPLC-QTOF-ESI-MS) and high resolution Nuclear Magnetic Resonance (NMR) analytical systems to identify their chemical metabolites. In total, more than 48 compounds were detected, out of which 21 were quantified, where their analytical standards were available in the laboratory. The identified compounds – metabolites that where present in all samples (n=38) belonged to 7 major categories: a) amino acids (n=13), b) fatty acyls/fatty acids and conjugates (n=6), c) organooxygen compounds/carbohydrates and carbohydrate conjugates (n=9), d) flavonols and flavonoids (n=7), e) organic acids (n=9), f) nucleotides and nucleotide derivatives (n=3), and g) alkaloids (n=1). The richest chemical category was that of amino acids, accounting for 27%, followed by organooxygen compounds/carbohydrates and carbohydrate conjugates and organic acids with 19%. Additionally, commercial carob syrup was rich in alanine (mean concentration=16.2 mg kg^{-1}) and D-glucose (mean concentration=12.4 mg kg^{-1}), while the corresponding traditional syrup showed higher levels of alanine (mean concentration=14.7 mg kg^{-1}) and L-serine (mean concentration=10.8 mg kg^{-1}). NMR complemented the results of Mass Spectrometry and primarily highlighted the distinctiveness of the aromatic and aliphatic region, as well as the abundance of sugars (sorbitol, D-xylose, D-glucose, D-lactose, sucrose). The increase in temperature (20-100°C) during the preparation of traditional carob syrups led to an increase in specific metabolites. Further chemometric processing of the NMR results, after sugar removal, allowed for the differentiation between commercial and traditional carob syrups. This led to the conclusion that the aromatic and aliphatic regions can provide significant information regarding the authenticity of the

product. In summary, the chemical characterization of Cypriot carob syrup through the UHPLC-QTOF-ESI/MS analytical systems, in combination with NMR, contributes to the direct recording and quantification of its metabolites, as well as a better understanding of its unique nutritional value.

SUBJECT AREA: Chemical Analysis-Foods

KEYWORDS: carob, carob syrup, traditional product, metabolites, Liquid Chromatography (LC), High Resolution Mass Spectrometry (HRMS), Nuclear Magnetic Resonance (NMR), foodomics

ACKNOWLEDGMENTS

I would like to sincerely thank Dr. Nikolaos Thomaidis, Professor at the National and Kapodistrian University of Athens (NKUA), who accepted me in the Laboratory of Analytical Chemistry and for all his valuable advice and scientific guidance. Furthermore, I would like to thank Dr. Thomas Mavromoustakos, Professor of Organic Chemistry and Dr. Gikas Evangelos, Professor of Analytical Chemistry, for their valuable help throughout the experimental preparation and implementation of current MSc thesis, as well as for the NMR and UHPLC-ESI-Q-TOF/MS access and training, respectively. I thank them for their time, as well as for their patience and willingness to solve every difficulty I faced. Also, I would like to express a big thank to Dr. Anthi Panara for her valuable help and for teaching me the use of HRMS, as well as to the other members of the Trace analysis and mass spectrometry (TrAMS) research team. Additionally, I feel deeply grateful to Dr. Antigoni Cheilari for transmitting on to me her knowledge in the field of NMR and statistical techniques and her valuable advice. Finally, I would like to say a big thank you to my parents, Agapios and Panagiota, who are always by my side, support me and give me strength to continue.

Table of Contents

Figures

Tables

1. Chapter 1
Introduction

Less known Mediterranean fruit species, such as carobs, have gained much more attention recently. The reason is the need for foodstuff having high content of bioactive compounds, such as flavonoids, phenolics, anthocyanins, phenolic acids, as well as nutritional compounds such as sugars, essential oils, carotenoids, vitamins, and minerals. Moreover, there they should present distinct flavor and taste, excellent medicinal value, and health care functions [1].

Carob seems to fully satisfy the above recent consumer's needs, as well as the following sustainable development goals (SDGs) of the Food and Agriculture Organization of the United Nations (FAO), which highlights the role of sustainable land use in mitigating climate change and promoting food security and nutrition [2] (Figure 1):

- SDG1: Eradicate extreme poverty
- SDG2: End poverty
- SDG3: Achieve good health and well-being
- SDG5: Empower women in economic life
- SDG12: Responsible consumption and production
- SDG15: Protecting, restoring, and promoting the sustainable use of terrestrial ecosystems and reversing land degradation
- SDG17: Partnerships to achieve the goal

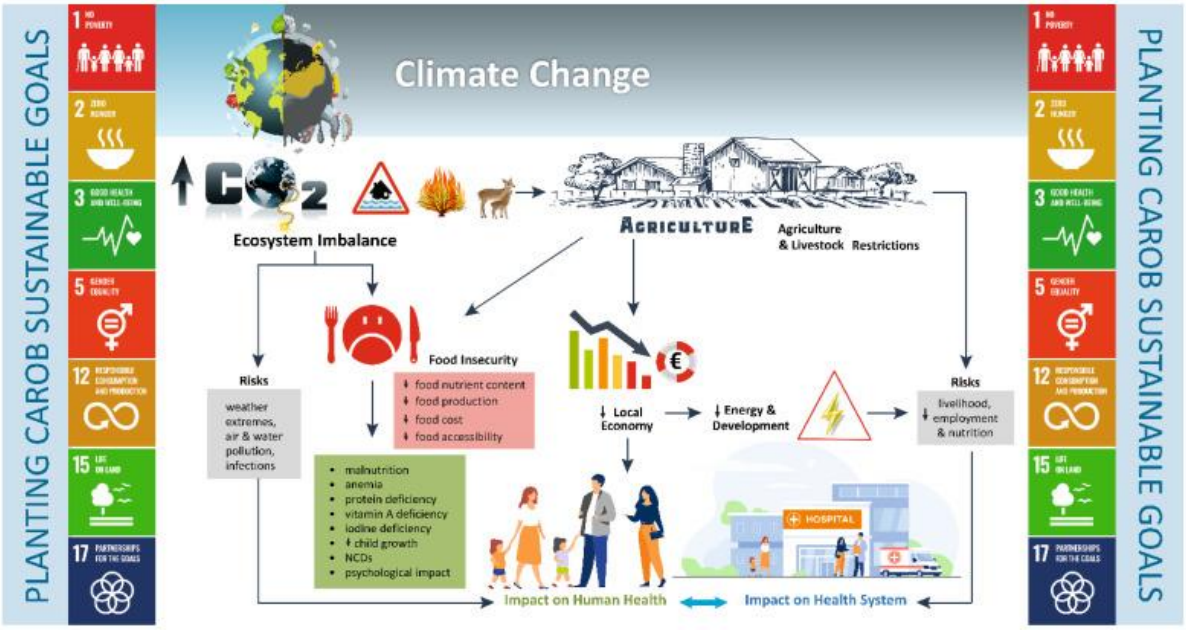


Figure 1. Interplay between climate change, food security, nutrition, and human health. Planting carob trees might be a promising course of action to achieve the following sustainable development goals (SDGs) across the Mediterranean area [2].

In addition, carob's recognized nutritional and medicinal value next to its unique agriculture importance is associated with an array of social, economic, and cultural activities [1].

1.1 Carobs ecology

The carob tree (*Ceratonia siliqua L*.) belongs to the family of *Leguminosae*, in the genus of Ceratonia (Figure 2). The thermophilus nature of the tree allows it to thrive in environments with mild and drought climate conditions, such as in most Mediterranean countries: Greece, Italy, Spain, Portugal, Cyprus, Morocco, Turkey, Lebanon and Algeria. Its official name, *'Ceratonia siliqua'* (*Κερατόνια η κερατέα*), derives from the Greek word for horn (*κέρας*) and the Latin '*siliqua*' (pod, carob), denoting the shape and hardness of the horn [3].

In the past, in Greece, especially during the years of famine and the period of occupation, it proved an extremely valuable food. Together with raisins and threpsins, replaced sugar, which was a hard-to-find and very expensive. The carob is a hardy and long-lived tree, which has the ability to sprout, with a height of up to 10 meters. Its trunk is rough with thick bark, it has spreading branches, which form an umbrella-shaped gum, like that of the olive. Its leaves consist of 2-5 pairs of shiny leaflets with a leathery texture. Its male, female or sometimes hermaphrodite flowers are small, greenish with red hues and grow

in short-stemmed clusters, on the buds of the branches. Males and hermaphrodites have a characteristic heavy smell, similar to that of human semen, which explains its medicinal use in infertility problems [1].

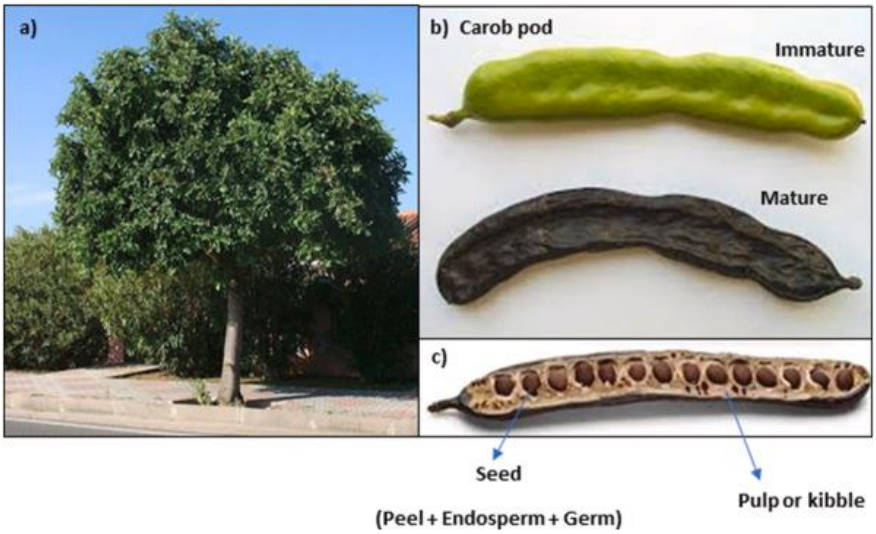


Figure 2. Carob tree (Ceratonia siliqua L.), b) Immature and mature carob pod, c) Carob pod constituents [4].

The carob is an evergreen, long-lived and fruitful tree. Fruit color and other organoleptic characteristics are evidence of fruit's maturity. Due to its slow growth, it bears fruit fully between its 8th and 10th year, while it has a productive life of 80-100 years. It is worth mentioning that a productive tree in its 6 years of life can yield 2.25 kg, while when it reaches 12 years it can yield up to 45 kg of carobs. The average production is 20 kg, with a maximum production of 70 kg. The ripening of the carob fruit takes place almost half a year and maturity is judged by the color of the carob pod (from green bean to dark brown with hard texture pulp) and the respective dehydration. Its flowering begins in October and the fruits ripen in August, when their collection begins, until the end of September. Carob is a tree with limited soil requirements, while it can withstand temperatures even above 40 °C. Therefore, it is an ideal species for the climate of Mediterranean countries as Cyprus [5]. It thrives in various types of soil, apart from very clayey and wet soils, it even thrives in rocky, dry, and sloping soils provided they are deep and light so that they can be penetrated by its root system. This is because it has a deep root system, which allows to finding water at greater depths, but also the fact that it contains D-Pinitol, which protects it from lack of water and high levels of salinity [3]. Its leaves and bark present tanning properties, but the unripe fruits have even more, as well as dyeing properties and

were used in the past to dye fabrics. Its wood provides excellent quality charcoal, while its heartwood is used in furniture, wood carving, turning and barrel making.

1.2 Worldwide production

For the period 2013-2016, according to the Food and Agriculture Organization of the United Nations (FAO) the average production of carob pod in the world was estimated from 113,000 to 166,000 tons per year. However, this number is likely underestimated since the nine top producing countries (considering national data inventories) exceeded 170,000 tons in 2016 and 2017. Based on the above results, Spain and Portugal come into first places. They produced around 44,000 tons per year for 2013–2018, whereas Italy produced approximately 29,000 tons per year for 2013–2019, coming into third place, as presented in Table 1. In the mid-20th century, a dramatic decline in total carob production was observed (72% from about 600,000 tons to 170,000 tons per year. This is mainly due to the decrease in Spanish carobs in 1950-52).

The cultivation and production of carob has been neglected and underestimated. However, in recent years, it has made a resurgence due to its nutritional composition and industrial value, being a rich source of minerals and fiber combined with low content in lipids).

It is worth noting that concerning of Cyprus there was an increase in the yield of crops during the period 2009 to 2019, from 1.68 tons to 8.88 tons per hectare (YSTAT, 2021) [6].

Table 1. Carob production (tons) of the nine major producing countries [7].

Year	Spain	Portugal	Italy	Morocco	Turkey	Greece	Cyprus	Lebanon	Algeria	Sum
2013	38,900	42,841	9,445	22,109	14,261	12,026	6,178	2,261	3,053	151,074
2014	60,400	64,895	31,486	22,169	13,985	13,473	11,034	2,922	3,655	224,019
2015	36,900	34,398	31,522	21,916	12,851	13,609	7,413	3,437	4,624	166,670
2016	39,600	40,087	28,925	21,666	13,405	13,163	6,369	4,186	3,257	170,658
2017	36,400	41,329	28,910	21,611	15,016	12,287	7,475	4,431	4,042	171,50
2018	53,500	41,734	36,951	21,556	15,506	–	6,616	4,376	2,880	183,119

2019	–	–	35,936	21,501	16,256	–	9,295	3,511	3,526	90,025
Average	44,283	44,214	29,025	21,790	14,469	12,912	7,769	3,589	3,577	

1.3 Historical data and the case study of Cyprus

The fruit is a flattened, leathery-textured, and sweet-tasting pod. Its shape resembles a goat's horn, hence the name "wood horn", while the name carob comes from the Arabic name of the tree "Kharrub".

Carob seeds have the property of being uniform in size and keeping their weight unchanged, regardless of weather conditions (0.205 g). Therefore, Venetian goldsmiths and silversmiths used them as a unit of weight for measurement of the weight of precious stones and metals. Hence, the name "carat", from the horn, i.e., the seed of the carob [8].

In general, there are 2 varieties:

- the wild, which has brown, short and woody fruits
- the tame, which has long, wide, black, and more sugary fruits.

Specifically, in Cyprus, there are several varieties, which are known by local names, depending mainly on the characteristics of the tree or fruit: *Arkoteracia, Pastoteracia, Arsenitzi, Apostolitzi, Hemeri, Kotsina, Komboti*, etc.

The carob has been cultivated in the Cyprus region for almost four thousand years. The island has been closely associated with the promotion of the carob in Mediterranean civilization, therefore we assume that carobs have existed there since prehistoric times. The carob was also characteristically called the "black gold of Cyprus", as it was the product with the largest agricultural exports of the island, and for some villages, it was the main agricultural activity and the main source of income. Magda Ohnefalsch-Richter, German archaeologist, who lived in Cyprus, between 1894 and 1912, commented that: "without this gift of nature, the island could have declared bankruptcy a long time ago". [9]. Throughout Cyprus history, carobs have been exported in many Middle East countries, as well as in countries that occupied the island in the past, such as Italy (Venice) and England (Figure 3).

Cyprus is world-wide known for the production of best quality carobs and carob products. Cyprus limestone soils favour the growth of carob trees which results to its particularly

attractive taste. The demand for Cyprus carobs in the international market is remarkably high as they are particularly rich in natural sugar content.

Figure 3. A stamp depicting the Queen of England and the importance of carob.

1.4 Beneficial effects of carob and its products

The pods are rich in carbohydrates and contain more than 55% sugars, such as glucose, sucrose, fructose and more than 30-40% fiber, among which about 18% is cellulose and hemicellulose. In addition, they contain 4% protein and a very low-fat content of 0.21%-0.23%. Its seeds are also rich in proteins, tannins and cellulose. From them, carob bean gum (E410) is extracted, a natural gum, which is widely used as a thickening agent in food technology and industry, with many uses in: ice creams, soups, sauces, cheeses, canned meats, jellies, products, bakery products, etc. Other technical applications include cosmetics, pharmaceuticals, paints, adhesives, etc. Carobs are also rich in antioxidant compounds. The main classes of phenolic compounds found in carobs are phenolic acids, flavonoids and gallotannins. In the first category, gallic acid dominates, a substance with anti-cancer properties, which can only be found in higher amounts than in carobs in cloves and chestnuts. In the second place, quercetin dominates, which also has anti-cancer properties, and myricetin. The last category includes the tannins, that are responsible for the astringent taste. Additionally, they contain important amino acids, such as histidine (His), isoleucine (lle), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), threonine (Thr), tryptophan (Trp), valine (Val), arginine (Arg) etc. Finally, they contain minerals, such as Na, K and large amounts of Ca, Fe, Mg, Mn, P, Zn, Cu, etc., as well as vitamins i.e.: E, D, C, B6, niacin, folic acid, etc. [10].

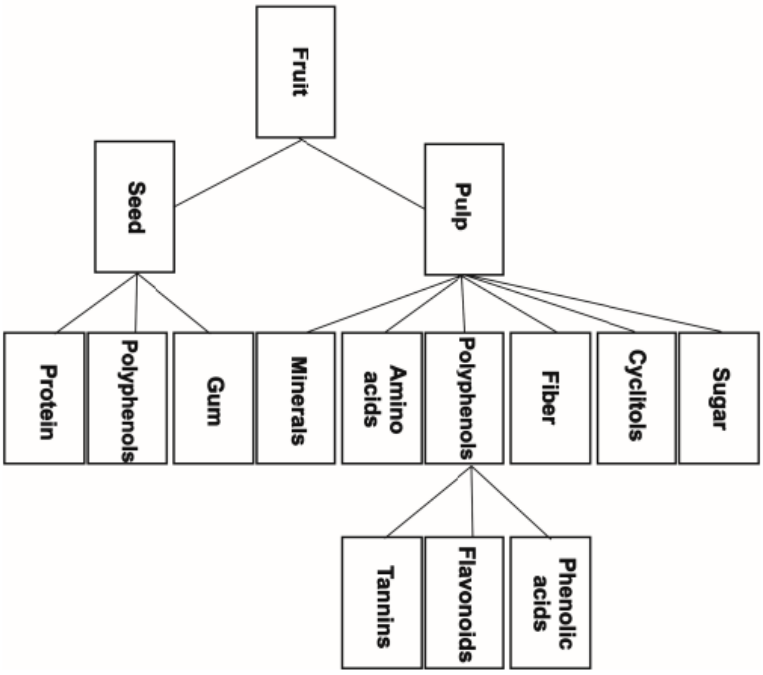


Figure 4. Main chemical constituents in carob pulp and seed with nutritional and health-promoting properties [11].

Carob has a great potential to be used in the food industry, not only for its health benefits and also of its unique and enduring aroma, which remains present even after undergoing processing. Acids, esters and aldehydes/ketones emitted from carob fruit and powder, are biogenic volatile organic compounds (VOCs) that contribute to the plant growth, breeding, defense and their wide presence in Cypriot carob strongly contributes to its resilient aroma. On the other side, for the food industry's diversification and innovative development, brand new technological modernization, and product-line expansion strategies are needed. In their study, Lobanov et al. (2018) examined the financial implications of incorporating vegetal ingredients, specifically carob, in the production of flour-based functional foods. The researchers focused on creating a carob bun with 4% carob content and evaluated its impact on the nutritional composition. The inclusion of carob led to significant changes in the bun's nutritional profile. It resulted in increased protein, fat, and fiber content, while reducing the amount of carbohydrates compared to a lecithin bun without carob. These findings suggest that carob can serve as a valuable raw material for enhancing the nutritional value of gluten-free bread and flours by enriching them with vitamins, minerals, and proteins [1].

The benefits on the prevention and treatment of chronic diseases of the carob fruit and its products, have been in the scope of multiple studies. For example, D-pinitol has several properties, such as anti-cancer, insulin-mimetic, antioxidant, hepatoprotective, preventive of osteoporosis, anti-aging, preventive and ameliorating of Alzheimer's disease.

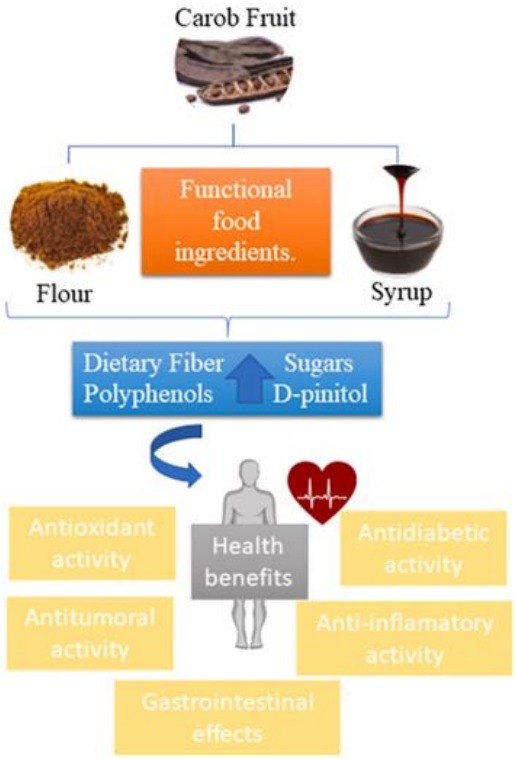


Figure 5. Carob beneficial health effects [1].

Carobs have also been used in traditional medicine. For example, in the "Iatrosofikon", which was written at the beginning of the 19th century by Mitrophanis, custodian of the Holy Monastery of Machairas (Cyprus), he mentioned that "*teratsia*" (Carob tree) together with green or dry roses, sumac and sweet wine can cure bleeding. On the other hand, according to testimonies from the last century, carob syrup was given to children to fight cough, diarrhea, and gastroenteritis [12].

The carob chemical components and their biological properties associated with health benefits are presented in more detail at the Table 2.

Table 2. Carob chemical constituents and their biological properties associated with health benefits [11].

Group of Chemical Constituents/ Individual Substances	Biological Evaluation of Constituents/ Disease	Carob Parts/ Fraction
LBG/galactomannan	Gastrointestinal effects	Seed endosperm
D-Pinitol	Anti-diabetic activity	Carob pulp
Soluble and Insoluble Dietary Fiber Polyphenols/Gallic acid, Gallotannins, Flavonol Glycosides	Glycemic control, Enhanced lipid metabolism, Lowers total and LDL cholesterol	Carob Pulp
Insoluble Dietary Fiber Polyphenols/Tannins, Cellulose, Semicellulose, Lignin, Pectin	Cholesterol metabolism, Enhances lipid oxidation, Lowers postprandial acylated ghrelin	Carob Fiber
Polyphenols/ Gallic acid, Catechin, Myricetin rhamnoside, Eriodictyol glucoside, Quercetin glucoside, Quercetin rhamnoside	Anticancer effects	Carob Fiber
Polyphenols—Alkaloids/(+)-Catechin; Gentisic acid; Chlorogenic acid; Catechol; Ferulic acid; Gallic acid; Myricetin; Methyl gallate; Quercetin; Rutin; Syringic acid; Theophylline; Vanillin	Cytotoxic activities	Germ Flour Extracts (seed)
Fiber	Nutritional utilization, Induction of lipodemia	Carob Fiber
Fiber	Hyperlipidemia effects	Carob fiber
Tannins—Polyphenols	Anti-diarrheal effects	Carob pod
Tannins—Pectin	Anti-diarrheal effects	Carob bean juice

Carob syrup is made from the carob pods after grinding. In particular, the carobs are boiled with water and the extracted syrup continues boiling for approximately four hours. When most of the water is evaporated, a golden-black thick syrup is produced with a distinctive caramel-like flavour and aroma. [13].

Carob syrup has been used in Cypriot cuisine for centuries as a natural sweetener and flavouring agent for a variety of dishes, including desserts, sauces, and marinades. It is often used as a substitute for honey or molasses and can be drizzled over pancakes, waffles, ice cream, or added to coffee, tea, cocktails.

In addition to its culinary uses, carob syrup is also believed to have medicinal properties and sometimes was used as a natural remedy for coughs, sore throats, and digestive issues. It is high in antioxidants, fiber, and minerals such as calcium, magnesium, and potassium making it a healthy alternative to processed sugars.

Today, carob syrup is still a popular and beloved ingredient in Cypriot cuisine, and it is often used in traditional dishes such as pasteli (a sweet sesame and honey bar) and "*daktyla"* (a sweet pastry filled with ground almonds and cinnamon). It is also exported and enjoyed in many other countries as a unique and flavourful sweetener.

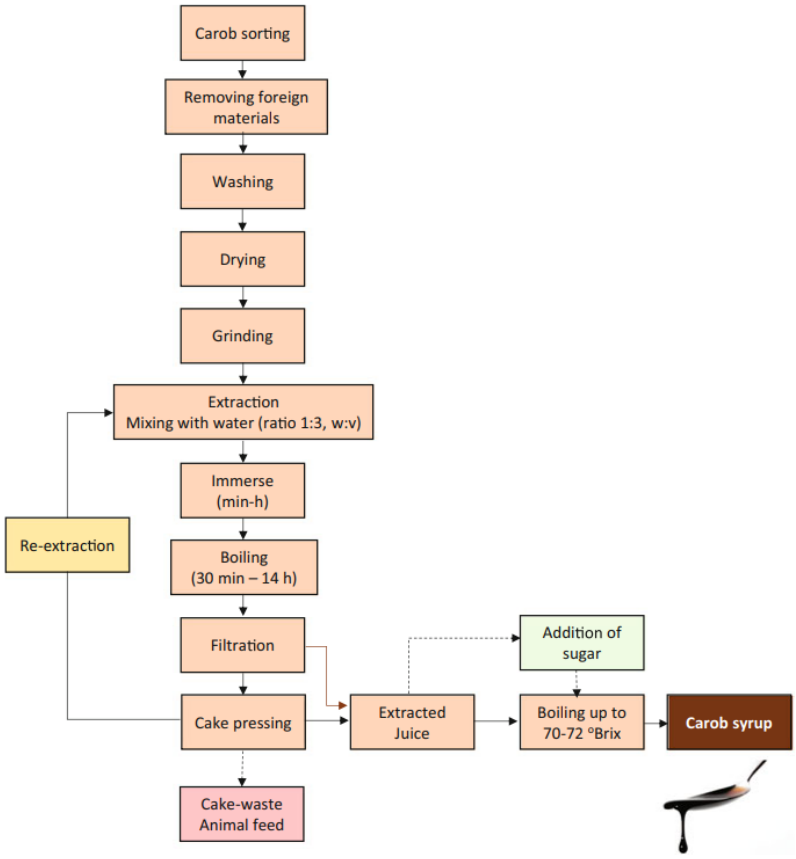


Figure 6. Industrial and homemade production of carob syrup [14].

Aziz and Hicham (2014) examined the optimal conditions to produce carob syrup from different Moroccan carob populations. They used an experimental design where the effect of three parameters (extraction temperature, extraction time and water to pulp ratio) on syrup yields were examined. As a result, the optimal conditions were found to be that of extraction temperature 43.45 °C, extraction time 2.40 h, and extraction solvent (water) to pulp ratio 2.27 (v/w). The syrup yield from the different regions varied between 28.76 and 37.22 g/100 g (% w/w) [1].

The hydroxymethylfurfural (HMF) content of the carob syrup samples must not exceed the limit value set by the Food and Drug Administration (FDA), which 40 mg/kg is the maximum permissible level in the EU for table syrup [15].

According to Dhaouadi et al. (2014), the phenolics compounds found in carob syrups, exhibit antioxidant, antibacterial and antifungal properties. The sugar supplementation during the syrup processing can have an impact on these phenolic compounds. If the addition of sugar is avoided during the preparation of the syrup, it will preserve its bioactive properties, as sucrose is associated with physiological disorders (mainly diabetes and obesity). In addition, carob is known to be the only raw material from which D-pinitol can be isolated in sufficient quantities and can be exploited for commercial purposes (López-Sánchez et al., 2018) and Christou et al., 2019) quantified D-pinitol in 13 carob syrup samples originating from Cyprus by developing a gas chromatography–mass spectrometry (GC–MS) method and the mean concentration was found to be 68.58 ± 4.80 g/kg. Compared to other plants or legumes, carob is distinguished as a highly concentrated source of D-pinitol, underscoring its role as a valuable functional organic food (López-Sánchez et al., 2018) [16].

1.5 Chemical characterization

Generally, the term chemical characterization focuses on the process of defining and measuring the chemical properties and chemical compound's structure. In addition, the definition provided by International Organization for Standardization *(ISO-10993-18)* constitutes the: "recognition of a material and the quantification of its chemical components as part of an evaluation of total biological safety" [94, 95].

Therefore, the chemical characterization of carob syrup can be rendered as the overall determination of its constituents using various analytical methods such as mass spectrometry, gas chromatography, NMR, etc.

1.6 Foodomics-Metabolomics

The globalization of the food market and the constant changes of modern consumer preferences have raised new challenges and significant concerns for raw materials and food products, due to the possibility of their entry from multiple points into the EU. Therefore, the need for food traceability, throughout the production process, is considered imperative, to ensure their quality, safety and nutritional value. The challenge of -omics in the field of food safety and quality concerns the development of new food production technologies and innovative trends in nutrition, with the main goal of facilitating the emergence of new risks and at the same time the suspension of older ones, issues of microbial resistance and Economically Motivated Adulteration (EMA).

Foodomics is defined as the comprehensive, high-performance approach to harnessing food science to improve human nutrition. It is a new approach that studies the food sector and with the nutrition sector, as an entity, improve human health and well-being. The *foodomics* approach is much more than the application of different analytical methods, but rather a new way of thinking, i.e., a new science of food and nutrition and represents a significant scientific challenge. Researchers have the ability to link together food components, nutrition, humans, health and disease, which requires not only the application of advanced analytical methods and techniques, but above all, the ability to face a different perspective and approximation [17].

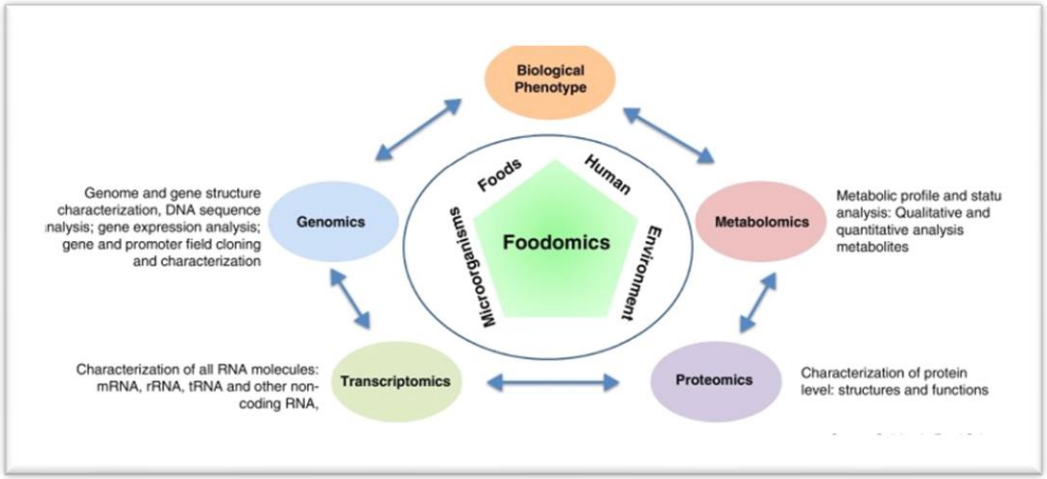


Figure 7. Foodomics involves the use of multiple tools to deal with the different applications [18].

Metabolomics is a new emerging field in metagenomic research that concerns the identification and quantification of small molecules, that are intermediates and end products of metabolism, named metabolites. Metabolites are defined as the full set of endogenous or exogenous of low molecular weight (< 1000 Da). The metabolome was deemed as the terminal downstream product of the genome, transcriptome, and proteome, and thus reflects most closely the phenotypes. Metabolomics is the systematic and thorough examination, both qualitatively and quantitatively, of the metabolome of an organism or a biological system. It involves analyzing and studying the complete set of small molecular species, i.e., the metabolites present in a biological sample, such as cells, tissues, or biofluids. By identifying and measuring the metabolites, metabolomics provides valuable insights into the biochemical and physiological status of the biological system under investigation. There are three basic approaches used in a metabolomic study, i.e., target analysis, metabolic profiling, and metabolic fingerprinting. Target

analysis focusing on specific metabolites representative of key pathways, is a useful tool for evaluation metabolic modifications out coming from genetic mutation, altered gene expression, and protein dysfunction. Metabolic profiling focuses on the analysis of a group of metabolites. The comparison of patterns of metabolites that change in response to a treatment, environmental or genetic alterations, among others, is defined as metabolic fingerprinting [19].

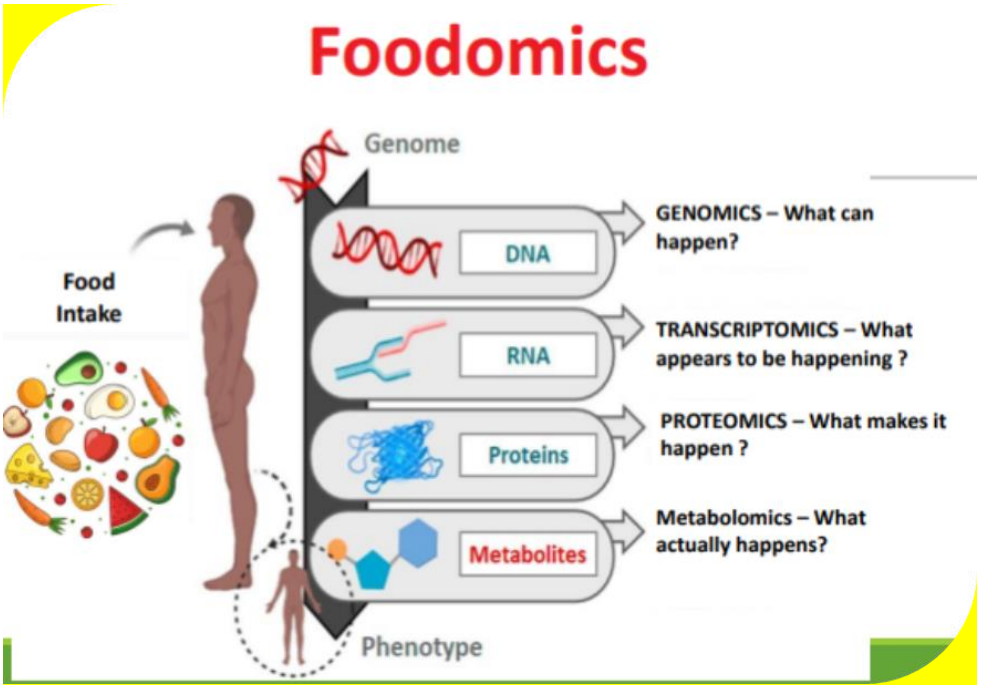


Figure 8. The questions that each -omics field must answer [20].

Nowadays, two state of the art analytical methods are currently used in metabolomic analyses: NMR spectroscopy and MS. NMR is a non-destructive technique that does not require prior separation of the compounds in the sample. It provides detailed information on the molecular structure and requires minimal sample preparation but has a low sensitivity. On the other hand, MS is much more sensitive than NMR. The traditional approach is MS to be coupled to a separation technique, such as chromatography-based methods, to identify and quantify metabolites with good sensitivity and specificity. Others utilize surface-based mass analysis with or without matrices, and with ultra-high-accuracy mass analyzers to further increase sensitivity, minimize background and reduce the sample preparation [21].

2. Chapter 2
Analytical Methods

2.1 Mass Spectrometry (MS)

A mass spectrometer is used to establish the mass of a molecule. The mass-to-charge ratio (*m/z*) of the molecule's ionized forms is measured. More specifically, ions are generated by losing or gaining a charge from neutral species, before being electrostatically directed into the mass analyzer in which they are analyzed based on *m/z* and detected. The process of molecular ionization, separating ions, and detecting them results in a spectrum that furnishes both molecular mass and structural details. Generally, all mass spectrometers consist of the following basic components: a sample inlet, an ionization source, a mass analyzer, and an ion detector as presented in Figures 9 and 10 [22].

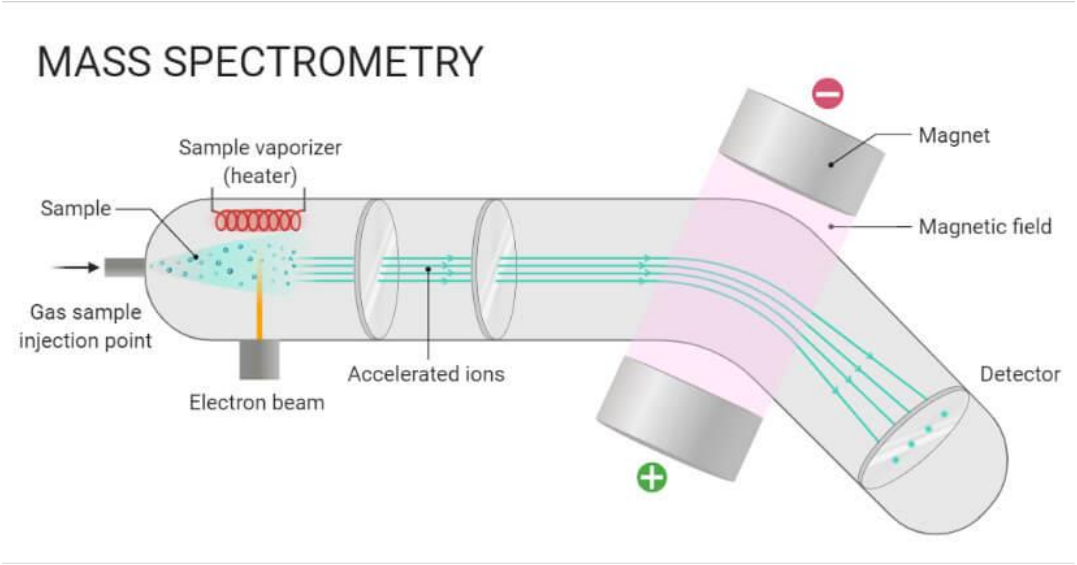


Figure 9. Principle of Mass Spectrometry (MS) [23].

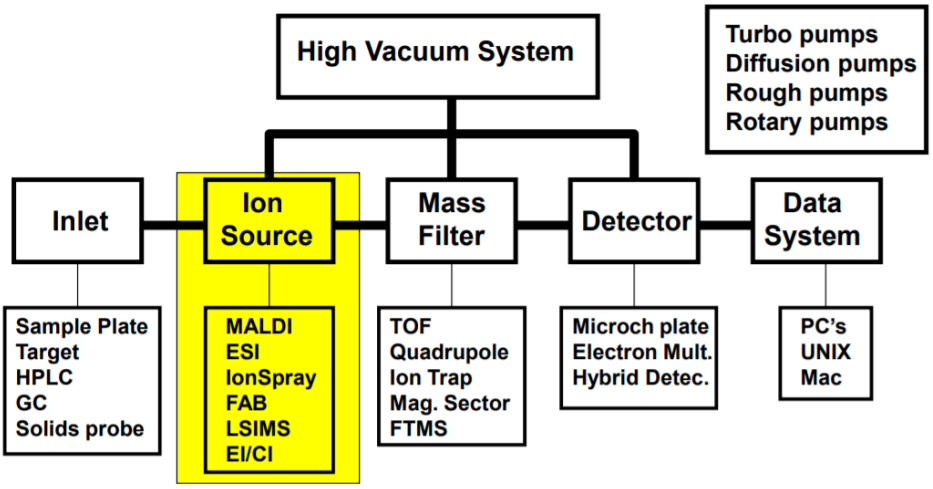


Figure 10. Types of ionization sources [24].

2.1.1 Types of ionization sources

In hyphenated liquid chromatography (LC-MS) the two most used sources are the Electrospray Ionization (ESI) and the Atmospheric Pressure Chemical Ionization (APCI). The desorption electrospray ionization (DESI) and the matrix-assisted laser desorption/ionization (MALDI) are two other metabolite desorption/ionization methods for natural product analysis [25].

2.1.1.1 Electrospray Ionization (ESI)

The ESI interface is described in the following process. Firstly, the eluent is passed through a sharp steel needle in which high voltage (2-6kV) is applied and therefore charged eluent droplets are produced. Subsequently, the eluent droplets are heated and evaporated forming the so-called Taylor cone under the influence of a strong electric field. As a result of the increased tension on the droplet surface, coulomb fission occurs, and the droplets repeatedly decompose, forming smaller droplets. Accordingly, gas phase analyte ions are released from the droplets before being transferred by a flow of dry nitrogen and electrostatic forces into the MS analyzer [25, 26].

At the ESI interface the produced charged ions are typically protonated $[M+H]^+$ or deprotonated $[M-H]^-$, but they might also form a variety of other adducts with elements such as sodium or potassium. The ionization efficiency of an analyte is related to mass spectrometry detection limit. Therefore, the MS signal of two compounds with the same concentration in a sample can be differentiated, due to different sensitivity. The ESI response is affected by many other factors such as flow rate, solvent additives, and analyte physicochemical properties. Inordinate amounts of salt and high flow rates should

be avoided, as they lead to signal suppression, whereas derivatization improves ionization efficiency.

Lastly, the quantitative structure-property relationships (QSPR) studies are executed. These studies correlated the compound properties with ESI sensitivities and discovered that molecular volume, logP, pKa, and surface activities are more relevant to ESI ionization efficiency [28].

2.1.1.2 Atmospheric Pressure Chemical Ionization (APCI)

The APCI interface has many similarities with ESI. Yet, in APCI high voltage is applied on a corona discharge needle to an eluent being sprayed by the influence of high flow rate and dry nitrogen. In contrast to ESI in APCI, analyte ionization happens in a vapor state rather than in liquid droplets. Particularly, the analyte is initially vaporized to the gas phase and the H_3O^+, $CH_3OH_2^+$ vapor is usually used as proton donors. Analytes with high ionization efficiency in atmospheric pressure chemical ionization (APCI) are typically volatile and possess a significant gas-phase basicity [29].

Rebane et al., investigated the ionization of 40 compounds over a wide range of pKa and logP. From latter research it was shown that some highly non-polar compounds, for example pyrene and anthracene, could only be ionized using the APCI source. Moreover, ionization efficiency was also found to be positively correlated with weighted average negative sigma, logP, and molar volume [30].

In the study of Olmo-Garcia et al., it was discovered that polar metabolites, such as phenolic compounds, and flavonoids, are more efficiently ionized in positive mode ESI, whereas organic acids and coumarins in negative mode ESI. However, non-polar sterols and tocopherols revealed the highest efficiency using APCI [31].

2.1.2 Mass analyzers

There are two types of mass analyzers: 1) low resolution (e.g., quadrupole) and 2) high resolution (e.g., orbitrap, time of flight).

High resolution mass spectrometry (HRMS), as defined by the Royal Society of Chemistry, encompasses all forms of mass spectrometry that enable the determination of the precise mass of molecular ions in a sample, as opposed to the "nominal" mass, which is determined by multiplying the number of atoms for each element by the closest whole number of the isotope mass of the most prevalent isotope of that element, which is rounded to the nearest integer value.

The resolution of the instruments is utilized to describe the performance characteristics of high-resolution mass analyzers. The respective mass analyzers in this category are the Time-Of Flight (TOF), the Fourier Transform Ion Cyclotron Resonance (FTICR) and the Orbitrap (OT) [32].

2.1.2.1 Quadrupole

As all mass analyzers, quadrupole mass analyzer is a type of mass spectrometer that is used to separate and analyze ions based on their mass-to-charge ratio. It consists of four parallel metal rods arranged in a square pattern, with an oscillating radio frequency (RF) voltage applied to the rods. The RF voltage creates an oscillating electric field that traps and guides ions through the analyzer. As ions enter the quadrupole mass analyzer, they are separated based on their mass-to-charge ratio. The oscillating electric field causes ions of different masses to follow different trajectories through the analyzer, leading to their separation. Only ions with a specific mass-to-charge ratio can pass through the quadrupole mass analyzer and reach the detector (Figure 11). The quadrupole mass analyzer is commonly used in applications such as drug discovery, environmental monitoring, and forensics, where the identification and quantification of small molecules with high sensitivity is required [33].

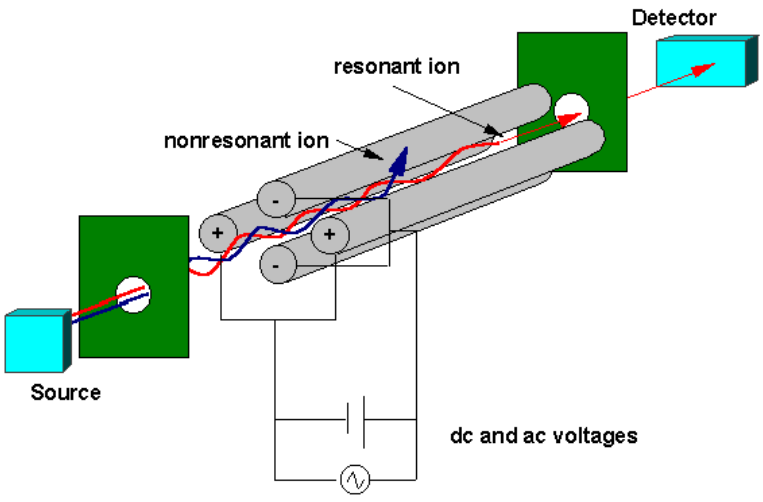


Figure 11. Schematic of a quadrupole [34].

2.1.2.2 Orbitrap

The orbitrap is an electrostatic ion trap which generates mass spectra using Fourier transform. The external part consists of a barrel-shaped electrode divided into two equal

parts with a small interval. The central electrode has the shape of a spindle and a maximum diameter of 8 mm, whereas the outer electrode's maximum diameter is 20 mm (Figure 12). Through the interstice between the two parts of the external electrode, the ions are injected tangentially. The central electrode is met with an electrostatic voltage of several kilovolts, which is negative for positive ions and positive for negative ions, whereas the outer electrode is at zero voltage. The ions are injected with some kiloelectronvolts of kinetic energy. Consequently, they begin to move in spiral trajectories inside the trap, around the central electrode under the impact of an electrostatic field which is a result of the DC voltage and trap geometry. The voltage of the inner and outer electrode is set to about -3200 V for positive ions, and close to zero respectively. Ions follow a circular or oval path around the center electrode with kinetic energies around 1600 eV. They begin to oscillate along the z axis as they rotate around the inner electrode. The detection of the ions is done via their image current over an electrometer and their conversion to mass spectra by the Fourier Transform [35]. A scheme of an orbitrap MS is illustrated in Figure 12.

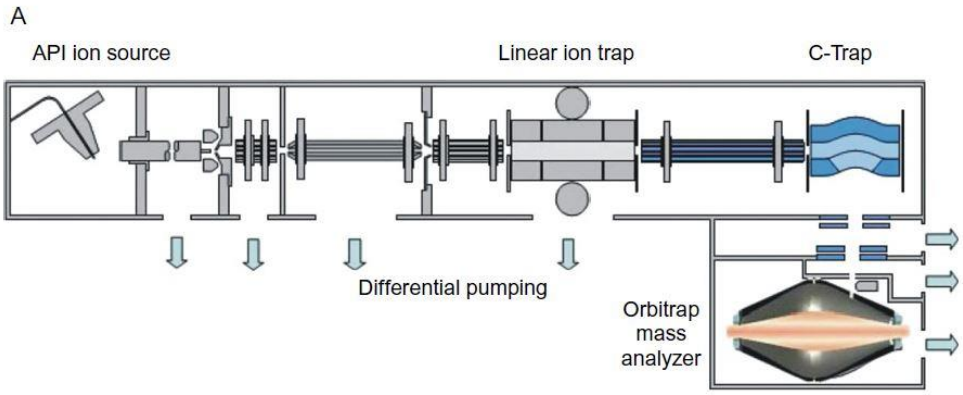


Figure 12. Scheme of the orbitrap instrument [35].

2.1.2.3 Time of flight

A Time-of-flight (TOF) mass analyzer is a type of mass spectrometer that uses the TOF technique to separate and detect ions based on their mass-to-charge ratio (m/z). In a TOF mass analyzer, ions are accelerated by an electric field and then travel through a flight tube where they are separated based on their m/z ratio. The basic components of a TOF mass analyzer include an ion source, an acceleration region, a flight tube, and a detector. The ion source produces ions from a sample, and the ions are accelerated into the flight tube by an electric field. As the ions travel down the flight tube, they separate

based on their m/z ratio, with lighter ions reaching the detector faster than heavier ions. The detector records the time of arrival of the ions, which is used to calculate their m/z ratio (Figure 13). The resulting data is a mass spectrum, which shows the distribution of ions based on their m/z ratio. The mass spectrum can be used to identify the chemical composition of a sample, determine the molecular weight of a compound, and study the fragmentation patterns of ions. TOF mass analyzers are commonly used in proteomics, metabolomics, and other fields of analytical chemistry, due to their high mass resolution, sensitivity, accuracy and especially due to their speed. They are also used in many other applications, such as laser mass spectrometry and particle physics research [36].

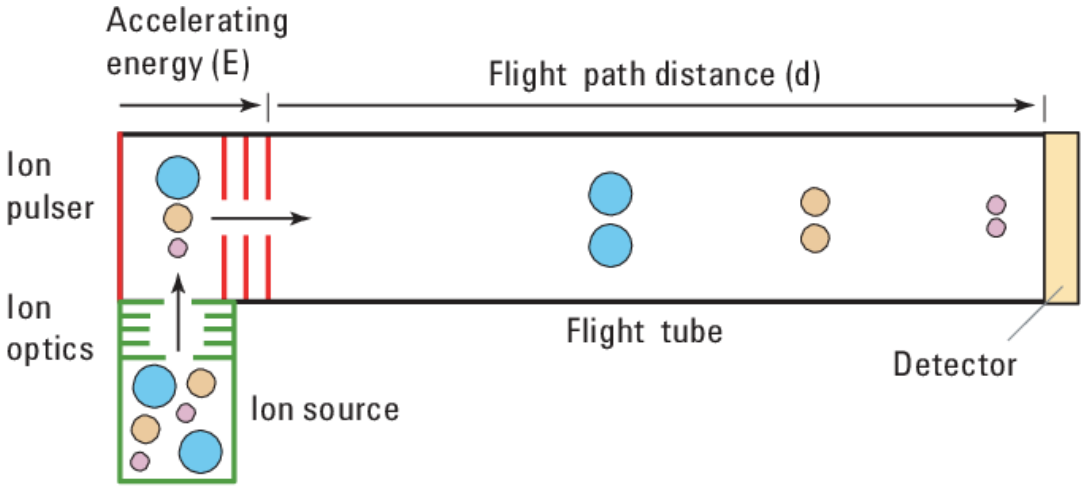


Figure 13. Time-of-flight analysis [37].

2.1.2.4 Comparison between Orbitrap and TOF

In order for the two mass analyzers Orbitrap and TOF to be compared, their advantages and disadvantages need to be established. The orbitrap can achieve high mass accuracy in the range of 1-2 ppm and high resolution up to 500.000. Moreover, its dynamic *m/z* range is quite broad, and it is very sensitive as compounds with sub femtomole concentrations can be determined. On the other hand, it is a very expensive instrument with mostly high vacuum requirements. Another drawback of Orbitrap is its low speed, due to the effort required to obtain highly accurate spectra [38, 39].

The mass accuracy range is between 10 to 100 ppm and resolution up to 60.000. These values are restricted in comparison with Orbitrap. However, the scan rate of the TOF is very high, reaching 20.000 scan/s. Therefore, it offers the possibility to collect information for the average spectra and gain better accuracy. Also, its mass range is theoretically unlimited. Another benefit of TOF is that it is more compatible with UPLC than Orbitrap.

Lastly, it has the same sensitivity as the Orbitap, as it reaches the concentrations of the analyte in the magnitude of subfemtomole [40].

2.1.2.5 Hybrid Instruments (Q-TOF)

The Q-TOF-MS is a 'hybrid' instrument, a combination of quadrupoles (Q) and a time-of-flight mass spectrometry (TOF) analyzer. The first quadrupole (Q_1) works as a mass filter, which selects specific ions based on their mass-to-charge ratio (*m/z*). The second quadrupole (Q_2) functions as a collision cell, fragmenting ions by collisions with neutral gas molecules such as nitrogen or argon, causing ion fragmentation via a process known as collision induced dissociation (CID). After leaving the quadrupole, ions are reaccelerated into the TOF analyzer's ion modulator region, where they are pulsed by an electric field and speeded up orthogonally to their original orientation. The flight tube, a field-free drift region is entered by all ions with the same kinetic energy and mass separation is performed. Ions with a lower mass have shorter flight time towards the detector, whereas ions with higher mass take longer to travel towards the detector. Owning to Q-TOF-MS's quadrupole technology, data can be acquired using two unique scan types in conjunction with a time-of-flight analyzer. The first mode (single MS mode) uses the first and second quadrupole in RF-only mode to perform an accurate mass scan of the unfragmented precursor ion. The Q_1 can be used to pick a particular mass or range of masses to transmit to the TOF analyzer. The second mode (MS/MS mode) transfers ions into the collision cell (Q_2), in which CID occurs by employing the Q_1 in RF-only or mass filter mode.

The TOF analyzer is used to perform precise mass measurement on both the resulting product ions and any precursor ions that remain unfragmented. The Q-TOF-MS (Figure 14) can collect both precursor and product ion data at the same time, by switching between the two modes. For the ion detection to be accomplished, a detector system (time-to-digital converter) is used, which converts the flight time of the ion into a mass signal. A two-stage ion mirror (reflectron) is utilized to compensate the different kinetic energies enhanced mass accuracy [41].

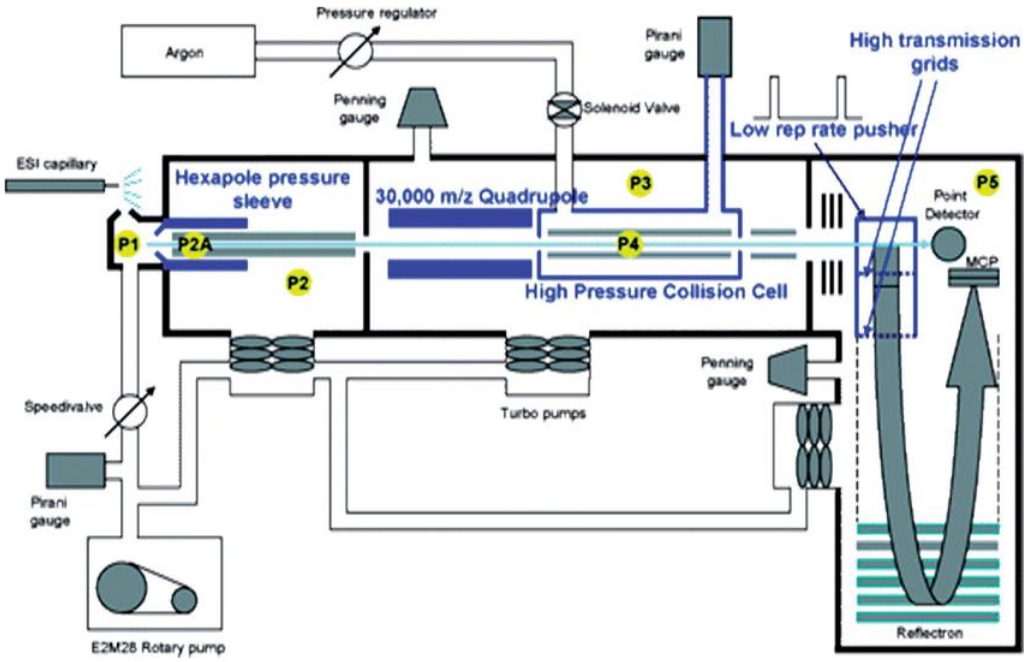


Figure 14. Q–TOF instrumentation [42].

2.2 Mass spectrometry applications

Mass spectrometry finds extensive use across various applications in many fields (food, environmental, environmental epidemiology, pharmaceutical, health, etc). Firstly, mass spectrometry in conjunction with liquid or gas chromatography is used to detect contaminants in foods (i.e., veterinary drugs) [43], mycotoxins, and pesticides which are highly important for food quality assurance [44, 45].

Also, mass spectrometry is essential for food authentication and traceability, as it manages to determine individual food components or food's chemical fingerprints [46]. What is important to point out is that mass spectrometry imaging (MSI), will possibly become the future prospect for the evaluation of food quality, as the distribution of its constituents will be depicted, and food-related factors based on their molecular weights will be identified [47]. The field of environmental chemistry is benefited from the investigation of emerging contaminants (ECs) using LC-MS/MS, which facilitate monitoring studies and the comprehensive regarding the occurrence, fate, and ecological consequences of emerging contaminants (ECs) in both wastewater and the natural environment [48].

Additionally, in environmental sciences the use of HRMS for a variety of tasks such as target analysis, retrospective analysis, metabolite and transformation product discovery, and non-target analysis, with thousands of different organic contaminants [49] has gained a growing interest. The approach of wastewater epidemiology was employed to examine

compounds associated with chemical exposure, public health, habits, and lifestyle factors. In the recent decades, LC-MS/MS was extensively used in waste-water analysis, and HRMS with advanced chemometric tools provides a better insight of their composition. In waste-water epidemiology, metabolomics are becoming popular, as endogenous compounds can be identified and linked to the psychological and physical well-being of individuals living in the catchment area. [50].

In the field of health, LC-MS/MS and LC-HRMS are expected to bring additional benefits from emergency toxicology testing, to bring additional benefits from emergency toxicology testing, to provide patient diagnosis and contribute to public health. Also, applying mass spectrometry-based toxicological tests in the treatment of poisoned patients could also be beneficial [51], and metabolomics implementation in drug metabolism studies could help bridge the gap between drug metabolism and drug-induced toxicity. Lastly, LC-HRMS due to its strength in characterization and screening has been growing in popularity for rapid chemical profiling of herbal medicines [52, 53].

2.3 Data processing - MS/MS aacquisition modes

The processing tools typically used for generating tandem mass spectrometry (MS/MS) data in screening methods are called data-dependent (DDA) and data-independent acquisition (DIA). In DDA, precursor ions are selected in MS1 either automatically based on intensity or by the analyst using an inclusion list, followed by fragmentation in MS2 using collision-induced dissociation (CID) or high-energy collision-induced dissociation (HCD). The DDA mode is biased because it intentionally excludes other ions, resulting in mass spectra that may not represent the entire sample composition. Moreover, the stochastic selection of precursor ions leads to less repeatable results. However, the DDA mode is commonly used for compound elucidation since the spectra obtained are clear and easy to interpret [96].

On the other hand, DIA involves initially acquiring two consecutive MS1 scans (one in low collision energy and one in high collision energy) without precursor selection. An improved version of this approach called SWATH acquisition creates *m/z* isolation windows to cover the entire *m/z* range in MS1, systematically collecting MS/MS data for every ion detected in MS1. Unlike DDA, DIA is unbiased and detects and fragments all ions in a sample, providing a more comprehensive and repeatable analysis of a sample by collecting data for a wide range of known and unknown ions. However, the resulting spectra are more complex and may be more difficult to interpret.

Each acquisition mode has its advantages and disadvantages. Therefore, for comprehensive characterization, both acquisition modes should be used to generate complementary information [97].

2.4. Analytical screening strategies

There are 3 main screening strategies in chemical analysis:

- Target Analysis, based on the identification of known metabolites and carried out using corresponding standard solutions.
- Suspect Screening is performed by scanning for a list of metabolites where it has been recorded after a reference review and the samples are analyzed for their possible detection in the samples.
- Non-target screening, where scanning is performed to identify unexpected compounds, since it is difficult to select the individual ions, especially in a complex matrix, or the concentration of the analyte is very low. Under these conditions, it is necessary to make use of sophisticated software with chemometric capabilities, to identify the presence of many compounds and produce spectra for each of the compounds separately [54].

2.4.1 Suspect screening methodology

In current MSc thesis, the suspect screening methodology was followed. Suspect screening is a recently developed methodology. Its purpose is to search for compounds, where there is prior knowledge of their presence in the investigated samples [55]. Lists of these compounds are compiled using their molecular formula (and the resulting calculated exact mass), while corresponding reference standards are not available. Along with the exact mass, the predicted retention time of the suspect compound is another piece of information that is included in a suspect list.

This information is useful for prioritizing the acquired findings, avoiding any false positives, and peak selection when numerous chromatographic peaks are eluted in the same chromatogram. By uploading the canonical SMILES of each compound to our group's online "retention time prediction" program (available at http://rti.chem.uoa.gr/), the retention time of each compound can be predicted. For the correct determination of the prediction time, a mixture of selected compounds "called RTI mix" should be included for the creation of the retention time calibration curve in the same sequence of the analyzed samples. Finally, fragments of the suspect compounds derived from the literature or mass

spectral libraries (e.g., Mass Bank of North America (MoNA), MassBank, and mz Cloud) should be included in the analysis.

2.4.1.1 Identification confidence

There are five identification confidence levels to express the certainty of a candidate identification, according to Schymanski et al. (2014):

a) Level 1: In confidence level 1, the MS/MS spectra and retention time value are in agreement between the candidate and the corresponding reference standard [56].

b) Level 2: Here there are two options, 2a and 2b. Thus, 2a is reached when MS/MS spectrum of online available libraries and the experimental spectra of the candidate are consistent with a similarity score above 0.7 [56], and on the other hand 2b is fulfilled for probable structures when diagnostic ions exist. Diagnostic ions are used as valuable markers for the screening of the compounds, whose formation reveals structural or compositional information about its precursor ion [57]. Furthermore, Fraski et al. [60] provide valuable information about specific *m/z* values, which served as diagnostic fragment ions, and their correlation to structural information [58].

c) Level 3: To reach this level, evaluation of the MS/MS fragments via in silico fragmentation tools by using relevant softwares (MetFrag or CFM-ID) is followed. In addition, retention time prediction and ionization efficiency could assist in the prioritization of the candidates [60, 61, 62, 63, 64, 65].

d) Level 4 is attained when the candidate's molecular formula can be substantiated [56].

e) Level 5 is achieved when the precise mass of interest can be corroborated without any corresponding molecular formula match. [56].

2.5 Nuclear magnetic resonance (NMR)

NMR spectroscopy is a physicochemical method used for determining the structural properties of molecules. It is based on the interaction of externally applied radio frequency radiation with atomic nuclei, which causes an exchange of energy that leads to a change in a nuclear property called nuclear spin. Only those nuclei with $J \neq 0$ are detectable in NMR, such as ^{1}H, ^{2}H, ^{13}C, and ^{15}N, and they act as magnetic dipoles that can align with an external magnetic field, resulting in the magnetization effect. The strength of these microscopic magnets is determined by a constant called the gyromagnetic ratio (γ).

Certain NMR-active nuclei have the ability to adopt two different orientations when aligned with an external magnetic field, one with the lowest energy level and the other with the highest (parallel and antiparallel, respectively). The sensitivity of NMR spectroscopy is influenced by the difference in energy levels (ΔE), which relies on the magnetic field and the gyromagnetic ratio. (Figures 15 and 16). Since its first demonstration in 1938 by Isidor I. Rabi, NMR spectroscopy has advanced alongside other scientific fields, with researchers developing new methodologies for studying complex systems such as membrane proteins, metabolically complex samples, and biological tissues. It is considered one of the most powerful methods for the structural determination of chemical species, as well as for the study of molecular dynamics and interactions [66].

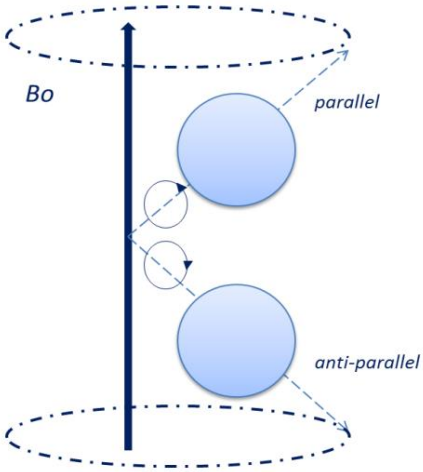


Figure 15. Nuclear magnetic moment of a nucleus in relation to an external magnetic field Bo [67].

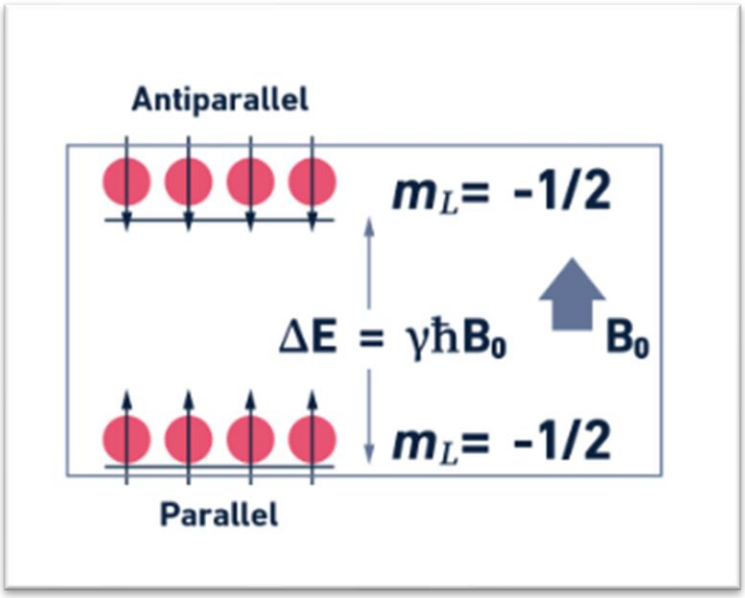


Figure 16. Distribution of nuclear spin populations in the two possible energy levels in nuclei with J=±1/2 [66].

Magnetic resonance is achieved when nuclei are irradiated with radio frequencies and are under the influence of a magnetic field, transitions between energy levels and thus changes in the orientation of the nuclear spins are induced, causing the nuclei to perform a spinning top-like motion (transitional motion) around from an axis parallel to the direction of the field (B_0), with a frequency of this transition motion, called the Larmor frequency (vL).

In an NMR experiment, it is not possible to measure the signal in the z direction, as the magnetic field (B_0) is too strong in that direction. Therefore, a magnetic pulse containing frequencies close to the Larmor frequency is applied perpendicular to the magnetic field to reach the resonance of the nuclear spins, and then a relaxation process takes place. As a consequence of this relaxation, the energy is emitted as radio frequency. This produces a characteristic signal called the free induction decay (FID) which is recorded by the detector. This FID is then transformed into a plot of intensities versus frequencies known as an NMR spectrum [68, 69].

The magnet creates a strong magnetic field that aligns the nuclear spins of the atoms present in the sample. Generally, in NMR spectroscopy, magnets based on superconducting materials (e.g., He, N_2) are used, as presented in Figure 17. Therefore, they require very low temperatures to function (around 4 K). For this reason, NMR spectrometers contain a cooling system consisting of an inner shell filled with liquid helium

which is cooled by an additional shell filled with liquid nitrogen and several layers of thermal insulating materials to keep the temperature constant.

The superconducting magnet surrounds a cylindrical chamber, the "detector", into which the sample is introduced and thus placed under the influence of the magnetic field. In addition, there is a series of magnetic coils also located around the sample to radiate the RF pulses to detect and collect the NMR signal emitted by the sample, but also to check the homogeneity of the magnetic field and apply pulse gradients. Therefore, all experimental conditions are controlled by the electronic system of the spectrometer. Thus, it is possible to set and modify each parameter of the NMR experiment through the computer, acquiring data. The NMR spectrum is obtained through the mathematical Fourier transform, with peaks of different intensities as a function of the chemical shift derived from the Larmor frequency of the different atomic nuclei present in the sample [70].

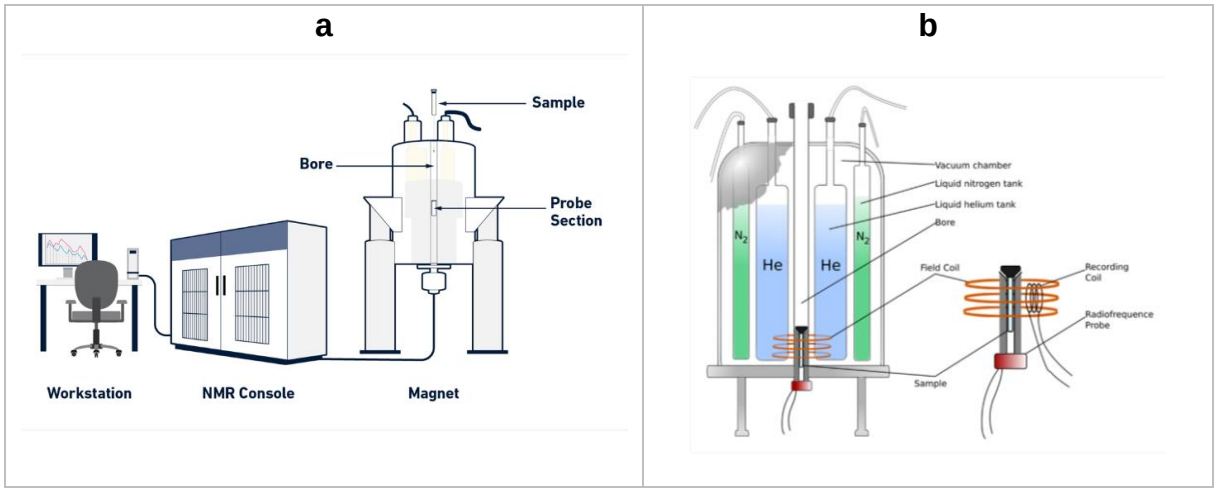


Figure 17. a) NMR instrumentation, b) Internals data of the NMR spectrometer and representation of the detector [71].

The resonance frequencies of the nuclei are measured and converted into an NMR spectrum that shows all the correct frequencies as peaks on a graph. The height of each peak represents the number of nuclei tuned to each particular frequency, the signal strength. Therefore, the more tuned nuclei, the higher the intensity, while the value of each frequency gives information about the environment of that person. such as its neighboring atoms and their relative positions.

2.5.1 ^{1}H vs ^{13}C NMR

The ^{1}H NMR spectra are more complex than the ^{13}C spectra as indicated by the relative intensity peak heights of each chemical treatment. These relative intensities correspond

to the number of hydrogens in that particular environment within a molecule, shown as a number above the peak.

In the case of hydrogen, the most abundant isotope is ^{1}H (99.98%, J= ½) and is an NMR-active nucleus, while in the case of carbon, its most abundant isotope is not NMR-active (^{12}C, 98, 89 %, J = 0). NMR spectrometers can only detect the ^{13}C isotope, which has an abundance of 1.11%, so ^{13}C NMR is significantly less sensitive than ^{1}H NMR and makes the ^{13}C experiment more time-consuming. Moreover, the chemical shift of ^{1}H nuclei commonly fall within the range of 0-14 ppm, whereas the chemical shifts of ^{13}C nuclei (Figure 18) span a significantly broader range, typically 10-220 ppm. These increased shifts in ^{13}C NMR result in better resolution compared to ^{1}H NMR, as the signals are typically more dispersed [72].

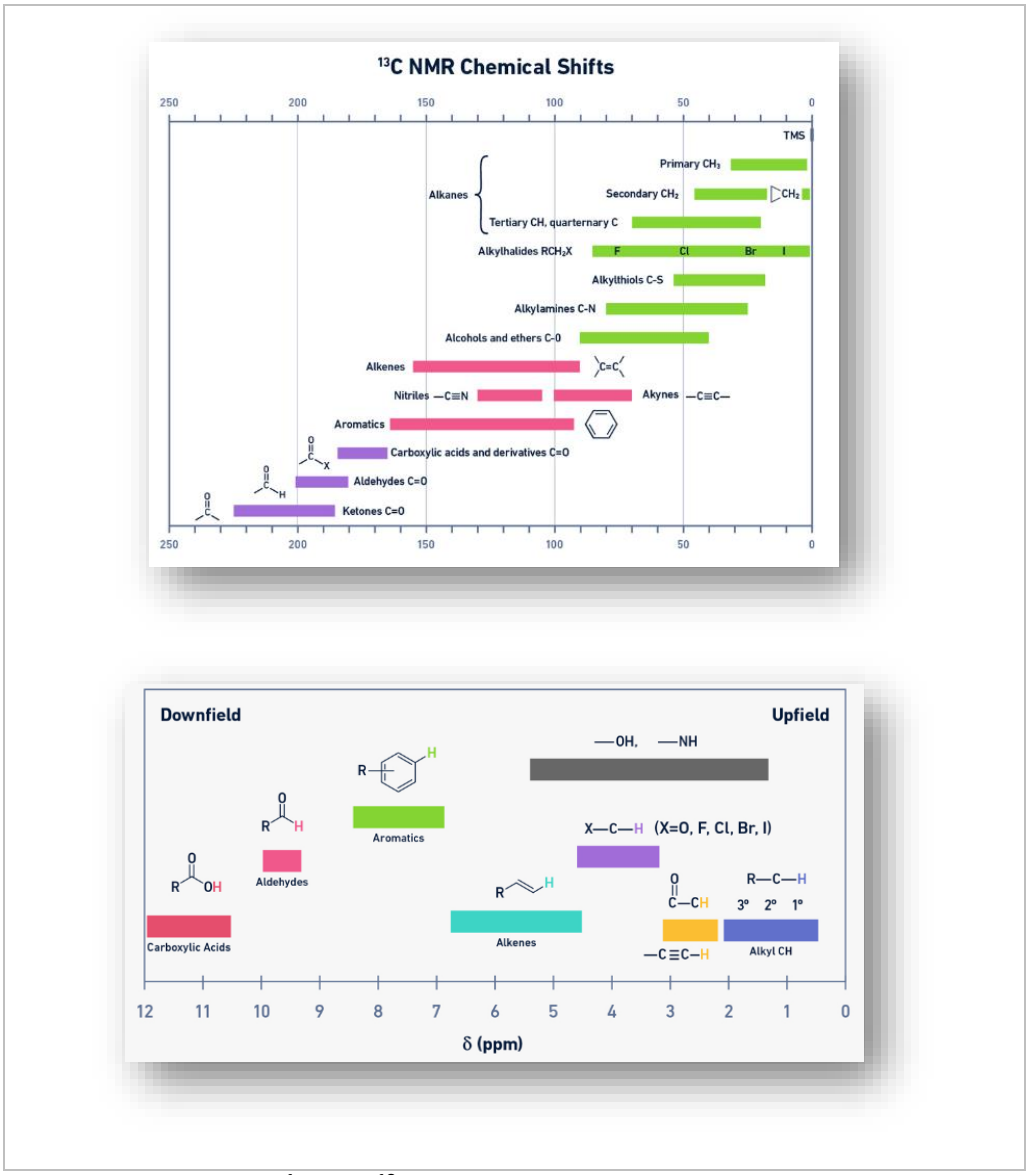


Figure 18. H^{1} and C^{13} NMR typical chemical shift values [73].

2.5.2 2D NMR

The magnetic pulses used in NMR have been developed since its discovery. Therefore, in the case of macromolecules like proteins, which contain numerous active NMR nuclei resulting in complex spectra with overlapping broad peaks, the application of 2D NMR can be utilized to tackle these challenges. 2D NMR experiments generate spectra defined by two chemical shift axes (rather than one, as in 1D spectra), with signals correlating pairs of different nuclei. The most important of these are:

a) COSY, (Correlated Spectroscopy) where spectra show peaks associated with pairs of nuclei separated by at most three chemical bonds.

b) TOCSY (Total Correlated Spectroscopy) is a technique in which spectra reveal signals indicating correlations between pairs of nuclei belonging to the same spin system.

c) NOESY, (Nuclear OverHauser Effect Spectroscopy), where spectra contain peaks that correlate pairs of nearby nuclei (typically, separated by less than 5–6 Å). Unlike COSY, NOESY correlation arises from the nuclear Overhauser effect, in which interaction occurs when two nuclei are spatially close, regardless of the number of chemical bonds separating them [73].

2.5.3 NMR applications

The main NMR applications include the structural identification and quantification of organic, organometallic, and biochemical molecules. Also, NMR works complementary to data obtained from other techniques such as mass spectrometry, infrared spectroscopy, and elemental analysis. A few different NMR applications are presented in Table 3.

Table 3. NMR applications [74-82].

Field	Application	Reference
Chemistry	Structural determination of new compounds, product quality control and purity determination	[74]
Pharmaceuticals	Study of structure, dynamics and molecular interactions for drug discovery, quality control	[75]

	and determination of drug purity	
Petrochemistry	Analysis of rock materials to test the suitability of an oil reservoir for exploitation, solid-state NMR compositional analysis of petroleum derivatives, product quality control	[76]
Materials	Characterization of new materials by solid-state NMR	[77]
Molecular biology and biophysics	Study of structure, dynamics and molecular interactions of peptides, proteins, nuclei acids, carbohydrates, and other biomolecules	[78, 79]
Health sciences	Analysis of biological fluids to obtain metabolic profiles related to diseases (metabolomics), use of NMR imaging techniques for medical diagnosis	[80, 81]
Food science	NMR fingerprint analysis to check quality or authenticity of food samples	[82]

2.5.3.1 NMR-based metabolomics

In metabolomics there are many areas with spectral overlap. So, a quite good resolution is necessary to discriminate the different signals. The coefficients of variation of FWHM and of the S/N ratio are also significant to ensure that the acquisition quality during the

experiment remains stable. As in a metabolomic study, groups of spectra are compared, the spectral quality must remain constant. It is important to check that the S/N ratio remains relatively uniform across the different levels of the biological factors in the experiment.

Using a variety of stable isotope labeled precursors, NMR can trace metabolic processes and measure metabolic fluxes. It can detect metabolites by using one or more types of atomic nuclei, such as ^{1}H, ^{13}C, ^{31}P, or ^{15}N [83].

Additionally, NMR can export useful conclusions in nutrition research for dietary biomarker discovery and diet related disease studies [84].

The medical field is another field that NMR can play a significant role, such as in the analysis of blood samples from patients [85].

2.6 Chemometrics (PCA and PLS-DA techniques)

A very useful tool for finding the metabolic profile of complex mixtures such as carob syrup is the development of advanced chemometric methods, which help to better understand this topic, in combination with data obtained from LC-HRMS and NMR. Initially, the spectroscopic information should be sorted out, meaning that there is the need to transform each NMR or MS signal in a numeric value. To achieve that and according to the analytical platform used, there is a list of different software providing the possibility to convert the raw data into a single data matrix ready to be utilised for statistical analysis. The next step in the statistical analysis of metabolic profiling data is to perform both unsupervised and supervised analysis, with the aim of identifying and characterizing important differences between classes resulting from the presence of external or internal stimuli to the organism. This includes the development of advanced chemometric methods depending on the purpose of the study and the nature of the dataset. On one hand, univariate statistical methods, like the t-test, can be applied and on the other hand multivariate data statistical analyses are usually used in the demanding and complex multidimensional metabolomic datasets. Classical multivariate methods employed in metabolic profiling analysis include principal components analysis (PCA) and partial least squares (PLS) regression analysis [67].

PCA and PLS-DA are both statistical techniques used in data analysis, particularly in the fields of chemometrics and bioinformatics. PCA is a method used for data reduction and dimensionality reduction. It involves the transformation of a set of correlated variables into a set of uncorrelated variables called principal components. The first principal component

accounts for the largest possible variance in the original data, the second for the next largest variance, and so on. PCA is often used to visualize patterns in high-dimensional data, reduce the number of variables, and identify key relationships between variables.

PLS-DA, on the other hand, is a supervised method used for classification and discrimination. It involves building a linear model that maximizes the separation between classes in a dataset. PLS-DA works by finding latent variables that are linear combinations of the original variables, and these latent variables are used to model the relationship between the predictor variables and the class variable. PLS-DA is commonly used to identify differences between groups in large datasets, such as in the identification of biomarkers for a disease or for drug discovery.

In summary, PCA is an unsupervised method used for data reduction and dimensionality reduction, while PLS-DA is a supervised method used for classification and discrimination (Figure 19). Both techniques have a wide range of applications in various fields, including chemistry, biology, and engineering [86].

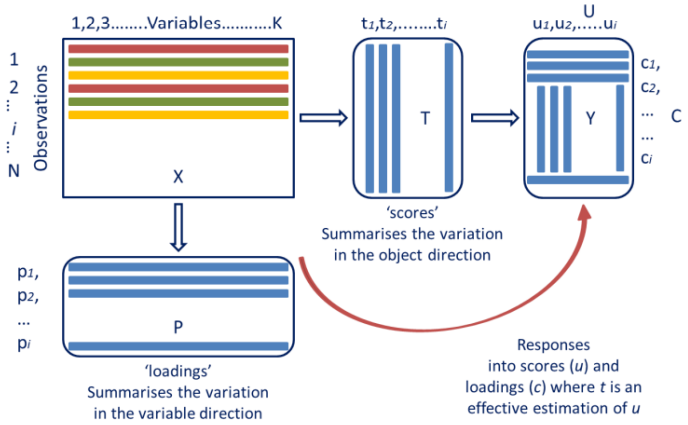


Figure 19. Presentation of PCA (X) and PLS (Y) projection methods. Visualisation of PCA scores (T) and loadings plot (P) [67].

3. Chapter 3
Experimental Part

3.1 UHPLC-QTOF-MS

3.1.1 Samples

The samples analysed in the term of the current work are listed below.

- 11 Cypriot commercial samples
- 1 Cretan commercial sample
- 1 Australian commercial sample
- 22 Cypriot carob fruits were collected from different altitudes and carob syrup was traditionally prepared. They were derived from 3 different Cyprus regions and prepared under various temperatures, as detailed presented on Table 6.
- 1 Slovenian and 1 Croatian carob syrup samples were also prepared under the same Cypriot traditional way.

3.1.2 Reagents and Standards

All standards and reagents were of high-purity grade (>95%) and were described below:

1. Methanol (MeOH) (LC–MS grade**,** Merck, Darmstadt, Germany**)**
2. Formic acid 99% (LC–MS grade, Fluka, Buchs, Switzerland)
3. Isopropanol (Sigma Aldrich, Merck, Darmstadt, Germany)
4. Ammonium acetate (Fisher Scientific, Geel, Belgium)
5. Ammonium formate (Fisher Scientific, Geel, Belgium).
6. Distilled water (Milli-Q purification apparatus, Millipore Direct-Q UV, Bedford, MA, USA)

Reference Standards**:**

Thirty reference standards were applied for qualitative and quantitative purposes. These reference standards and their corresponding class are tabulated in Table 4:

Table 4. Reference standards and their corresponding chemical class.

Chemical class	Reference standards
Amino acids	Glutamic acid, L-Tyrosine, L-Threonine, Aspartic acid, Alanine, Valine, *γ*-aminobutyric acid, Proline, Phenylalanine, L-isoleucine, L-Serine, L-Tryptophan, L-Histidine, L-Methionine, Pyroglutamic acid
Sugars	D-Glucose

Organic acids	a-Hydroxybenzoic acid, Ascorbic acid, Chlorogenic acid, p-Coumaric acid, Syringic acid, trans Cinnamic acid, Gallic acid, 2,5-Dihydroxybenzoic acid
Flavonols and flavonoids	Apigenin, Naringenin, Hesperitin*
Vitamins	Nicotinamide, Nicotinic acid
Alkaloid	Theobromine

*Hesperitin was used as internal standard

Standard stock solutions of individual compounds (1000 mg L^{-1}) were prepared in MeOH (LC-MS grade) and stored at −20 °C in amber glass bottles. Standard stock solutions of individual compounds of amino acids (10000 mg L^{-1}) were diluted in HCl 0.1 M or HCl 1M depending on the analyte and stored at 4 °C in amber glass vials.

Mixtures of working solutions of various concentrations ranging from 0.5 to 25 mg L^{-1} of each analyte were prepared by subsequent dilution of the stock solutions in MeOH: H_2O (50:50 v/v).

Consumables

- Acclaim RSLC 120 C18 column (2.2 µm, 2.1 x 100 mm), Thermo Fisher Scientific (Dreieich, Germany) equipped with the Van guard Acquity UPLC BEH C18 (1.7 µm, 2.1 x 5 mm, Waters, Ireland)
- Regenerated cellulose syringe filters (RC filters, pore size 0.2 µm, diameter 15 mm), Macherey-Nagel (Düren, Germany)
- Centrifuge tubes, 15 mL, and 50 mL (nerbe plus, Germany)
- Eppendorf tubes 1.5 mL (nerbe plus, Germany)

Apparatus

- Ultrasonic bath (Ultrasons H-D, J.P. selecta)
- Centrifugal Hettich Rotofix 32A

Analytical Instruments

Ultra-High-Performance Liquid Chromatography (UHPLC) with an HPG-3400 pump (Dionex UltiMate 3000 RSLC, Thermo Fisher Scientific, Dreieich, Germany), Hybrid Quadrupole Time of Flight Mass Analyzer (QToF-MS) (Maxis Impact, Bruker Daltonics, Bremen, Germany)

3.1.3 Preliminary experiments - Fractional experimental design

Design of experiments (DOEs) is a very effective and powerful statistical tool that can help understand and improve processes and design better products. The main objectives of DOE are to reduce experimental costs, decreasing the number of tests while studying as many factors as possible to find the most important ones. Therefore, in the initial stages of the experimental process the effects of 4 two-level factors (Solvent, Sonication, Shaking, Centrifugation) were investigated. The parameters were tabulated in the following Table (Table 5). DOE below was performed to evaluate the response of each factor [87].

Table 5. Experimental design.

Std	Run	Factor 1 A:MeOH % %	Factor 2 B:AcOH 0.1% %	Factor 3 C:Sonication min	Factor 4 D:Shaking min	Factor 5 E:Centrifugation min	Factor 6 F:F
9	1	50	100	30	0	5	-1
12	2	0	0	0	0	5	-1
10	3	0	100	30	30	5	-1
2	4	0	100	30	0	10	1
5	5	0	0	30	0	10	1
6	6	0	0	0	30	5	1
11	7	50	0	30	30	10	-1
7	8	50	0	0	0	10	-1
8	9	50	100	0	0	5	1
1	10	50	100	0	30	10	1
4	11	0	100	0	30	10	-1
3	12	50	0	30	30	5	1

Therefore, through DOE and MS dial software (*5.1.221218*), the best possible experimental procedure for carob syrup sample preparation was found based on the maximum number of features.

3.1.4 Sample pretreatment

3.1.4.1 Carobs processing

150 g of carobs (from Cyprus, Slovenia, and Croatia) weighted, cut in small pieces using a pair of scissors and placed in 50 mL centrifuge tubes. Afterwards, they were transferred in a 1000 mL beaker and 500 mL potable H_2O were added. After that, they were boiled for 5 minutes. The samples were kept in the water for 36 hours and then filtered and further heated until a final volume of 50 mL was reached.

3.1.4.2 Effect of temperature in Cyprus carob syrup samples

Several samples (n=19), following the same pretreatment as above, were heated at different temperatures to examine the effect of temperature in the process, as presented in detail in Table 6.

Table 6. Samples preparation in different temperatures.

Sample/Region	Saint Fyla (altitude 100 m)	Dieronas (altitude 460 m)	Klonari (Altitude 480 m)
Sample 1	Water extraction for 36-40 h at 20 °C and filtration	Water extraction for 36-40 h at 20 °C and filtration	Water extraction for 36-40 h at 20 °C and filtration
Sample 2 (5' boiling)	Boiling sample 1 at 80 °C for 5', water extraction for 36-40 h, and filtration	Boiling sample 1 at 80 °C for 5', water extraction for 36-40 h, and filtration	Boiling sample 1 at 80 °C for 5', water extraction for 36-40 h, and filtration
Sample 3 (10' boiling)	Boiling sample 2 at 82 °C for 5', water extraction for 36-40 h and filtration	Boiling sample 2 at 88 °C for 5', water extraction for 36-40 h and filtration	Boiling sample 2 at 90-100 °C for 5', water extraction for 36-40 h and filtration
Sample 4 (15' boiling)	Boiling sample 3 at 82 °C for 5', water extraction for 36-40 h and filtration	Boiling sample 3 at 88 °C for 5', water extraction for 36-40 h and filtration	Boiling sample 3 at 90-100 °C for 5', water extraction for 36-40 h and filtration
Sample 5 (20' boiling)	Boiling sample 4 at 82 °C for 5', water extraction for 36-40 h, and filtration	Boiling sample 4 at 90-100 °C for 5', water extraction for 36-40 h, and filtration	Boiling sample 4 at 90-100 °C for 5', water extraction for 36-40 h, and filtration

Sample 6 (25’ boiling)	Boiling sample 5 at 82 °C for 5’, water extraction for 36-40 h, and filtration	Boiling sample 5 at 95 °C for 5’, water extraction for 36-40 h, and filtration	Boiling sample 5 at 96 °C for 5’, water extraction for 36-40 h, and filtration
Sample 7 (30’ boiling)			Boiling sample 6 at 96 °C for 5’, water extraction for 36-40 h, and filtration

One hundred mg of all the samples (traditional and commercial) were weighted in centrifuge tubes of 15 mL. The standard working solution of Hesperitin was added as internal standard (C= 8.0 mg kg $^{-1}$). The samples remained for 15 min, to facilitate the better incorporation of the analytes in the matrix. Afterwards, 5 mL of a mixture of methanol: water in the proportion of 50:50 was added. The solutions were vortexed for 30 sec and shaken vigorously on a table shaker for 15 min. Then, the solutions were centrifuged at 4000 rpm for 10 minutes and 500 μL from the supernatants were transferred to eppendorf tubes and the samples were diluted with a mixture of methanol: water in a proportion of 50:50 (v/v) for quantification purposes. The solutions were filtered through RC syringe filters 0.22 μm to glass autosampler vials for analysis. The abovementioned sample preparation was performed for the analytical procedural blank, to test for contaminations by the sample preparation procedure, as well as for the spiked sample for quantification purposes.

3.1.5 Sample analysis

The chemical characterization of carob syrup samples was achieved using an Ultra High-Performance Liquid Chromatography (UHPLC) apparatus with an HPG-3400 pump (Dionex UltiMate 3000 RSLC, Thermo Fisher Scientific, Dreieich, Germany) equipped with Hybrid Quadrupole Time of Flight Mass Analyzer (QToF-MS) (Maxis Impact, Bruker Daltonics, Bremen, Germany). An Acclaim RSLC 120 C18 column (2.2 μm, 2.1 x 100 mm), Thermo Fisher Scientific (Dreieich, Germany) equipped with a Van guard Acquity UPLC BEH C18 (1.7 μm, 2.1 x 5 mm, Waters, Ireland) was used for the chromatographic analysis. The column temperature was set at 30 ^{o}C throughout the analysis, and the injection volume was 5 μL. The instrument was operated in both positive and negative electrospray ionization modes (± ESI mode). The mobile phase was consisted of (A) 5

mM ammonium formate in water:MeOH (90:10, v/v) acidified with 0.01% formic acid and (B) 5 mM formate in MeOH acidified with 0.01% formic acid for the positive ionization mode. The mobile phase was consisted of (A) 10 mM ammonium acetate in water:MeOH l (90:10, v/v) and (B) 10 mM acetate in MeOH for the negative ionization mode. The elution program and the flow rate for the respective solvents were gradient and are presented in Table 7. A picture of the UHPLC-QTOF-ESI-MS system that was used is presented in Figure 20.

Table 7. The elution program and the flow rate for the solvents.

Time (min)	**Flow rate (mL min ^{1})**	**% A**	**% B**
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	0.2	99.0	1.0

Two acquisition MS modes were employed for the samples analysis, one data-independent fragmentation method and one data-dependent method. In bbCID mode, the Q-TOF/MS system was operated in broadband collision-induced dissociation (bbCID, data-independent) and spectra were recorded in the *m/z* range 50-1000, with scan rate 2 Hz. This function provides MS and MS/MS spectra simultaneously, with application of at least two collision energies. At low collision energy (4 eV), the full scan MS spectra were acquired. At high collision energy (25 eV), no isolation was performed on the quadrupole and the ions from the default region of mass were fragmented in the collision cell. The main advantage of this mode is that it provides high sensitivity to the MS function and that the chromatograms include sample overall profile ("fingerprint"). In autoMS mode, the five abundant ions per MS scan, are selected and fragmented. The applied collision energy was set to predetermined values, depending on the mass and charge of each ion. This method operation, provides a selective MS/MS spectrum for the compounds and is used

for confirmation of their structure, based on the identification of individual fragments. The Q-TOF/MS was daily calibrated with a mixture solution of isopropanol, ultrapure water, sodium hydroxide, formic acid). For internal calibration, one section is used in each chromatogram (0.1-0.25 min), where the calibration solution was injected before sample analysis. The theoretical exact masses of the calibration ions with types of Na $(NaCOOH)_{1\text{-}14}$ in the range 40-800 Da. The instrument provides a typical resolving power of 36000-40000 during the calibration.

The chromatograms are divided into 4 segments, as shown in Figure 21, where:

1) From 0-0.1 min: No MS logging (direct to waste)

2) From 0.1-0.25 minutes: Calibration of the MS with the calibration solution.

3) From 0.25-16 min: Separation of analytes (scan range: 50-1000 Da)

4) From 16-20 minutes: Cleaning – column re-equilibration

Data processing and evaluation was performed with *DataAnalysis 5.0* program and TASQ 1.4 (Bruker Daltonics, Bremen, Germany) program.

Figure 20. UHPLC-ESI-Q-TOF/MS instrument setup.

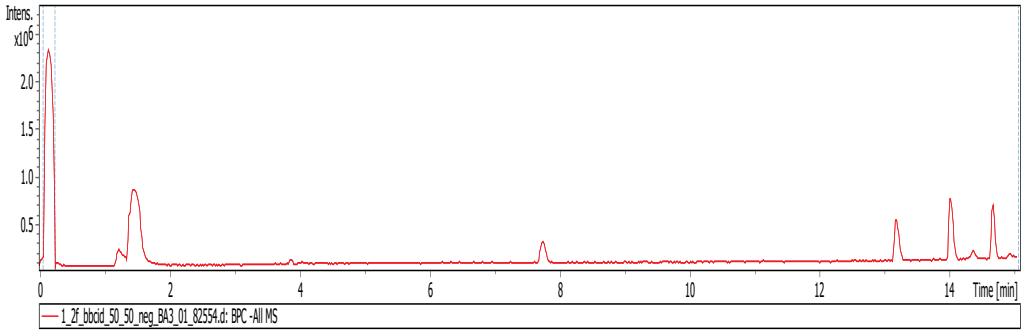


Figure 21. Example of full scan MS chromatogram of a commercial carob syrup sample.

3.1.6 Methodology

3.1.6.1 Identification confidence

There are five levels of identification confidence according to the Metabolomics Standards Initiative [56]:

Achieving definitive (level 1) identification necessitates the comparison of two or more independent properties. When MS/MS spectra match, the retention time and fragmentation mass spectrum have the same properties with the candidate and the corresponding reference standard observed under identical analytical conditions.

Putative (levels 2a, 2b, and 3) annotation primarily relies on the utilization of a single or a pair of properties. In order to achieve identification confidence level 2a, MS/MS spectra need to exhibit a degree of similarity to the reference spectra, with a similarity score exceeding 0.7. These can be found online from public libraries and compared to the experimental ones. To attain identification confidence level 2b, it is necessary to have potential structures available in cases where there is diagnostic evidence (e.g., detection of diagnostic ion). Furthermore, at confidence level 3, the determination of a plausible candidate can be achieved by evaluating MS/MS fragments using computational tools for in silico fragmentation, such as Met Frag [88] or CFM-ID [89].

When it comes to identification confidence level 4, the sole requirement is to verify the molecular formula of the candidate [90]. Conversely, at confidence level 5, confirmation can be established solely based on the absence of results for the corresponding molecular type, with the exact mass of interest serving as the determining factor [91].

3.1.6.2 Suspect screening methodology

A suspect list containing various chemical compounds was generated based on the literature [11, 98-102], employing of PubChem library. This list contained compounds

found in all parts of the carob tree, where their identification was achieved using the analytical methods of HRMS and NMR. The assembled list is presented in Table 8 containing information such as compound name, IUPAC name, molecular formula, molecular weight and $[M \pm H]^{\pm}$ exact mass ions and canonical smiles.

Table 8. Literature based suspect list.

	Compound Name	IUPAC Name	Molecular Formula (MF)	$[M+H]^+$	$[M-H]^-$	Molecular weight (MW) [g/mol]	Canonical SMILES
1	Gallic acid	3,4,5-trihydroxybenzoic acid	$C_6H_2(OH)_3COOH$	171.0288	169.0142	170.12	C1=C(C=C(C(=C1O)O)O)C(=O)O
2	Ellagic acid	6,7,13,14-tetrahydroxy-2,9-dioxatetracyclo [6.6.2.04,16.011,15] hexadeca-1(15),4,6,8(16),11,13-hexaene-3,10-dione	$C_{14}H_6O_8$	303.0135	300.9990	302.19	C1=C2C3=C(C(=C1O)O)OC(=O)C4=CC(=C(C(=C43)OC2=O)O)O
3	Isoquercetin (Quercetin glucoside)	2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl] oxychromen-4-one	$C_{21}H_{20}O_{12}$	465.1028	463.0882	464.40	C1=CC(=C(C=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)OC4C(C(C(C(O4)CO)O)O)O)O)O
4	Apigenin	5,7-dihydroxy-2-(4-hydroxyphenyl)chromen-4-one	$C_{15}H_{10}O_5$	271.0601	269.0455	270.24	C1=CC(=CC=C1C2=CC(=O)C3=C(C=C(C=C3O2)O)O)O
5	Naringenin	5,7-dihydroxy-2-(4-hydroxyphenyl)-2,3-dihydrochromen-4-one	$C_{15}H_{12}O_5$	273.0758	271.0612	272.25	C1C(OC2=CC(=CC(=C2C1=O)O)O)C3=CC=C(C=C3)O

6	Syringic acid	4-hydroxy-3,5-dimethoxybenzoic acid	$C_9H_{10}O_5$	199.0601	197.0455	198.17	COC1=CC(=CC(=C1O)OC)C(=O)O
7	Catechin hydrate	(2R,3S)-2-(3,4-dihydroxyphenyl)-3,4-dihydro-2H-chromene-3,5,7-triol; hydrate	$C_{15}H_{16}O_7$	309.0969	307.0823	308.28	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C=C3)O)O)O.O
8	Kaempferol	3,5,7-trihydroxy-2-(4-hydroxyphenyl)chromen-4-one	$C_{15}H_{10}O_6$	287.0550	285.0405	286.24	C1=CC(=CC=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O
9	Caffeic acid	(E)-3-(3,4-dihydroxyphenyl) prop-2-enoic acid	$C_9H_8O_4$	181.0495	179.0350	180.16	C1=CC(=C(C=C1C=CC(=O)O)O)O
10	Resorcinol	benzene-1,3-diol	$C_6H_6O_2$	111.0441	109.0295	110.11	C1=CC(=CC(=C1)O)O
11	Epicatechin-3-o-gallate	[(2R,3R)-2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3,4-dihydro-2H-chromen-3-yl] 3,4,5-trihydroxybenzoate	$C_{22}H_{18}O_{10}$	443.0973	441.0827	442.40	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C=C3)O)O)OC(=O)C4=CC(=C(C(=C4)O)O)O
12	p-Coumaric acid	(E)-3-(4-hydroxyphenyl) prop-2-enoic acid	$C_9H_8O_3$	165.0546	163.0401	164.16	C1=CC(=CC=C1C=CC(=O)O)O
13	Luteolin	2-(3,4-dihydroxyphenyl)-5,7-dihydroxychromen-4-one	$C_{15}H_{10}O_6$	287.0550	285.0405	286.24	C1=CC(=C(C=C1C2=CC(=O)C3=C(C=C(C=C3O2)O)O)O)O

14	Ferulic acid	(E)-3-(4-hydroxy-3-methoxyphenyl) prop-2-enoic acid	$C_{10}H_{10}O_4$	195.0652	193.0506	194.18	COC1=C(C=CC(=C1)C=CC(=O)O)O
15	Chlorogenic acid	(1S,3R,4R,5R)-3-[(E)-3-(3,4-dihydroxyphenyl) prop-2-enoyl] oxy-1,4,5-trihydroxycyclohexane-1-carboxylic acid	$C_{16}H_{18}O_9$	355.1024	353.0878	354.31	C1C(C(C(CC1(C(=O)O)O)OC(=O)C=CC2=CC(=C(C=C2)O)O)O)O
16	trans-Cinnamic acid	(E)-3-phenylprop-2-enoic acid	$C_9H_8O_2$	149.0597	147.0452	148.16	C1=CC=C(C=C1)C=CC(=O)O
17	Sinapic acid	(E)-3-(4-hydroxy-3,5-dimethoxyphenyl) prop-2-enoic acid	$C_{11}H_{12}O_5$	225.0758	223.0612	224.21	COC1=CC(=CC(=C1O)OC)C=CC(=O)O
18	Catechol	benzene-1,2-diol	$C_6H_6O_2$	111.0441	109.0295	110.11	C1=CC=C(C(=C1)O)O
19	Kaempferol-3-O-rutinoside	5,7-dihydroxy-2-(4-hydroxyphenyl)-3-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-[[(2R,3R,4R,5R,6S)-3,4,5-trihydroxy-6-methyloxan-2-yl] oxymethyl] oxan-2-yl] oxychromen-4-one	$C_{27}H_{30}O_{15}$	595.1657	593.1512	594.5	CC1C(C(C(C(O1)OCC2C(C(C(C(O2)OC3=C(OC4=CC(=CC(=C4C3=O)O)O)C5=CC=C(C=C5)O)O)O)O)O)O)O
20	Cirsimaritin	5-hydroxy-2-(4-hydroxyphenyl)-6,7-dimethoxychromen-4-one	$C_{17}H_{14}O_6$	315.0863	313.0718	314.29	COC1=C(C(=C2C(=C1)OC(=CC2=O)C3=CC=C(C=C3)O)O)OC
21	Luteolin-7-O-glucoside	2-(3,4-dihydroxyphenyl)-5-hydroxy-7-	$C_{21}H_{20}O_{11}$	449.1078	447.0933	448.4	C1=CC(=C(C=C1C2=CC(=O)C3=C(C=C(C=C3O2)OC4C(C(C(O4)CO)O)O)O)O)O

		[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl] oxychromen-4-one					
22	Epigallocatechin	(2R,3R)-2-(3,4,5-trihydroxyphenyl)-3,4-dihydro-2H-chromene-3,5,7-triol	$C_{15}H_{14}O_7$	307.0812	305.0667	306.27	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C(=C3)O)O)O)O
23	Myricetin 3-glucoside	5,7-dihydroxy-3-[(2S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl] oxy-2-(3,4,5-trihydroxyphenyl) chromen-4-one	$C_{21}H_{20}O_{13}$	481.0977	479.0831	480.40	C1=C(C=C(C(=C1O)O)O)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)OC4C(C(C(C(O4)CO)O)O)O
24	Catechin gallate	[(2S,3R)-2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3,4-dihydro-2H-chromen-3-yl] 3,4,5-trihydroxybenzoate	$C_{22}H_{18}O_{10}$	443.0973	441.0827	442.40	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C=C3)O)O)OC(=O)C4=CC(=C(C(=C4)O)O)O
25	Epigallocatechin gallate	[(2R,3R)-5,7-dihydroxy-2-(3,4,5-trihydroxyphenyl)-3,4-dihydro-2H-chromen-3-yl] 3,4,5-trihydroxybenzoate	$C_{22}H_{18}O_{11}$	459.0922	457.0776	458.40	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C(=C3)O)O)O)OC(=O)C4=CC(=C(C(=C4)O)O)O
26	(+)-Catechin	(2R,3S)-2-(3,4-dihydroxyphenyl)-3,4-dihydro-2H-chromene-3,5,7-triol	$C_{15}H_{14}O_6$	291.0863	289.0718	290.27	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C=C3)O)O)O

27	Betaine	2-(trimethylazaniumyl)acetate	$C_5H_{11}NO_2$	118.0862	116.0717	117.15	C[N+](C)(C)CC(=O)[O-]
28	γ-aminobutyric acid	4-aminobutanoic acid	$C_4H_9NO_2$	104.0706	102.0561	103.12	C(CC(=O)O)CN
29	Pyroglutamic acid	(2S)-5-oxopyrrolidine-2-carboxylic acid	$C_5H_7NO_3$	130.0499	128.0353	129.11	C1CC(=O)NC1C(=O)O
30	Cellobiose	(2R,3S,4S,5R,6S)-2-(hydroxymethyl)-6-[(2R,3S,4R,5R,6R)-4,5,6-trihydroxy-2-(hydroxymethyl) oxan-3-yl] oxyoxane-3,4,5-triol	$C_{12}H_{22}O_{11}$	343.1235	341.1089	342.30	C(C1C(C(C(C(C(O1)OC2C(OC(C(C2O)O)O)CO)O)O)O)O)O
31	Thiamine	2-[3-[(4-amino-2-methylpyrimidin-5-yl) methyl]-4-methyl-1,3-thiazol-3-ium-5-yl] ethanol	$C_{12}H_{17}N_4OS^+$	266.1196	264.1050	265.36	CC1=C(SC=[N+]1CC2=CN=C(N=C2N)C)CCO
32	Proline	(2S)-pyrrolidine-2-carboxylic acid	$C_5H_9NO_2$	116.0706	114.0561	115.13	C1CC(NC1)C(=O)O
33	Maltose	(2R,3S,4S,5R,6R)-2-(hydroxymethyl)-6-[(2R,3S,4R,5R,6R)-4,5,6-trihydroxy-2-(hydroxymethyl) oxan-3-yl] oxyoxane-3,4,5-triol	$C_{12}H_{22}O_{11}$	343.1235	341.1089	342.30	C(C1C(C(C(C(C(O1)OC2C(OC(C(C2O)O)O)CO)O)O)O)O)O
34	Sorbitol	(2R,3R,4R,5S)-hexane-1,2,3,4,5,6-hexol	$C_6H_{14}O_6$	183.0863	181.0718	182.17	C(C(C(C(C(CO)O)O)O)O)O

35	Sucrose	(2R,3R,4S,5S,6R)-2-[(2S,3S,4S,5R)-3,4-dihydroxy-2,5-bis(hydroxymethyl)oxolan-2-yl] oxy-6-(hydroxymethyl) oxane-3,4,5-triol	$C_{12}H_{22}O_{11}$	343.1235	341.1089	342.30	C(C1C(C(C(C(O1)OC2(C(C(C(O2)CO)O)O)CO)O)O)O)O
36	Lactose	(2R,3R,4S,5R,6S)-2-(hydroxymethyl)-6-[(2R,3S,4R,5R,6R)-4,5,6-trihydroxy-2-(hydroxymethyl) oxan-3-yl] oxyoxane-3,4,5-triol	$C_{12}H_{22}O_{11}$	343.1235	341.1089	342.30	C(C1C(C(C(C(O1)OC2C(OC(C(C2O)O)O)CO)O)O)O)O
37	Trehalose	(2R,3S,4S,5R,6R)-2-(hydroxymethyl)-6-[(2R,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl] oxyoxane-3,4,5-triol	$C_{12}H_{22}O_{11}$	343.1235	341.1089	342.30	C(C1C(C(C(C(O1)OC2C(C(C(C(O2)CO)O)O)O)O)O)O)O
38	Theobromine	3,7-dimethylpurine-2,6-dione	$C_7H_8N_4O_2$	181.0720	179.0574	180.16	CN1C=NC2=C1C(=O)NC(=O)N2C
39	Melibiose	(3R,4S,5S,6R)-6-[[(2S,3R,4S,5R,6R)-3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl] oxymethyl]oxane-2,3,4,5-tetrol	$C_{12}H_{22}O_{11}$	343.1235	341.1089	342.30	C(C1C(C(C(C(O1)OCC2C(C(C(C(O2)O)O)O)O)O)O)O)O
40	Benzoic acid	benzoic acid	$C_7H_6O_2$	123.0441	121.0295	122.12	C1=CC=C(C=C1)C(=O)O

41	Alanine	(2S)-2-aminopropanoic acid	$C_3H_7NO_2$	90.0550	88.0404	89.09	CC(C(=O)O)N
42	Leucine	(2S)-2-amino-4-methylpentanoic acid	$C_6H_{13}NO_2$	132.1019	130.0874	131.17	CC(C)CC(C(=O)O)N
43	Phenylalanine	(2S)-2-amino-3-phenylpropanoic acid	$C_9H_{11}NO_2$	166.0863	164.0717	165.19	C1=CC=C(C=C1)CC(C(=O)O)N
44	Valine	(2S)-2-amino-3-methylbutanoic acid	$C_5H_{11}NO_2$	118.0863	116.0717	117.15	CC(C)C(C(=O)O)N
45	Lactic acid	2-hydroxypropanoic acid	$C_3H_6O_3$	91.0390	89.0244	90.08	CC(C(=O)O)O
46	Monoisoamylamine	3-methylbutan-1-amine	$C_5H_{13}N$	88.1121	86.0975	87.16	CC(C)CCN
47	Choline	2-hydroxyethyl(trimethyl)azanium	$C_5H_{14}NO^+$	105.1148	103.1003	104.17	C[N+](C)(C)CCO
48	Nicotinamide	pyridine-3-carboxamide	$C_6H_6N_2O$	123.0553	121.0407	122.12	C1=CC(=CN=C1)C(=O)N
49	Nicotinic acid	pyridine-3-carboxylic acid	$C_6H_5NO_2$	124.039305	122.0248	123.11	C1=CC(=CN=C1)C(=O)O
50	Uridine	1-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl) oxolan-2-yl] pyrimidine-2,4-dione	$C_9H_{12}N_2O_6$	245.0768	243.0623	244.20	C1=CN(C(=O)NC1=O)C2C(C(C(O2)CO)O)O
51	Guanine	2-amino-1,7-dihydropurin-6-one	$C_5H_5N_5O$	152.0567	150.0421	151.13	C1=NC2=C(N1)C(=O)NC(=N2)N

52	Fructose	(3S,4R,5R)-2-(hydroxymethyl) oxane-2,3,4,5-tetrol	$C_6H_{12}O_6$	181.0707	179.0561	180.16	C1C(C(C(C(O1)(CO)O)O)O)O
53	Glucose	(3R,4S,5S,6R)-6-(hydroxymethyl) oxane-2,3,4,5-tetrol	$C_6H_{12}O_6$	181.0707	179.0561	180.16	C(C1C(C(C(C(O1)O)O)O)O)O
54	Adenosine	(2R,3R,4S,5R)-2-(6-aminopurin-9-yl)-5-(hydroxymethyl) oxolane-3,4-diol	$C_{10}H_{13}N_5O_4$	268.1040	266.0895	267.24	C1=NC(=C2C(=N1)N(C=N2)C3C(C(C(O3)CO)O)O)N
55	Pyridoxine	4,5-bis(hydroxymethyl)-2-methylpyridin-3-ol	$C_8H_{11}NO_3$	170.0812	168.0666	169.18	CC1=NC=C(C(=C1O)CO)CO
56	Acetyl-carnitine	(3R)-3-acetyloxy-4-(trimethylazaniumyl)butanoate	$C_9H_{17}NO_4$	204.1230	202.1085	203.24	CC(=O)OC(CC(=O)[O-])C[N+](C)(C)C
57	a-Hydroxyisovaleric acid	2-hydroxy-3-methylbutanoic acid	$C_5H_{10}O_3$	119.0703	117.0557	118.13	CC(C)C(C(=O)O)O
58	Itaconic acid	2-methylidenebutanedioic acid	$C_5H_6O_4$	131.0339	129.0193	130.10	C=C(CC(=O)O)C(=O)O
59	Hypoxanthine	1,7-dihydropurin-6-one	$C_5H_4N_4O$	137.0458	135.0312	136.11	C1=NC2=C(N1)C(=O)NC=N2
60	Xanthine	3,7-dihydropurine-2,6-dione	$C_5H_4N_4O_2$	153.0407	151.0261	152.11	C1=NC2=C(N1)C(=O)NC(=O)N2
61	Ribose	(3R,4R,5R)-oxane-2,3,4,5-tetrol	$C_5H_{10}O_5$	151.0601	149.0455	150.13	C1C(C(C(C(O1)O)O)O)O

62	Xylose	(3R,4S,5R)-oxane-2,3,4,5-tetrol	$C_5H_{10}O_5$	151.0601	149.0455	150.13	C1C(C(C(C(O1)O)O)O)O
63	Ascorbic acid	(2R)-2-[(1S)-1,2-dihydroxyethyl]-3,4-dihydroxy-2H-furan-5-one	$C_6H_8O_6$	177.0394	175.0248	176.12	C(C(C1C(=C(C(=O)O1)O)O)O)O
64	Uracil	1H-pyrimidine-2,4-dione	$C_4H_4N_2O_2$	113.0346	111.0200	112.09	C1=CNC(=O)NC1=O
65	Mannose	(3S,4S,5S,6R)-6-(hydroxymethyl) oxane-2,3,4,5-tetrol	$C_6H_{12}O_6$	181.0707	179.0561	180.16	C(C1C(C(C(C(O1)O)O)O)O)O
66	Cotinine	(5S)-1-methyl-5-pyridin-3-ylpyrrolidin-2-one	$C_{10}H_{12}N_2O$	177.1022	175.0877	176.21	CN1C(CCC1=O)C2=CN=CC=C2
67	Cytidine	4-amino-1-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl) oxolan-2-yl] pyrimidin-2-one	$C_9H_{13}N_3O_5$	244.0928	242.0782	243.22	C1=CN(C(=O)N=C1N)C2C(C(C(O2)CO)O)O
68	Thymidine	1-[(2R,4S,5R)-4-hydroxy-5-(hydroxymethyl) oxolan-2-yl]-5-methylpyrimidine-2,4-dione	$C_{10}H_{14}N_2O_5$	243.0975	241.0830	242.23	CC1=CN(C(=O)NC1=O)C2CC(C(O2)CO)O
69	Pyrocatechol	benzene-1,2-diol	$C_6H_6O_2$	111.0441	109.0295	110.11	C1=CC=C(C(=C1)O)O
70	2,5-Dihydroxybenzoic acid	2,5-dihydroxybenzoic acid	$C_7H_6O_4$	155.0339	153.0193	154.12	C1=CC(=C(C=C1O)C(=O)O)O

71	Chrysoeriol	5,7-dihydroxy-2-(4-hydroxy-3-methoxyphenyl) chromen-4-one	$C_{16}H_{12}O_6$	301.0707	299.0561	300.26	COC1=C(C=CC(=C1)C2=CC(=O)C3=C(C=C(C=C3O2)O)O)O
72	Eriodictyol	(2S)-2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-2,3-dihydrochromen-4-one	$C_{15}H_{12}O_6$	289.0707	287.0561	288.25	C1C(OC2=CC(=CC(=C2C1=O)O)O)C3=CC(=C(C=C3)O)O
73	Genistein	5,7-dihydroxy-3-(4-hydroxyphenyl)chromen-4-one	$C_{15}H_{10}O_5$	271.0601	269.0455	270.24	C1=CC(=CC=C1C2=COC3=CC(=CC(=C3C2=O)O)O)O
74	Isorhamnetin	3,5,7-trihydroxy-2-(4-hydroxy-3-methoxyphenyl) chromen-4-one	$C_{16}H_{12}O_7$	317.0656	315.0510	316.26	COC1=C(C=CC(=C1)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O
75	Kaempferol 7-O-rhamnoside	3,5-dihydroxy-2-(4-hydroxyphenyl)-7-[(2S,3R,4R,5R,6S)-3,4,5-trihydroxy-6-methyloxan-2-yl] oxychromen-4-one	$C_{21}H_{20}O_{10}$	433.1129	431.0984	432.40	CC1C(C(C(C(O1)OC2=CC(=C3C(=C2)OC(=C(C3=O)O)C4=CC=C(C=C4)O)O)O)O)O
76	Myricetin rhamnoside	5,7-dihydroxy-3-(3,4,5-trihydroxy-6-methyloxan-2-yl) oxy-2-(3,4,5-trihydroxyphenyl) chromen-4-one	$C_{21}H_{20}O_{12}$	465.1028	463.0882	464.40	CC1C(C(C(C(O1)OC2=C(OC3=CC(=CC(=C3C2=O)O)O)C4=CC(=C(C(=C4)O)O)O)O)O)O
77	Tricetin 3',4' dimethyl ether	5,7-dihydroxy-2-(3-hydroxy-4,5-dimethoxyphenyl) chromen-4-one	$C_{17}H_{14}O_7$	331.0812	329.0667	330.29	COC1=CC(=CC(=C1OC)O)C2=CC(=O)C3=C(C=C(C=C3O2)O)O

78	Myricetin-pentoside	5,7-dihydroxy-3-(3,4,5-trihydroxyoxan-2-yl) oxy-2-(3,4,5-trihydroxyphenyl) chromen-4-one	$C_{20}H_{18}O_{12}$	451.0871	449.0726	450.30	C1C(C(C(C(O1)OC2=C(OC3=CC(=CC(=C3C2=O)O)O)C4=CC(=C(C(=C4)O)O)O)O)O)O
79	Arginine	(2S)-2-amino-5-(diaminomethylideneamino)pentanoic acid	$C_6H_{14}N_4O_2$	175.1190	173.1044	174.20	C(CC(C(=O)O)N)CN=C(N)N
80	Aspartic acid (Asp)	(2S)-2-aminobutanedioic acid	$C_4H_7NO_4$	134.0448	132.0302	133.10	C(C(C(=O)O)N)C(=O)O
81	Cysteine (Cys)	(2R)-2-amino-3-sulfanylpropanoic acid	$C_3H_7NO_2S$	122.0270	120.0125	121.16	C(C(C(=O)O)N)S
82	Glycine (Gly)	2-aminoacetic acid	$C_2H_5NO_2$	76.0393	74.0248	75.07	C(C(=O)O)N
83	Glutamic acid (Glu)	(2S)-2-aminopentanedioic acid	$C_5H_9NO_4$	148.0604	146.0459	147.13	C(CC(=O)O)C(C(=O)O)N
84	L-Histidine (His)	(2S)-2-amino-3-(1H-imidazol-5-yl) propanoic acid	$C_6H_9N_3O_2$	156.0768	154.0622	155.15	C1=C(NC=N1)CC(C(=O)O)N
85	L-isoleucine (Ile)	(2S,3S)-2-amino-3-methylpentanoic acid	$C_6H_{13}NO_2$	132.1019	130.0874	131.17	CCC(C)C(C(=O)O)N
86	L-Lysine (Lys)	(2S)-2,6-diaminohexanoic acid	$C_6H_{14}N_2O_2$	147.1128	145.0983	146.19	C(CCN)CC(C(=O)O)N
87	L-Methionine (Met)	(2S)-2-amino-4-methylsulfanylbutanoic acid	$C_5H_{11}NO_2S$	150.0583	148.0438	149.21	CSCCC(C(=O)O)N

88	L-Serine (Ser)	(2S)-2-amino-3-hydroxypropanoic acid	$C_3H_7NO_3$	106.0499	104.0353	105.09	C(C(C(=O)O)N)O
89	L-Threonine (Thr)	(2S,3R)-2-amino-3-hydroxybutanoic acid	$C_4H_9NO_3$	120.0655	118.0600	119.12	CC(C(C(=O)O)N)O
90	L-Tryptophan (Trp)	(2S)-2-amino-3-(1H-indol-3-yl) propanoic acid	$C_{11}H_{12}N_2O_2$	205.0972	203.0826	204.22	C1=CC=C2C(=C1)C(=CN2)CC(C(=O)O)N
91	L-Tyrosine (Tyr)	(2S)-2-amino-3-(4-hydroxyphenyl) propanoic acid	$C_9H_{11}NO_3$	182.0812	180.0666	181.19	C1=CC(=CC=C1CC(C(=O)O)N)O
92	Quercetin 7-(rhamnosylglucoside)	7-[(2S,3R,4S,5S,6R)-4,5-dihydroxy-6-(hydroxymethyl)-3-[(2S,3R,4R,5R,6S)-3,4,5-trihydroxy-6-methyloxan-2-yl] oxyoxan-2-yl] oxy-2-(3,4-dihydroxyphenyl)-3,5-dihydroxychromen-4-one	$C_{27}H_{30}O_{16}$	611.1607	609.1461	610.50	CC1C(C(C(C(O1)OC2C(C(C(OC2OC3=CC(=C4C(=C3)OC(=C(C4=O)O)C5=CC(=C(C=C5)O)O)O)CO)O)O)O)O)O
93	D-pinitol	(1*S*,2*S*,4*S*,5*R*)-6-methoxycyclohexane-1,2,3,4,5-pentol	$C_7H_{14}O_6$	195.0863	193.0707	194.18	COC1C(C(C(C(C1O)O)O)O)O
94	m-coumaric acid	(*E*)-3-(3-hydroxyphenyl) prop-2-enoic acid	$C_9H_8O_3$	165.0546	163.0401	164.16	C1=CC(=CC(=C1)O)C=CC(=O)O
95	Myricetin	3,5,7-trihydroxy-2-(3,4,5-	$C_{15}H_{10}O_8$	319.0448	317.0292	318.23	C1=C(C=C(C(=C1O)O)O)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O

		trihydroxyphenyl)chromen-4-one					
96	Quercetin	2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxychromen-4-one	$C_{15}H_{10}O_7$	303.0499	301.0343	302.23	C1=CC(=C(C=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O)O
97	Protocatechuic acid	3,4-dihydroxybenzoic acid	$C_7H_6O_4$	155.0339	153.0182	154.12	C1=CC(=C(C=C1C(=O)O)O)O
98	Hydroxytyrosol	4-(2-hydroxyethyl)benzene-1,2-diol	$C_8H_{10}O_3$	155.0703	153.0546	154.16	C1=CC(=C(C=C1CCO)O)O
99	4-hydroxybenzoic acid	4-hydroxybenzoic acid	$C_7H_6O_3$	139.0390	137.0233	138.12	C1=CC(=CC=C1C(=O)O)O
100	(Iso)schaftoside-4′-*O*-glucoside	5,7-dihydroxy-8-[(2*S*,4*R*,5*S*)-3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl]-2-[4-[(2*S*,4*R*,5*S*)-3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl]oxyphenyl]-6-[(2*S*,4*S*,5*S*)-3,4,5-trihydroxyoxan-2-yl]chromen-4-one	$C_{32}H_{38}O_{19}$	727.2080	725.1919	726.60	C1C(C(C(C(O1)C2=C(C(=C3C(=C2O)C(=O)C=C(O3)C4=CC=C(C=C4)OC5C(C(C(C(O5)CO)O)O)O)C6C(C(C(C(O6)CO)O)O)O)O)O)O)O
101	Trigalloyl glucose	[(2*R*,3*R*,4*S*,5*R*,6*S*)-3,5-dihydroxy-4,6-bis[(3,4,5-trihydroxybenzoyl) oxy] oxan-2-yl] methyl	$C_{27}H_{24}O_{18}$	637.1035	635.0869	636.50	C1=C(C=C(C(=C1O)O)O)C(=O)OCC2C(C(C(C(O2)OC(=O)C3=CC(=C(C(=C3)O)O)O)O)OC(=O)C4=CC(=C(C(=C4)O)O)O)O

		3,4,5-trihydroxybenzoate					
102	Gallocatechin	[(2*S*,3*R*)-5,7-dihydroxy-2-(3,4,5-trihydroxyphenyl)-3,4-dihydro-2*H*-chromen-3-yl] 3,4,5-trihydroxybenzoate	$C_{22}H_{18}O_{11}$	459.0922	457.0770	458.40	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C(=C3)O)O)O)OC(=O)C4=CC(=C(C(=C4)O)O)O
103	Schaftoside (Apigenin-6-*C*-glucoside-8-*C*-arabinoside)		$C_{26}H_{28}O_{14}$	565.1552	563.1401	564.08	
104	Tetragllayol glucose	[(2*R*,3*R*,4*S*,5*R*,6*S*)-3-hydroxy-4,5,6-tris[(3,4,5-trihydroxybenzoyl) oxy] oxan-2-yl] methyl 3,4,5-trihydroxybenzoate	$C_{34}H_{28}O_{22}$	789.1145	787.1002	788.60	C1=C(C=C(C(=C1O)O)O)C(=O)OCC2C(C(C(C(O2)OC(=O)C3=CC(=C(C(=C3)O)O)O)OC(=O)C4=CC(=C(C(=C4)O)O)O)OC(=O)C5=CC(=C(C(=C5)O)O)O)O
105	Chrysoeriol-*O*-deoxyheoxoside	5-hydroxy-2-(4-hydroxy-3-methoxyphenyl)-7-[(2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl] oxychromen-4-one	$C_{22}H_{22}O_{11}$	463.1235	461.1082	462.40	COC1=C(C=CC(=C1)C2=CC(=O)C3=C(C=C(C=C3O2)OC4C(C(C(C(O4)CO)O)O)O)O)O
106	Kampferol-*O*-deoxyhexoside	3,5-dihydroxy-2-(4-hydroxyphenyl)-7-[(2*S*,3*R*,4*R*,5*R*,6*S*)-3,4,5-trihydroxy-6-methyloxan-2-yl] oxychromen-4-one	$C_{21}H_{20}O_{10}$	433.1129	431.0976	432.40	CC1C(C(C(C(O1)OC2=CC(=C3C(=C2)OC(=C(C3=O)O)C4=CC=C(C=C4)O)O)O)O)O

107	Trihydroxy flavone (Apigenin isomer)	3,5,7-trihydroxy-2-phenylchromen-4-one	$C_{15}H_{10}O_5$	271.0601	269.0453	270.24	C1=CC=C(C=C1)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O
108	Kampferide	3,5,7-trihydroxy-2-(4-methoxyphenyl)chromen-4-one	$C_{16}H_{12}O_6$	301.0707	299.0556	300.26	COC1=CC=C(C=C1)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O
109	Dihydroxypalmitic acid	9,10-dihydroxyhexadecanoic acid	$C_{16}H_{32}O_4$	289.2373	287.2224	288.42	CCCCCCC(C(CCCCCCCC(=O)O)O)O
110	Dihydroxyoctadecatrienoic acid	(2*E*,4*E*,6*Z*)-6,13-dihydroxyoctadeca-2,4,6-trienoic acid	$C_{18}H_{32}O_4$	313.2373	311.2223	310.40	CCCCCC(CCCCCC=C(C=CC=CC(=O)O)O)O
111	Hydroxy oleic acid	(*Z*,2*R*)-2-hydroxyoctadec-9-enoic acid	$C_{18}H_{34}O_3$	299.2581	297.2430	298.50	CCCCCC(CCCCCC=C(C=CC=CC(=O)O)O
112	Linoleic acid	(9*Z*,12*Z*)-octadeca-9,12-dienoic acid	$C_{18}H_{32}O_2$	281.2475	279.2320	280.40	CCCCCC=CCC=CCCCCCCCC(=O)O
113	Palmitic acid	hexadecanoic acid	$C_{16}H_{32}O_2$	257.2475	255.2325	256.42	CCCCCCCCCCCCCCCC(=O)O
114	Oleic acid	(*Z*)-octadec-9-enoic acid	$C_{18}H_{34}O_2$	283.2632	281.2482	282.50	CCCCCCCCC=CCCCCCCCC(=O)O

The software *IsotopePattern* (Bruker, Bermen Germany) was utilized to generate $[M+H]^+$ ions for +ESI mode and $[M-H]^-$ ions for -ESI mode, along with their respective isotope patterns. Generally, when conducting suspect screening of a sample, the initial step involves aiming for identification confidence level 3. However, it is crucial to perform molecular formula evaluation. The assignment of molecular formulas was carried out using the Smart Formula Manually software (Bruker Daltonics, Bremen, Germany) with a mass tolerance of 2 mDa. At confidence level 3, the MS/MS fragments of plausible candidates were analyzed by employing *in silico* fragmentation tools, such as MetFrag. Subsequently, the MS/MS fragments of these candidates were compared to reference spectra using online databases. The comparison was based on a similarity score (above 0.7), leading to the attainment of identification confidence level 2a.

3.1.6.3 Quantification approach

Metabolites quantification was performed for compounds where their reference standards were available. Intermediate working solution of varied concentration was prepared in methanol: water in a ratio of 50:50 (v/v). Hesperitin (C=8.0 mg kg^{-1}), which did not exist in carob syrup samples, was used as an internal standard. The use of the internal standard is carried out to correct for possible variation resulting either from sample preparation or the instability of the instrument.

In addition, calibration curves were constructed for the analytes of interest using the normalized peak areas, which were calculated by dividing the peak area of each analyte with the peak area of the internal standard.

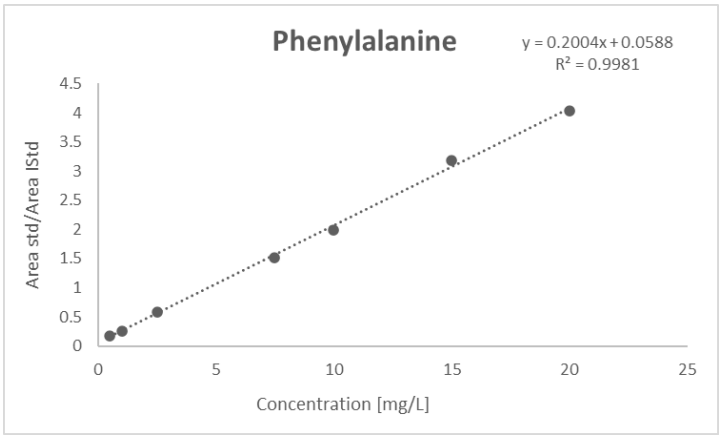
3.1.6.4 Calibration curves

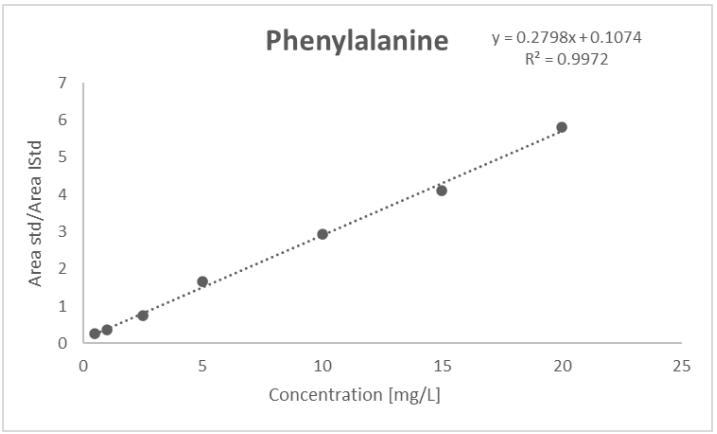
The calibration curves were constructed for the analytes of interest, which were identified in the Cypriot carob syrup samples, using (a) reference substances of high analytical purity (Table 5) diluted with methanol and HCl 0.1 M or HCl 1M depending on the analyte and (b) calibration matrix-matched curves diluted in the most dilute sample (matrix) in various concentrations of 0.5, 1, 2.5, 5, 7.5, 10, 15 and 20 mg L^{-1}, to achieve the desired concentrations of the analytes. The matrix-matched calibration corrects the signal suppression or enhancement due to the matrix components.

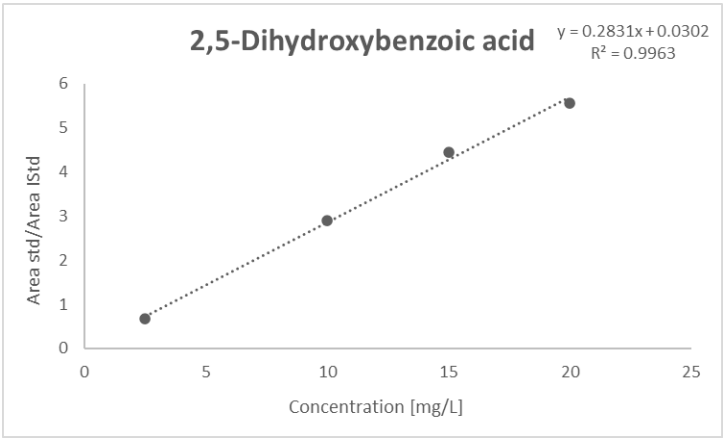
The linearity of the calibration curves for all the analytes was satisfactory, with coefficients of determination (R^2) ranging from 0.991 to 0.9990 indicatively. In Figure 22 the calibration

curves of reference standards of phenylalanine and 2,5-dihydroxybenzoic acid are presented. In Table 9 the respective Anova details are shown.

Standards curves	Matrix-matched curves







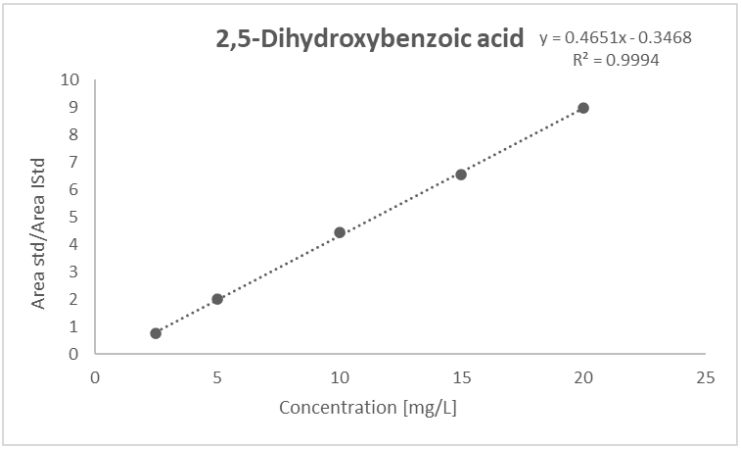


Figure 22. Standard calibration curves and matrix matched curves of phenylalanine and 2,5-dihydroxybenzoic acid.

Table 9. The Anova details of standard calibration and matrix matched curves of phenylalanine and 2,5-dihydroxybenzoic acid.

Standard calibration curves		
Anova details	Phenylalanine	2,5-Dihydroxybenzoic acid
Intercept	$5.9*10^{-2}$	$3*10^{-2}$
Standard Deviation	$4.1*10^{-2}$	$17*10^{-2}$
Intercept of the slope	$2.0*10^{-2}$	$28.3*10^{-2}$

Standard Deviation of the slope	$0.39*10^{-2}$	$1.2*10^{-2}$
t-Stat	1.44	0.18
P-value	0.21	0.87
Matrix matched curves		
Anova details	Phenylalanine	2,5-Dihydroxybenzoic acid
Intercept	$10.7*10^{-2}$	$34.7*10^{-2}$
Standard Deviation	$6.9*10^{-2}$	$8.0*10^{-2}$
Intercept of the slope	$0.7*10^{-2}$	$46.5*10^{-2}$
Standard Deviation of the slope	$28*10^{-2}$	$0.65*10^{-2}$
t-Stat	1.55	4.34
P-value	0.18	0.023

3.1.6.5 Identification workflow

The MS DIAL (4.92) software was used to annotate the compounds of the suspect list.

MS-DIAL is considered a valuable tool for leveraging DIA results and DDA spectra. Offline annotation of compounds is accomplished by comparing the deconvoluted spectra from DDA or DIA with corresponding spectra from the samples. MS-DIAL conducts search against pre-existing local libraries, such as the comprehensive ESI (+/-)-MS/MS list compiled from authentic standards, which is available for download through the software's page.

The replacement of the abovementioned library from a narrowed version, encompassing only specific compounds of interest (suspect list), based on a previous work of our group [93]. Hence, this flexibility permits the creation of customized libraries that can effectively reduce the number of potential candidates. The MSP files, which can be edited using a simple text editor like Notepad, contain entries containing the names of the candidates, their molecular formulas, exact masses, theoretical retention times, and MS/MS fragments. These files are publicly accessible through various sources such as GNPS, MS-DIAL, and other webpages. The MSP files were modified to focus specifically on the analytes of interest, thus serving as a targeted database or suspect list. This approach offers the added advantage of a more comprehensive perspective for spectral comparison. Importantly, the fragments included in this database are derived from experimental spectra, avoiding the common issue of using biased or curated fragments that often arise when compiling suspect lists. These files are compatible with the MS-DIAL software, enabling the processing of both DDA and DIA data. One novel list has been generated and imported to MS-DIAL. The

chemical characterization was searched with the aid of this suspect list, aiming to find the metabolites in the carob syrup.

3.1.6.6 Traceability methods

If the analyzed signal is not greater than the blank signal (S_{bl}), due to random errors, the signal cannot be detected with certainty. The minimum analytical signal (S_m) that can be detected is the sum of the mean blank signal ($\bar{S}_{bl}$) and the multiple (k=3) of the standard deviation of the blank (S_{bl}). The minimum analytical signal is given to Equation 1.

$$Sm = \bar{S}bl + ks_{bl} \quad \text{(Equation 1)}$$

From the slope (m) of Equation 1, S_m is converted to cm, which is the detection limit and is calculated in Equation 2.

$$Cm = \frac{Sm - \overline{S_{bl}}}{m} = k\frac{S_{bl}}{m} \quad \text{(Equation 2)}$$

Limit of Detection (LOD) is the lowest concentration of the analyte when detection is possible. This value (S) is usually 3 times higher than the noise level (N) of the analytical method. The detection limit is calculated in Equations 3 and 4.

$$Cm = \frac{3S_{bl}}{m} \quad \text{(Equation 3)}$$

$$LOD = 3.3\ x\ \frac{S}{N} = 3.3\ x\ \frac{SD}{b} \quad \text{(Equation 4)}$$

SD = Standard deviation

b = Calibration curve's slope

Limit of Quantitation (LOQ) is the lowest concentration of the analyte that can be quantified with accuracy and precision, as it shown in Equation 5.

$$LOQ = 10\ x\ \frac{S}{N} = 10\ x\ \frac{SD}{b} \quad \text{(Equation 5)}$$

3.2 NMR

3.2.1 Reagents and Standards

All standards and reagents were of high-purity grade (>95%) as follows:

- Deuterated solvent(D_2O) 99.8% Euriso-Top GmbH, Saarbrücken, Germany
- MeOD-d4 (99.8% D, Euriso-Top GmbH, Saarbrücken, Germany) with 0.02% v/v HMDSO (NMR grade, ≥ 99.5%, Sigma-Aldrich Corporation, St. Louis, MO, USA)

Consumables:

- 5 mm NMR tubes (D400-5-7) with a PTFE cap (both obtained by Deutero GmbH, Kastellaun, Germany).
- Commercial test tube containing 2 mM sucrose added with 0.5 mM DSS, 2 mM NaN_3 in 90–10% (v/v) H_2O-D_2O

Apparatus:

- Ultrasonic bath (Ultrasons H-D, J.P. selecta)
- Centrifugal Hettich Rotofix 32A

Analytical instruments:

The NMR spectra were acquired on a Bruker Avance NEO 500 MHz (Figure 23) system (magnet ASCEND 11.76 Tesla) equipped with a 5 mm Prodigy cryoprobe (BBO/H-F) detection probe. Experiments were performed in automation mode, using a SampleCase-24 sample changer operated by IconNMR. Data acquisition and processing were done with Bruker TopSpin 4.1.4. Metabolic profiling 1D NMR spectra were acquired using water suppression at the 1D NOESY pulse program with the following settings: 90°pulse; relaxation delay (d1) = 10 s; Acquisition time (AQ) = 3.28 s; FID (free induction decay) data points = 64k; spectral width (SW) = 20 ppm; Number of scans (ns) = 128. The transmitter offset was set manually in order to achieve optimal suppression of the residual water signal. FIDs were zero filled and multiplied by an exponential weighting function corresponding to a line broadening of 0.3 Hz prior to Fourier transformation. Chemical shift values were referenced to the residual TSP (0.00 ppm).

3.2.2 Sample pretreatment

Carobs were processed following procedure described in detail at sub-sections 3.1.4.1 and 3.1.4.2. The NMR samples are accordingly coded to distinguished from the UHPLC-QTOF-MS samples of Table 6. This information is illustrated in Table 10.

Table 10. NMR coding of carob samples' preparation in different temperatures.

Sample/Region	**Saint Fyla** **(altitude 100 m)**	**Dieronas** **(altitude 460 m)**	**Klonari** **(Altitude 480 m)**
Sample 1	Water extraction for 36-40 h at 20 °C and filtration **NMR code: B1**	Water extraction for 36-40 h at 20 °C and filtration **NMR code: D1on**	Water extraction for 36-40 h at 20 °C and filtration **NMR code: G1**

Sample 2 (5’ boiling)	Boiling sample 1 at 80 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: B2**	Boiling sample 1 at 80 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: D2on**	Boiling sample 1 at 80 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: G2**
Sample 3 (10’ boiling)	Boiling sample 2 at 82 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: B3**	Boiling sample 2 at 88 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: D3on**	Boiling sample 2 at 90-100 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: G3**
Sample 4 (15’ boiling)	Boiling sample 3 at 82 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: B4**	Boiling sample 3 at 88 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: D4on**	Boiling sample 3 at 90-100 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: G4**
Sample 5 (20’ boiling)	Boiling sample 4 at 82 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: B5**	Boiling sample 4 at 90-100 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: D5on**	Boiling sample 4 at 90-100 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: G5**
Sample 6 (25’ boiling)	Boiling sample 5 at 82 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: B6**	Boiling sample 5 at 95 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: D6on**	Boiling sample 5 at 96 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: G6**
Sample 7 (30’ boiling)			Boiling sample 6 at 96 °C for 5’, water extraction for 36-40 h, and filtration **NMR code: G7**

3.2.3 Sample preparation

1. Weighing a quantity of 100 mg (accuracy 4 decimal places) carob syrup samples.

If diluted too much it provides a low number of detected resonances but extracts with lower concentrations is easier for the shimming. On the other hand, if you concentrated too much it will be more difficult to shim with broader resonances due to higher viscosity.

2. The samples were further diluted with 600 µL deuterated solvent (D_2O) with an electronical pipette to distribute the accurate volume.
3. Homogenizing for 5 min with vortex.
4. Sonication (30 min at 30 °C).
5. Centrifugating to remove solid parts (10 min at 4 °C, 40.000 rpm).
6. Weighing on an electric balance of a quantity 0.003 g of 3-trimethylsilylpropanoic-2,2,3,3-d4 acid sodium salt (TMSP) and adding it in the exacts.

   The amount of TMSP was adjusted according to the sample type for chemical shift calibration. The transparent extracts were transferred to NMR tubes. High-quality tubes were re-used; these were prior cleaned and dried carefully. A constant volume was transferred to the NMR tube (600 µL for a 5 mm tube). One blank extraction was performed in the preparation workflow.

Figure 23. NMR instrument setup.

3.2.4 Sample analysis

3.2.4.1 NMR acquisition and instrument preparation

The NMR instrument was prepared before profiling of a sample series with three types of NMR tubes:

a) a calibration sample of MeOD-d4 for temperature calibration
b) a sucrose test-tube
c) a blank-extract tube for optimization of shim matrices.

The first step was that of temperature calibration. The temperature of the NMR spectrometer was checked and calibrated using a MeOD-d4 calibration sample. A stable temperature in the NMR probe is quite important because it will prevent variations of chemical shifts, especially in the water peak for water suppression. The stability of temperature was checked over time in the NMR probe and NMR spectra acquired at a temperature of 27 °C (300±0.03K). To ensure that acquisition was performed at the right temperature, a period of 5 min delay was followed in the magnet for temperature equilibration before spectra acquisition.

The following step was to run a 3D shimming with a commercial test tube containing 2 mM Sucrose. NMR spectrometers are certified annually but running a 3D shimming before NMR acquisition campaign with a commercial test tube containing 2 mM sucrose added with 0.5 mM 4,4-Dimethyl-4-silapentane-1-sulfonic acid (DSS) and 2 mM NaN_3 in 90–10% (v/v) H_2O-D_2O, is requisite to calibrate water suppression and spectrometer sensitivity. Then Shim matrices were optimized on the blank-extract tube and the spectral quality of this extract should be identical to that of the test-tube used for the specifications of the device (full width at half maximum (FWHM) of TMSP<0.65 Hz.

The NMR tubes were placed in the auto sampler for 24 h (stability assumed at room temperature). To standardize the acquisition, an automation routine was followed, which includes temperature homogenization, probe tuning and matching, locking solvent, shimming, and 90° calibration pulse determination, as presented in Figure 24.

After that, ^1H spectra (1D) spectra acquisition was checked for spectra quality to re-record or remove sample spectra that didn't meet quality criteria. It was assessed using the FWHM. Because of TMSP added as a spectral chemical shift reference in the extracts, it used to assess the quality of the spectrum. A FWHM of 0.65 Hz for TMSP, with a line broadening (lb) of 0.3 at a temperature of 27 °C, gives an excellent signal quality.

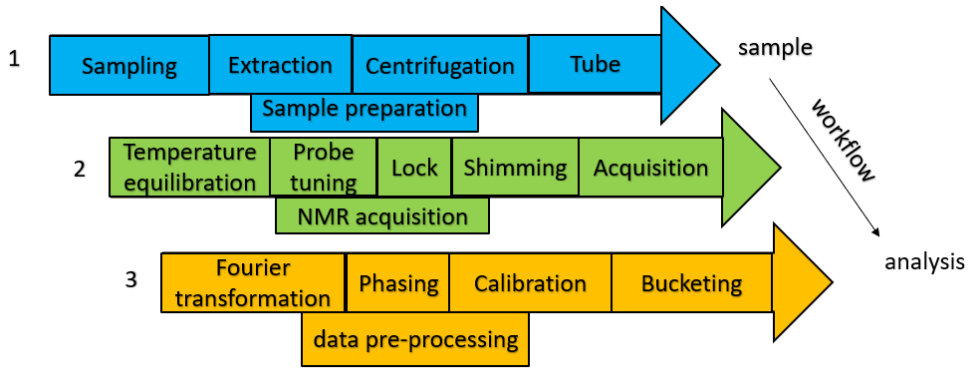


Figure 24. H-NMR metabolomic workflow of carob syrup extracts: 1) sample preparation; 2) NMR acquisition; 3) data pre-processing [92].

3.2.4.2 Workflow for data pre-processing

The spectra were obtained by the Fourier transformation (FT) of the free induction decay (FID) by applying an exponential multiplication with a line-broadening factor (lb) of 0.3 Hz and zero-filling (size = 128K) procedure. The resulting spectra were manually phased and baseline-corrected in the Bruker TopSpin program (version 4.1.4) using a polynomial function. Chemical shifts were measured with the TMSP set to 0 ppm.

Afterwards, NMR spectra were normalized to the total intensity and binned in equidistant size buckets (0.01-0.04 ppm) using Amix software (4.0.2). The peak region corresponding to water suppression (4.75 ppm) was not taken under consideration. Bins of 0.02 ppm in size were finally chosen as the most appropriate buckets for the spectra as they maintained a satisfactory resolution. Bins corresponding to the region lower than 10.40 ppm, water residue was filtered out in order to enhance the statistical power of the data analysis. At the end a table was generated with 76 rows corresponding to the individual carob syrup samples and 504 columns corresponding to the different variables (spectral bins). The NMR data table was imported into SIMCA 14.1 (Umetrics, Umea, Sweden), and variables were centered and pareto scaled to make the features more comparable. Data were analyzed using PCA.

4. Chapter 4
Results and Discussion

4.1 UHPLC-QTOF-ESI-MS analysis

In total, 48 compounds were identified in all samples, using suspect screening methodology reaching to different identification confidence levels depending on the available information. The levels of identification were based on the criteria set on the work of Schymanski et al. [56]. Specifically, 21 compounds reached at the level of identification 1, 18 compounds were identified at level of identification 2a, and 9 compound identified at level 3.

The identified compounds belonged to several categories. The compounds, as well as the categories they belong in, are presented in Table 11. Specifically, 9 organic acids (gallic acid, p-coumaric acid, chlorogenic acid, benzoic acid, 2,5-dihydroxybenzoic acid, trans-cinnamic acid, 4-hydroxy benzoic acid, lactic acid, pyroglutamic acid), 6 fatty acyls/fatty acids and their conjugates (palmitic acid, oleic acid, linoleic acid, beta-hydroxyisovaleric acid, itaconic acid, acetyl-carnitine), 13 amino acids (valine, proline, L-tyrosine, L-threonine, phenylalanine, aspartic acid, alanine, leucine, L-serine, L-isoleucine, glutamic acid, γ-aminobutyric acid, L-tryptophan), 9 organo-oxygen compounds/carbohydrates and their carbohydrate conjugates (trehalose, mannose, D-ribose, D-glucose, D-fructose, xylose, sucrose, sorbitol, D-pinitol), 7 flavonols and flavonoids (kaempferol, quercetin, naringenin, apigenin, genistein, luteolin-7-O-glucoside, (+)-catechin), 3 nucleosides and nucleotides derivatives (adenosine, hypoxanthine, uridine) and 1 alkaloid (theobromine) were identified.

Table 11. Name and categorization of the identified metabolites.

Amino acids	Fatty acyls/Fatty acids and conjugates	Organo-oxygen compounds/Carbohydrates and Carbohydrate conjugates	Flavonols and Flavonoids	Organic acids	Nucleosides and Nucleotides derivatives	Alkaloids
Valine	Palmitic acid	Trehalose	Kaempferol	Gallic acid	Adenosine	Theobromine

Proline L-Tyrosine L-Threonine (Thr) Phenylalanine Aspartic Acid Alanine Leucine L-Serine (Ser) L-isoleucine (Ile) Glutamic acid (Glu) γ-aminobutyric acid L-Tryptophan (Trp)	Oleic acid Linoleic acid beta-Hydroxyisovaleric acid Itaconic acid Acetyl-carnitine	Mannose D-Ribose D-Glucose D-Fructose Xylose Sucrose Sorbitol D-pinitol	Quercetin Naringenin Apigenin Genistein Luteolin-7-O-glucoside (+)Catechin	p-Coumaric acid Chlorogenic acid Benzoic acid 2,5-Dihydroxybenzoic acid trans-Cinnamic acid 4-Hydroxy benzoic acid Lactic acid Pyroglutamic acid	Hypoxanthine Uridine	

Additionally, the % percentage of its chemical category to the overall content of carob, as well as the number of identified compounds alongside with their chemical class are depicted in Figures 25 and 26, respectively.

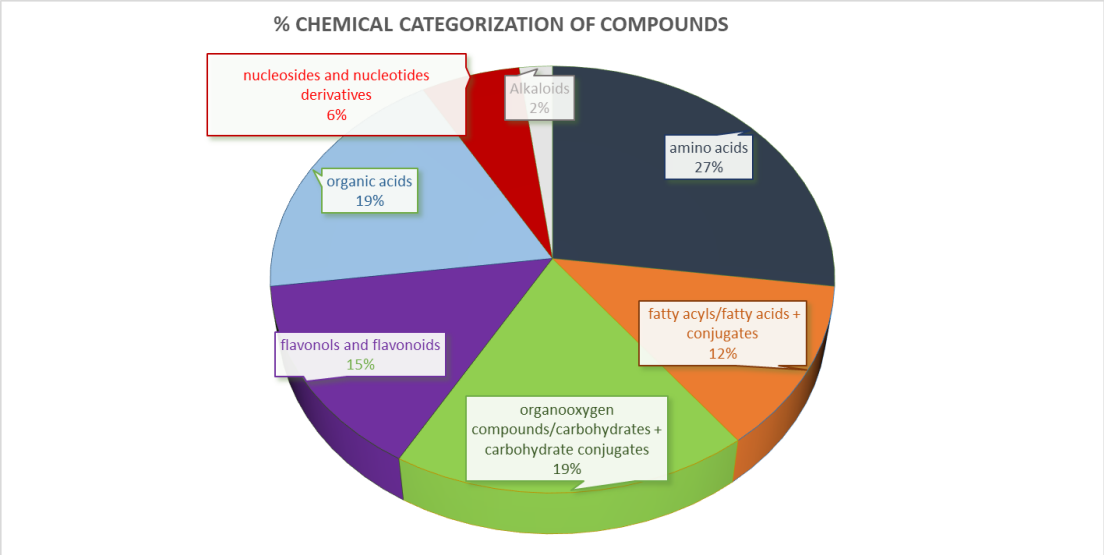


Figure 25. % Chemical categorization of compounds.

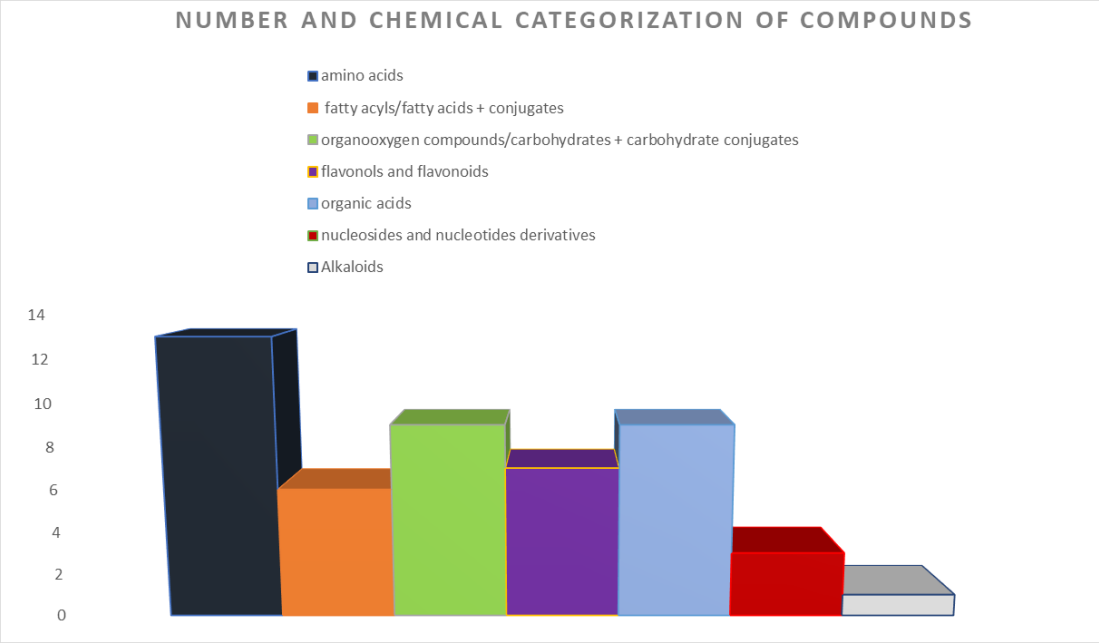


Figure 26. Number and chemical categorization of identified compounds.

All the compounds identified are presented in Table 12, where the compound name, the molecular formula, the experimental and the predicted retention time, the theoretical and experimental *m/z* value of the precursor ion, and the ionization mode are provided. The predicted retention time of the plausible candidates was estimated using the on-line available at http://rti.chem.uoa.gr. Additionally, the 5 most intense MS/MS fragments (if they existed) of the samples and their corresponding ones obtained either from the spectral library or the reference standards are presented. The acceptance identification criterion for the experimental *m/z* value was ± 0.005 Da error. The mass accuracy and the isotopic fit were evaluated for assigned molecular formula, using smart formula manually (Bruker Daltonics, Bremen, Germany). In Table 12, the cosine similarity scores of the investigated spectra for all samples are compared, as acquired from MS-DIAL, along with the corresponding level of identification.

Table 12. Compounds identified through suspect screening.

Compound Name	Chemical Formula	Exp. tR (min)	Pred. tR (min)	Application domain b	Exp. m/z values	Theo. m/z values	ESI mode	MS/MS fragments (most intense explained peaks) (sample)	Reference MS/MS Spectra (reference standard or mass spectral library)	DOT similarity score/ Total score	Level of Identification
Valine	$C_5H_{11}NO_2$	1.5	2.0	Box 1	116.0717	116.0717	-ESI	59.0140 116.0717 114.0560 101.0239	59.0145 116.0714 114.0557 101.0237	942.6152/ 795.8933	1
Trehalose	$C_{12}H_{22}O_{11}$	3.8	2.6	Box 2	341.1088	341.1089	-ESI	89.0239 71.0134	89.0240 71.0140	887.5275 / 906.2645	2a MassBank ID: PR311148
Pyroglutamic acid	$C_5H_7NO_3$	1.4	1.4	Box 1	128.0349	128.0353	-ESI	71.0137 88.0399 85.0290 104.0346	71.0133 88.0399 85.0295 104.0347	696.9209/ 710.5314	1
Proline	$C_5H_9NO_2$	1.5	1.7	Box 1	114.0559	114.0561	-ESI	114.0557 113.0241 71.0134	114.0557 113.0236 71.0137	662.4612 /759.1807	1

								59.0138	59.0142		
Palmitic acid	$C_{16}H_{32}O_2$	14.0	13.0	Box 1	255.2329	255.2325	-ESI	255.2332	255.2320	880 / 862.2629	2a MoNA ID: MetaboBASE1102
Oleic acid	$C_{18}H_{34}O_2$	14.1	13.1	Box 1	281.2489	281.2482	-ESI	281.2485 282.2471	281.2474 282.2509	668.9434/ 826.3914	2a MoNA ID: FiehnHILIC 002892
Naringenin	$C_{15}H_{12}O_5$	7.5	7.9	Box 1	271.0615	271.0612	-ESI	271.0609	271.0618	636.748 /814.0288	2a MassBank ID: BML00676
Mannose	$C_6H_{12}O_6$	1.5	1.7	Box 1	179.0559	179.0561	-ESI	179.0562 161.0452 59.0138 101.0239 71.0137	179.0550 161.0440 59.0120 101.0230 71.0120	557.8978 /760.7064	2a Metabolomics Core Facility spec89490
L-Tyrosine	$C_9H_{11}NO_3$	2.2	3.9	Box 2	180.0655	180.0656	-ESI	180.0663 137.0251	180.067 137.0235	576.3786 /648.6748	1

L-Threonine (Thr)	$C_4H_9NO_3$	1.4	1.3	Box 1	118.0508	118.0510	-ESI	89.0242 71.0137 101.0238 118.0505 113.0240	89.0238 71.0135 101.0239 118.0507 113.0242	627.8568 /711.0135	1
Linoleic acid	$C_{18}H_{32}O_2$	13.6	12.9	Box 1	279.2338	279.2330	-ESI	279.2325 241.2174 146.9661	279.2328 241.2173 146.9649	673.5611/ 854.1785	2a MassBank ID: EQ331654
Phenylalanine	$C_9H_{11}NO_2$	3.1	4.2	Box 1	164.0719	164.0712	-ESI	164.0713 147.0464 72.0107 148.0485 150.0552	164.0718 147.0450 72.0087 148.0469 150.0567	726.6403 /835.8843	1
Kaempferol	$C_{15}H_{10}O_6$	7.4	7.2	Box 1	285.0396	285.0405	-ESI	59.0139 285.0408	59.0139 285.0402	893.6086 /722.4335	2a 2a GNPS ID: VF-NPL-QEHF014174
Gallic acid	$C_7H_6O_5$	1.3	2.4	Box 2	169.0140	169.0143	-ESI	169.0145 125.0243	169.0130 125.0240	615.1254 /759.1628	2a RIKEN PLaSMA ID: RIKENPlaSMA 008154

D-Ribose	$C_5H_{10}O_5$	2.5	1.7	Box 1	131.0347	131.0337	-ESI	71.0137 131.0343 113.0242 75.0084	71.0120 131.0350 113.0230 75.0080	562.581/ 627.903	2a Metabolomics Core Facility spec311291
beta-Hydroxyisovaleric acid	$C_5H_{10}O_3$	1.9	2.6	Box 1	117.0554	117.0543	-ESI	59.0141	59.0120	700.2011 / 706.8716	2a MoNA ID: FiehnHILIC 001094
Aspartic Acid	$C_4H_7NO_4$	1.3	0.8	Box 1	132.0301	132.0302	-ESI	132.0299 71.0137 85.0290 129.0377	132.0302 71.0133 85.0294 129.0373	615.3026/ 631.8856	1
Apigenin	$C_{15}H_{10}O_5$	8.6	7.6	Box 1	269.0449	269.0450	-ESI	269.0459 63.02347 270.0485	269.045 63.0242 270.0484	640.8653/ 802.2033	1
D-Glucose	$C_6H_{12}O_6$	1.5	1.7	Box 1	161.0455	161.0443	-ESI	59.0140 89.0242 71.0125	59.0130 89.0230 71.0120	717.3643 /827.7443	2a Metabolomics Core Facility spec309857
Alanine	$C_3H_7NO_2$	1.4	1.3	Box 1	88.0400	88.0404	-ESI	71.0137 88.0399	71.0133 88.0399		1

								85.0300 74.0248	85.0295 74.0247	396.2613/ 711.1402	
Adenosine	$C_{10}H_{13}N_5O_4$	2.1	4.1	Box 2	266.0888	266.0900	-ESI	59.0139	59.0140	612.3462/ 628.1694	2a Global Natural Product Social Molecular CCMSLIB 00005720 762
Leucine	$C_6H_{13}NO_2$	2.6	2.5	Box 1	130.0873	130.0874	-ESI	130.0865 119.0496 120.0549	130.0875 119.0502 120.0543	635.3114/ 845.7513	1
D-Fructose	$C_6H_{12}O_6$	1.5	1.7	Box 1	179.0561	179.0560	-ESI	179.0562 161.0453 119.0349 59.0139 114.0557	179.0564 161.0457 119.0339 59.0141 114.0557	559.5324/ 640.7661	1
Xylose	$C_5H_{10}O_5$	3.6	1.7	Box 2	149.0448	149.0456	-ESI	59.0141	59.0120	926.0304/ 835.9714	2a Vaniya/Fiehn Natural Products QEHF011 799
Theobromine	$C_7H_8N_4O_2$	1.5	3.7	Box 2	179.0563	179.0574	-ESI				1

								179.0563 114.0556	179.0564 114.0557	612.4746/ 616.1194	
L-Serine (Ser)	$C_3H_7NO_3$	1.4	1.2	Box 1	104.0348	104.0353	-ESI	89.0242 71.0137 73.0291 104.0347	89.0238 71.0135 73.0289 104.0347	893.7769 /772.9861	1
Genistein	$C_{15}H_{10}O_5$	8.5	7.8	Box 1	269.0452	269.0455	-ESI	269.0467 132.0248	269.0450 132.0210	569.9865/ 661.4868	2a MassBank ID: AU241959
4-Hydroxy benzoic acid	$C_7H_6O_3$	3.7	2.8	Box 1	137.0244	137.0244	-ESI	137.0250 93.0341	137.0239 93.0340	615.3286/ 643.4152	1
L-isoleucine (Ile)	$C_6H_{13}NO_2$	2.3	2.5	Box 1	130.0875	130.0874	-ESI	130.0873 93.0299	130.0874 93.0337	893.569 /887.7137	1
Hypoxanthine	$C_5H_4N_4O$	1.4	3.3	Box 2	135.0291	135.0311	-ESI	133.0140 135.0296	133.0190 135.0320	596.9228/ 648.7356	2a Vaniya/Fiehn Natural Products QTOF008024

Sucrose	$C_{12}H_{22}O_{11}$|O	1.3	3.5	Box 2	341.1094	341.1084	-ESI	179.0564 341.1091 89.02419 101.024 71.01379	179.056 341.108 89.025 101.024 71.015	668.4019/ 854.2713	2a MassBank ID: PR100500
Glutamic acid (Glu)	$C_5H_9NO_4$	1.3	0.9	Box 1	146.0466	146.0459	-ESI	132.0297 146.0459 88.0400 129.0355 128.0349 118.0503	132.0302 146.0460 88.0400 129.0373 128.0352 118.0501	580.2543/ 714.5698	1
Luteolin-7-O-glucoside	$C_{21}H_{20}O_{11}$	6.4	6.5	Box 1	447.0931	447.0933	-ESI	447.0947	447.0940	595.8025 / 729.7804	2a RIKEN PLaSMA ID: RIKENPlaSMA 005617
Quercetin	$C_{15}H_{10}O_7$	6.5	7.0	Box 1	301.0349	301.0354	-ESI	271.0583 178.9941 177.0220	271.0617 178.9986 177.0174	541.3799/ 628.6583	1
p-Coumaric acid	$C_9H_8O_3$	2.8	2.6	Box 1	163.0398	163.0401	-ESI	61.9884 163.0390	61.9891 163.0404	596.31818/ 650.3761	1
Chlorogenic acid	$C_{16}H_{18}O_9$	2.9	3.5	Box 1	353.0877	353.0878	-ESI	353.0876 179.0347	353.0877 179.0337		1

								176.0404 191.0549	176.0400 191.0562	831.4444/ 787.3236	
(+)-Catechin	$C_{15}H_{14}O_6$	4.0	7.0	Box 3	289.0717	289.0718	-ESI	289.0714 61.9885	289.0718 61.9881	612.6068/ 703.1028	1
Benzoic acid	$C_7H_6O_2$	3	2.8	Box 1	121.0286	121.0295	-ESI	121.0293	121.0295	624.2576/ 742.3539	1
Sorbitol	$C_6H_{14}O_6$	1.4	2.3	Box 1	181.0714	181.0711	-ESI	181.0716 113.0241 119.0347 101.0239 71.01367	181.0710 113.0230 119.0340 101.0230 71.0120	607.8314/ 623.7103	2a Vaniya/Fiehn Natural Products QEHF011047
Lactic acid	$C_3H_6O_3$	1.4	1.4	Box 1	89.0242	89.0244	-ESI	59.0140 71.0137 73.0288 75.0081 72.9923	59.01386 71.01386 73.0295 75.0088 72.9931	In silico fragmentation	3
D-pinitol	$C_7H_{14}O_6$	1.4	1.7	Box 1	193.0722	193.0707	-ESI	89.0240 59.0139 71.0137 119.0351 179.0563	89.0244 59.0139 71.0139 119.0350 179.0561	In silico fragmentation	3

Acetyl-carnitine	$C_9H_{17}NO_4$	2.6	2.7	Box 1	202.1075	202.1085	-ESI	130.0873 158.0817 125.0248 172.0980	130.0874 158.0823 125.0244 172.0980	In silico fragmentation	3
Itaconic acid	$C_5H_6O_4$	1.4	1.1	Box 1	129.0189	129.0193	-ESI	59.0139 71.0137 101.0238 113.0244 85.0287	59.0139 71.0139 101.0244 113.0244 85.0295	In silico fragmentation	3
2,5-Dihydroxybenzoic acid	$C_7H_6O_4$	2.4	3.0	Box 1	153.0208	153.0193	-ESI	71.0130 87.0078 139.0393 108.0207 113.0235	71.01386 87.0088 139.0401 108.0217 113.0244	In silico fragmentation	3
trans-Cinnamic acid	$C_9H_8O_2$	4.7	2.8	Box 2	147.0453	147.0452	-ESI	132.0565	132.0581	In silico fragmentation	3
γ-aminobutyric acid	$C_4H_9NO_2$	1.4	1.9	Box 1	102.0550	102.0561	-ESI	59.0141 71.0137 85.0290 73.0291	59.0139 71.0139 85.0295 73.0295	In silico fragmentation	3
Uridine	C9H12N2O6	2.2	3.8	Box 2	243.0637	243.0623	-ESI	71.0137 87.0085 133.0135 89.0234 101.0236	71.0139 87.0088 133.0507 89.0244 101.0244	In silico fragmentation	3
L-Tryptophan (Trp)	C11H12N2O2	3.7	4.5	Box 1	203.0844	203.0826	-ESI	71.0136	71.01386	In silico fragmentation	3

Quantitative analysis

The compounds, with analytical standards being available, were quantified using the calibration curves of reference standards. Specifically, the concentrations of 6 organic acids (gallic acid, p-coumaric acid, 4-hydroxybenzoic acid, 2,5-dihydroxybenzoic acid, pyroglutamic acid, trans- cinnamic acid), 2 flavonoids (naringenin, apigenin), one alkaloid (theobromine), 11 amino acids (γ-aminobutyric acid, alanine, phenylalanine, valine, aspartic acid, glutamic acid, L-isoleucine, L-serine, L-threonine, L-tyrosine, L-tryptophan) and one sugar (D-glucose) were determined.

The compounds 4-hydroxybenzoic acid, γ-aminobutyric acid, phenylalanine, L-tryptophan (commercial samples), L-tyrosine (traditional samples) and L-isoleucine were identified through suspect screening, but their concentrations were below the method's limit of detection (LOD).

The mean concentrations and the maximum and minimum range values for the investigated samples (traditional and commercial) are presented in Table 13. Additionally, the corresponding calibration equations in the form of $y= (a\pm S_a)\ x+(b\pm S_b)$, as well as their correlation coefficients are also presented in the Table below.

Table 13. Quantitative analysis of compounds.

Analyte	Category	LOD (mg/kg)	LOQ (mg/kg)	Equation of the external matrix matched curve $y= (a\pm S_a)x+(b\pm S_b)$	Correlation coefficient (R^2)	Concentration Range (mean) (Commercial samples) (mg/kg)	Concentration Range (mean) (Traditional samples) (mg/kg)
Gallic acid	organic acids	0.11	0.30	$y=(33.89\pm2.7)10^2x+(88.67.\pm2.3)10^2$	0.9990	3.23-10.75 (6.46)	4.13-16.04 (8.53)
p-Coumaric acid	organic acids	$3.3*10^{-2}$	$9.9*10^{-2}$	$y=(261.09\pm2.3)10x+(233.62\pm4.3)10^2$	0.998	0.26-0.44 (0.33)	0.19-0.48 (0.27)
4-Hydroxybenzoic acid	organic acids	$2.3*10^{-2}$	$7.1*10^{-2}$	$y=(335.44\pm2.1)10x+(299.23\pm1.7)10^2$	0.991	<LOD	<LOD
2,5-Dihydroxybenzoic acid	organic acids	$0.19*10^{-2}$	$0.57*10^{-2}$	$y=(372.23\pm2.0)10x+(3491.76\pm1.5)10^2$	0.9990	0.61-0.64 (0.62)	0.61-0.62 (0.62)
trans-Cinnamic acid	organic acids	$6.2*10^{-2}$	0.19	$y=(16.20\pm1.6)10^2x+(85.37\pm1.5)10^2$	0.9990	0.73-0.98 (0.86)	0.62-0.86 (0.70)
Pyroglutamic acid (PCA)	organic acids	$7.7*10^{-2}$	0.23	$y=(7.09\pm1.0)10^2x+(42.79\pm9.5)10^2$	0.998	5.77-15.24 (11.40)	0.44-2.80 (1.51)
Naringenin	flavonols and flavonoids	1.1	3.3	$y=(2.78\pm4.4)10^2x+(13.24\pm3.7)10^2$	0.998	7.24-10.91 (8.50)	5.30-11.76 (9.27)
Apigenin	flavonols and flavonoids	$6.1*10^{-2}$	0.18	$y=(799.14\pm4.3)10x+(233.62\pm4.3)10^2$	0.9990	0.17-0.18 (0.17)	0.16-0.18 (0.17)
Theobromine	alkaloids	$1.1*10^{-2}$	$3.3*10^{-2}$	$y=(158.21\pm9.3)10x+(2813.99\pm8.7)10^2$	0.995	5.96-11.32 (8.75)	0.47-2.79 (1.23)
γ-Aminobutyric acid	amino acids	$0.98*10^{-2}$	$2.9*10^{-2}$	$y=(261.09\pm2.0)10x+(671.22\pm1.9)10^2$	0.997	<LOD	<LOD

(GABA)							
Alanine (Ala)	amino acids	$1.3*10^{-2}$	0.35	$y=(23.88\pm1.3)10^{2}x+(36.90\pm1.2)10^{2}$	0.997	13.45-19.58 (16.23)	4.81-24.40 (14.67)
Phenylalanine (Phe)	amino acids	$6.8*10^{-2}$	$2.1*10^{-2}$	$y=(725.48\pm5.1)10x+(2472.59\pm4.8)10^{2}$	0.997	<LOD	<LOD
Valine (Val)	amino acids	$8.2*10^{-2}$	0.26	$y=(348.99\pm6.7)10x+(271.33\pm6.3)10^{2}$	0.997	1.13-3.19 (2.40)	1.15-4.07 (2.03)
Aspartic acid (Asp)	amino acids	$9.3*10^{-2}$	0.28	$y=(122.41\pm3.5)10x+(124.46\pm3.1)10^{2}$	0.998	3.86-12.04 (6.96)	1.15-13.75 (5.31)
Glutamic acid (Glu)	amino acids	$6.4*10^{-2}$	0.19	$y=(111.81\pm5.6)10x+(289.73\pm5.3)10^{2}$	0.998	0.05-0.14 (0.10)	0.12-1.59 (0.42)
L-isoleucine (Ile)	amino acids	$0.38*10^{-2}$	0.011	$y=(317.43\pm4.6)10x+(4047.52\pm4.4)10^{2}$	0.9990	<LOD	<LOD
L-Serine (Ser)	amino acids	1.3	3.9	$y=(24.63\pm9.0)10^{2}x+(22.87\pm7.8)10^{2}$	0.995	2.01-12.35 (6.99)	4.70-20.76 (11.06)
L-Threonine (Thr)	amino acids	0.15	0.48	$y=(153.10\pm3.9)10x+(83.14\pm3.7)10^{2}$	0.998	0.47-2.67 (1.50)	0.78-7.13 (3.92)
L-Tyrosine (Tyr)	amino acids	$0.56*10^{-2}$	$1.7*10^{-2}$	$y=(275.09\pm1.7)10x+(1002.34\pm1.8)10^{2}$	0.9990	0.18-0.73 (0.46)	<LOD
L-Tryptophan (Trp)	amino acids	$0.87*10^{-2}$	$2.7*10^{-2}$	$y=(289.94\pm3.9)10x+(1472.46\pm3.4)10^{2}$	0.998	<LOD	0.30-0.34 (0.31)
Glucose	organo-oxygen compounds/carbohydrates + carbohydrate conjugates	$1.2*10^{-2}$	0.035	$y=(1731.83\pm8.3)1x+(2378.78\pm7.8)10^{2}$	0.995	8.47-16.62 (12.35)	0.29-3.87 (1.42)

Significant fluctuations were observed in the concentrations of specific metabolites in commercial carob syrups (an example is presented in Figure 27). More specifically, the concentrations of several metabolites presented elevated relative standard deviations (RSD), such as gallic acid (RSD = 29%), glutamic acid (RSD = 35%), L-serine (RSD = 45%), and aspartic acid (RSD = 33%). This may be attributed to the different preparation methods of carob syrups, emphasizing the importance of standardizing the preparation process.

Additionally, when comparing commercial and traditional carob syrups, it was found that the average concentration values of the chemical categories of amino acids (alanine, valine, aspartic acid, glutamic acid, L-serine, L-threonine, L-tyrosine, and L-tryptophan) and flavonoids (naringenin, apigenin) were 10% higher in traditional carob syrups than in the corresponding commercial ones. While commercial carob syrups contain 41% more total organic acids than traditional ones, there was a significant difference in the average concentration values of organic acids (gallic acid, p-coumaric acid, 2,5-dihydroxybenzoic acid, pyroglutamic acid, and trans-cinnamic acid). More specifically, the largest concentration difference was observed in pyroglutamic acid, with an average concentration of 11.40 mg kg^{-1} in commercial carob syrup, while in the corresponding traditional syrup, it was 1.51 mg kg^{-1}. Finally, significant differences were also found in the case of glucose and theobromine, as their average concentrations were higher by 89% and 86%, respectively, in commercial carob syrups.

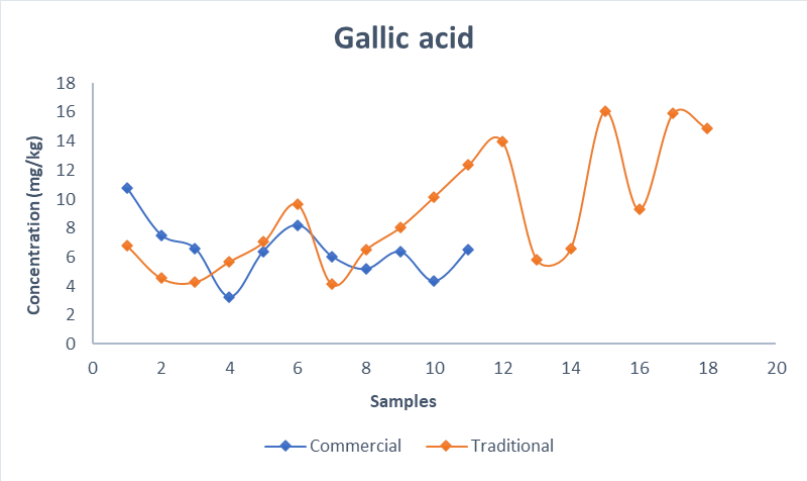


Figure 27. The fluctuation of gallic acid's concentration in commercial and traditional carob samples.

The extraction procedure, in particular the temperature (20-100°C) and extraction time (5'-35'), appear to be important factors in the extraction of several metabolites. On the other hand, some metabolites appear to be negatively affected by these variables, such as high temperature (>80°C). Particularly, the concentrations of theobromine, amino acids (alanine, valine, aspartic acid, glutamic acid, L-serine, L-threonine, L-tryptophan), and organic acids appear to rise with rising extraction time and temperature. These organic acids include gallic acid, p-coumaric acid, 2,5-dihydroxybenzoic acid, pyroglutamic acid, trans- cinnamic acid. The biggest shift happens after 25' of extraction time, before which the growth is steady but happens more slowly, as shown in Figure 28 for the instance of alanine at 82°C. Finally, the concentration of D-glucose drops as the temperature rises over 80°C.

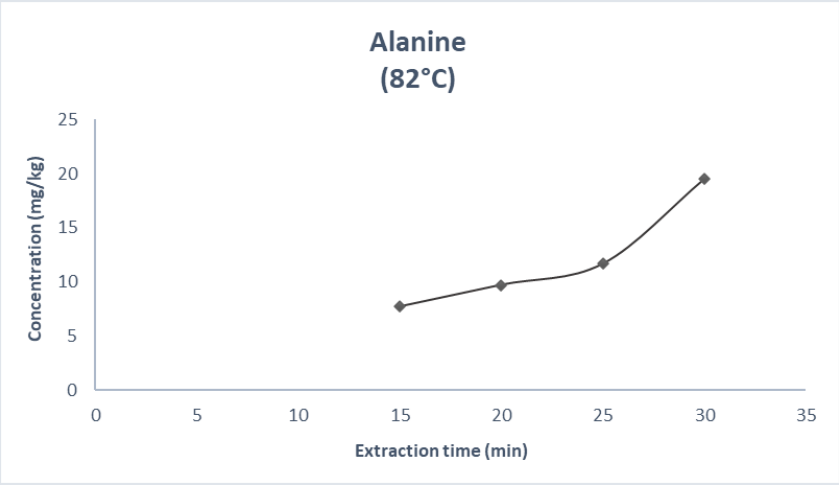


Figure 28. Increase in the concentration of alanine (especially after 25 minutes).

The concentrations of analytes for commercial (Cretan and Australian) and traditional (Slovenian and Croatian) samples are presented in Table 14, next to the Cypriot samples, respectively.

Table 14. The concentrations of detected metabolites for commercial (Cretan and Australian) and traditional (Slovenian and Croatian) samples compared to the Cypriot samples, respectively. Commercial samples colored grey.

Analyte	Category	Concentration Range (mean) (Commercial samples) (mg/kg)	Concentration (Commercial Cretan sample) (mg/kg)	Concentration (Commercial Australian sample) (mg/kg)	Concentration Range (mean) (Traditional samples) (mg/kg)	Concentration Traditional Slovenian sample) (mg/kg)	Concentration Traditional Croatian sample) (mg/kg)
Gallic acid	organic acids	3.23-10.75 (6.46)	3.78	13.57	4.13-16.04 (8.53)	-	14.86
p-Coumaric acid	organic acids	0.26-0.44 (0.33)	0.30	0.64	0.19-0.48 (0.27)	2.24	0.36
4-Hydroxybenzoic acid	organic acids	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD
2,5-Dihydroxybenzoic acid	organic acids	0.61-0.64 (0.62)	0.62	0.62	0.61-0.62 (0.62)	0.61	0.61
trans-Cinnamic acid	organic acids	0.73-0.98 (0.86)	0.51	2.15	0.62-0.86 (0.70)	1.15	0.75
Pyroglutamic acid (PCA)	organic acids	5.77-15.24 (11.40)	13.77	2.40	0.44-2.80 (1.51)	3.21	2.80
Naringenin	flavonols and flavonoids	7.24-10.91 (8.50)	9.05	10.40	5.30-11.76 (9.27)	3.63	11.06
Apigenin	flavonols and flavonoids	0.17-0.18 (0.17)	0.12	2.23	0.16-0.18 (0.17)	0.12	0.17
Theobromine	alkaloids	5.96-11.32 (8.75)	8.71	4.53	0.47-2.79 (1.23)	4.20	2.34

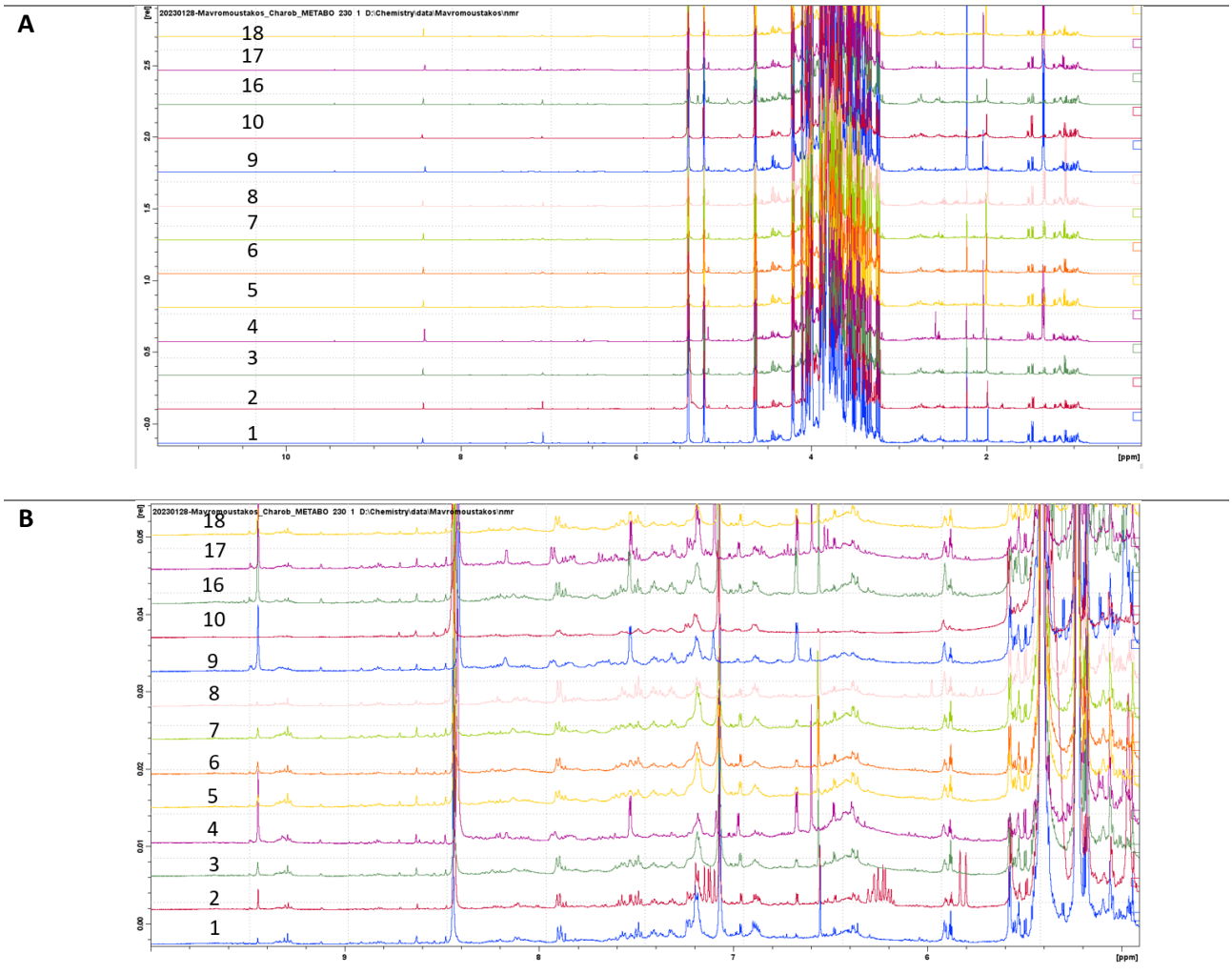
γ-Aminobutyric acid (GABA)	amino acids	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD
Alanine (Ala)	amino acids	13.45-19.58 (16.23)	17.56	19.16	4.81-24.40 (14.67)	17.91	18.40
Phenylalanine (Phe)	amino acids	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD
Valine (Val)	amino acids	1.13-3.19 (2.40)	5.03	0.89	1.15-4.07 (2.03)	0.27	3.56
Aspartic acid (Asp)	amino acids	3.86-12.04 (6.96)	11.85	2.00	1.15-13.75 (5.31)	10.23	13.75
Glutamic acid (Glu)	amino acids	0.05-0.14 (0.10)	0.06	0.08	0.12-1.59 (0.42)	0.23	0.44
L-isoleucine (Ile)	amino acids	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD
L-Serine (Ser)	amino acids	2.01-12.35 (6.99)	7.05	16.66	4.70-20.76 (11.06)	-	17.72
L-Threonine (Thr)	amino acids	0.47-2.67 (1.50)	1.08	3.56	0.78-7.13 (3.92)	0.32	6.64
L-Tyrosine (Tyr)	amino acids	0.18-0.73 (0.46)	0.27	0.12	<LOD	-	<LOD
L-Tryptophan (Trp)	amino acids	<LOD	<LOD	<LOD	0.30-0.34 (0.31)	<LOD	0.32
Glucose	organo-oxygen compounds/carbohydrates + carbohydrate conjugates	8.47-16.62 (12.35)	12.38	6.57	0.29-3.87 (1.42)	0.60	3.29

Although some differences are observed in the concentration of specific analytes, however the limited number of commercial (n=1 Cretan, n=1 Australian) and traditional carob syrups (n=1 Croatian, n=1 Slovenian) from other countries, do not allow to draw reliable conclusions.

4.2 NMR analysis

NMR, in combination with chemometric techniques (PCA), complemented the results of MS. Metabolic profiling 1D spectra were recorded to facilitate samples clustering and to provide useful information about contained metabolites and instructors. NMR spectra were acquired according to Bruker protocol metabolomic. The study of the NMR spectra revealed the presence of sugars, as it was expected, hence the focus of this study was set to explore other secondary metabolites. The oxygenated region of the spectra (5-3 ppm) was populated with broad overlapping peaks belonging to sugars, therefore in the following Figures are mainly highlighted the aromatic (8-6 ppm) and aliphatic region (2.5-0.5 ppm), that differentiate the samples. The NMR spectra of all samples (commercial, traditional) are shown in Figures 29, 30, 31, 32 and the specific focus is on these two spectral regions mentioned above. For the traditional samples, each series (B, G, D) corresponds to the same region.

Commercial



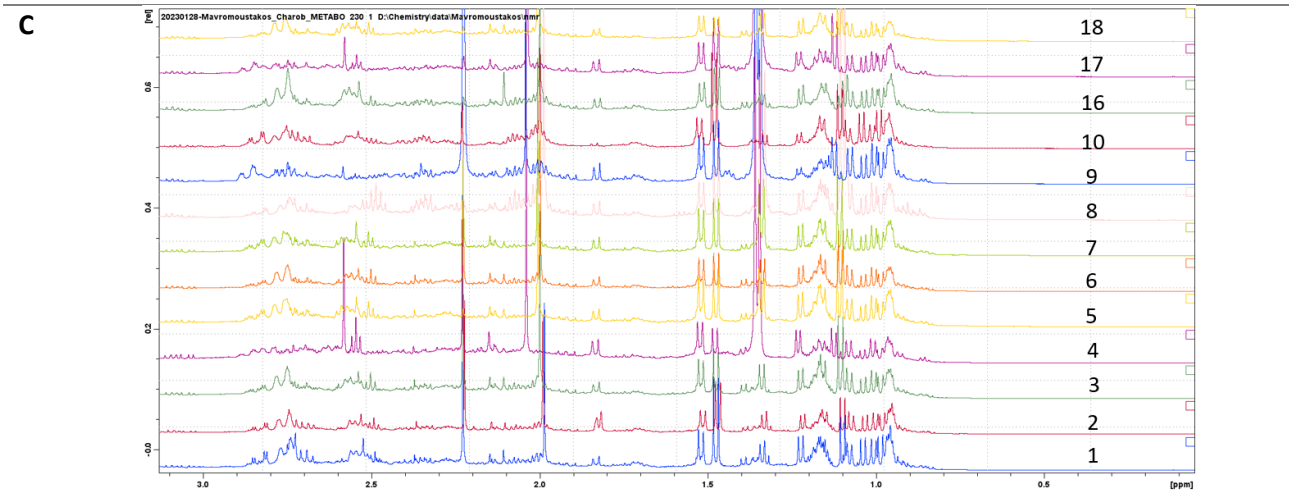
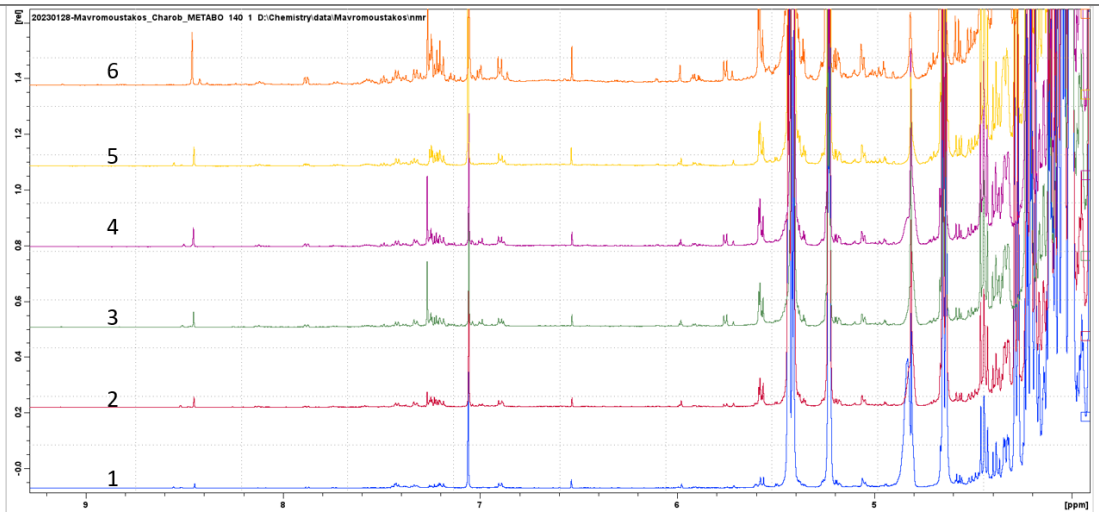


Figure 29. NMR spectra of commercial samples. A: The whole NMR spectra (10-0 ppm), B: Aromatic region (8-6 ppm), C: Aliphatic region (2.5-0.5 ppm).

In the commercial samples (Figure 29) the presence of sugars is dominated, and no distinctive differences can be found in the spectra in the other regions, even on the cases of Cretan and Australian carob syrup.

Samples from Saint Fyla (Series B) are shown in Figure 30. Carob syrup was prepared in different temperatures to monitor metabolites extraction. Influence of temperature is depicted in the spectra as sample 6 (100 °C) contain higher levels of secondary metabolites in comparison to the other samples, prepared in lower temperatures, while the spectra of all samples did not differ qualitatively, indicating their significance of the specific preparation conditions.

B Series (Traditional)



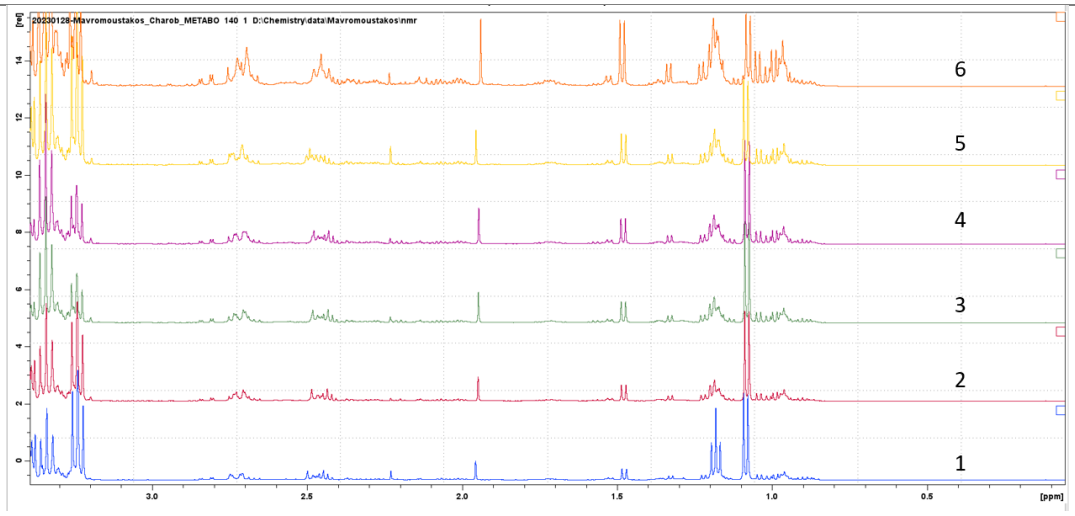
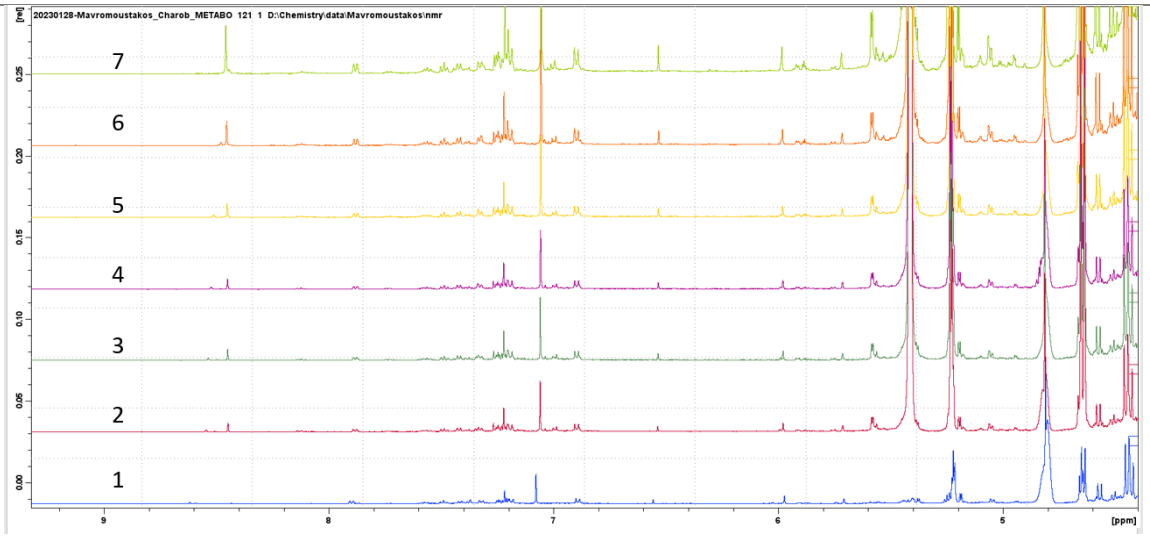


Figure 30. NMR spectra of traditional samples (B series). Upper image: Aromatic region (8-6 ppm), Lower image: Aliphatic region (2.5-0.5 ppm).

Likewise, NMR spectra of samples from Klonari (Series G) and Dieronas (Series D) show better capacity of metabolites extraction in higher temperatures, as presented in Figure 31 and 32, respectively.

G Series (Traditional)



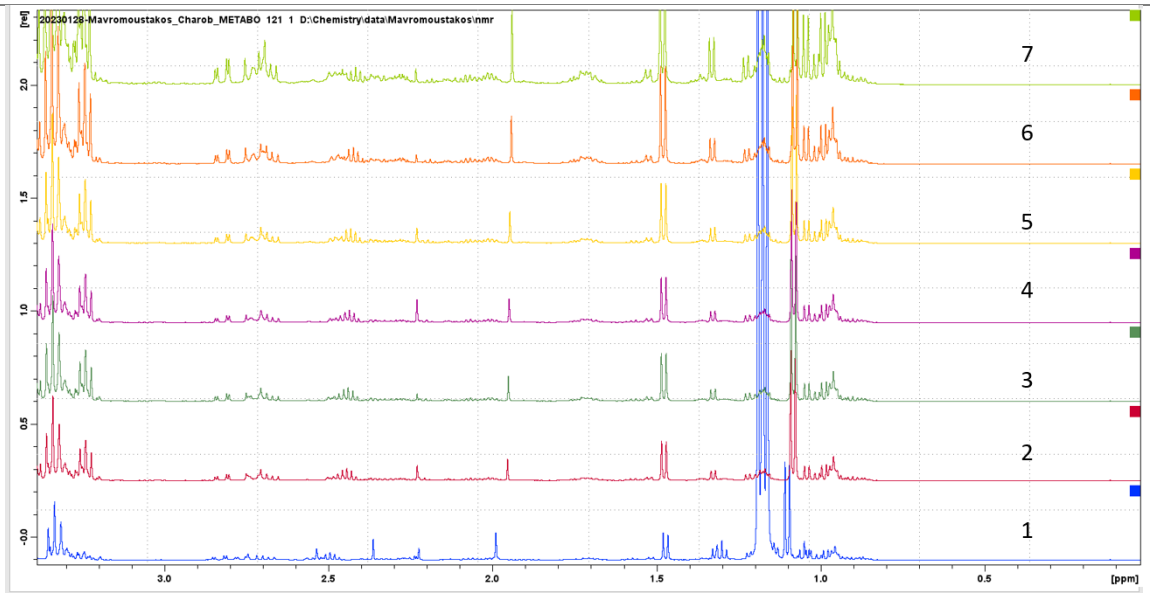
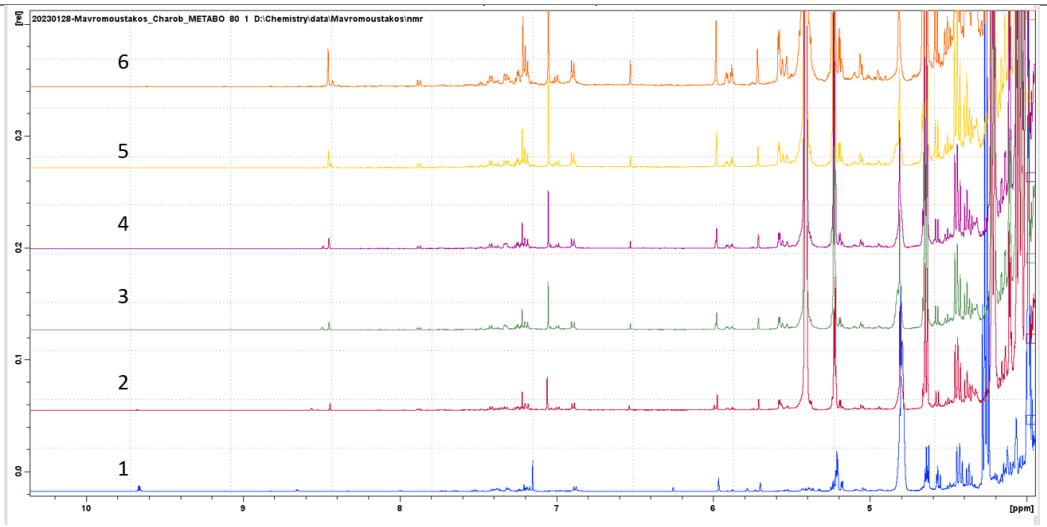


Figure 31. NMR spectra of traditional samples (G series). Upper image: Aromatic region (8-6 ppm), Lower image: Aliphatic region (2.5-0.5 ppm).

D Series (Traditional)



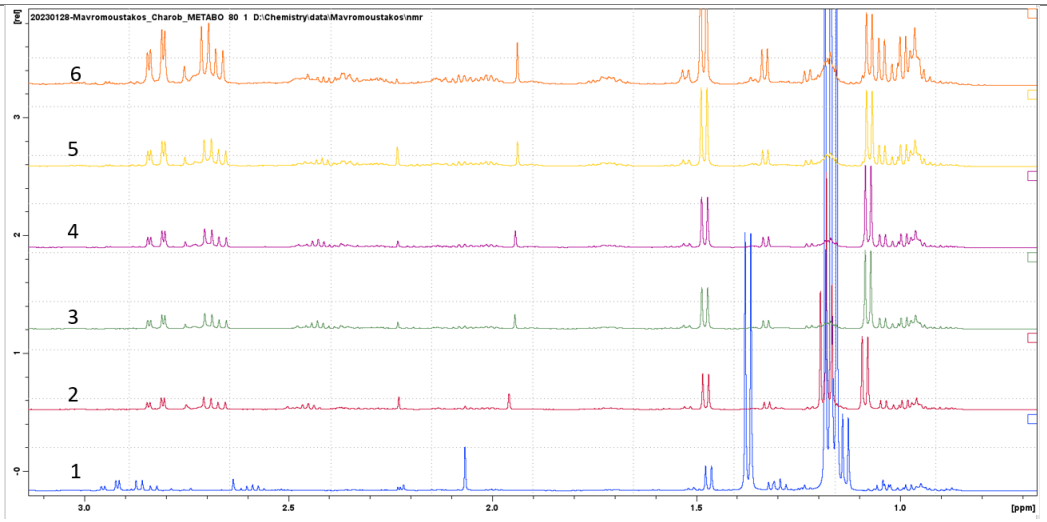
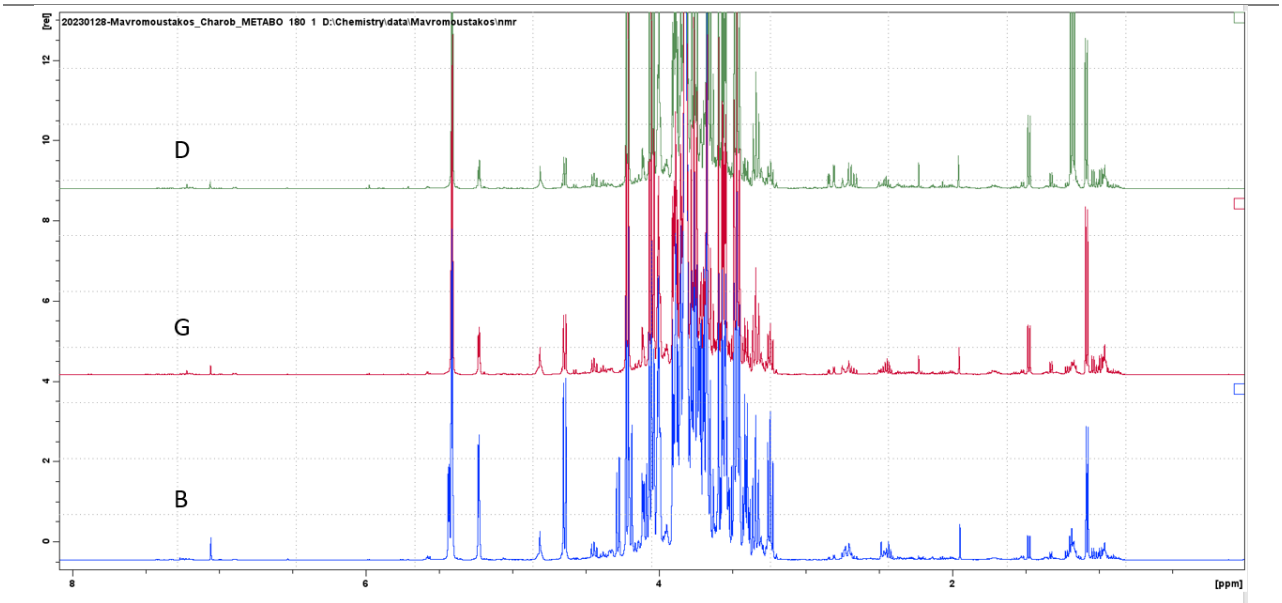


Figure 32. NMR spectra of traditional samples (D series). Upper image: Aromatic region (8-6 ppm), Lower image: Aliphatic region (2.5-0.5 ppm).

Figure 33 presents a comparison of NMR spectra for traditional samples in series 2 (B, D, G). These 3 samples have been prepared in the same way (extraction time=10', temperature=80°C). The observed differences are not significant, and the peaks in the aromatic and aliphatic regions are nearly identical, with some exceptions in the intensities of certain peaks highlighted in red boxes. From the above, it is evident that the geographical region does not significantly affect the presence or absence of a metabolite but influences the concentration variation of certain metabolites. This could be attributed to the environmental conditions, microenvironment, and soil agrochemical properties of the examined area.

Series 2



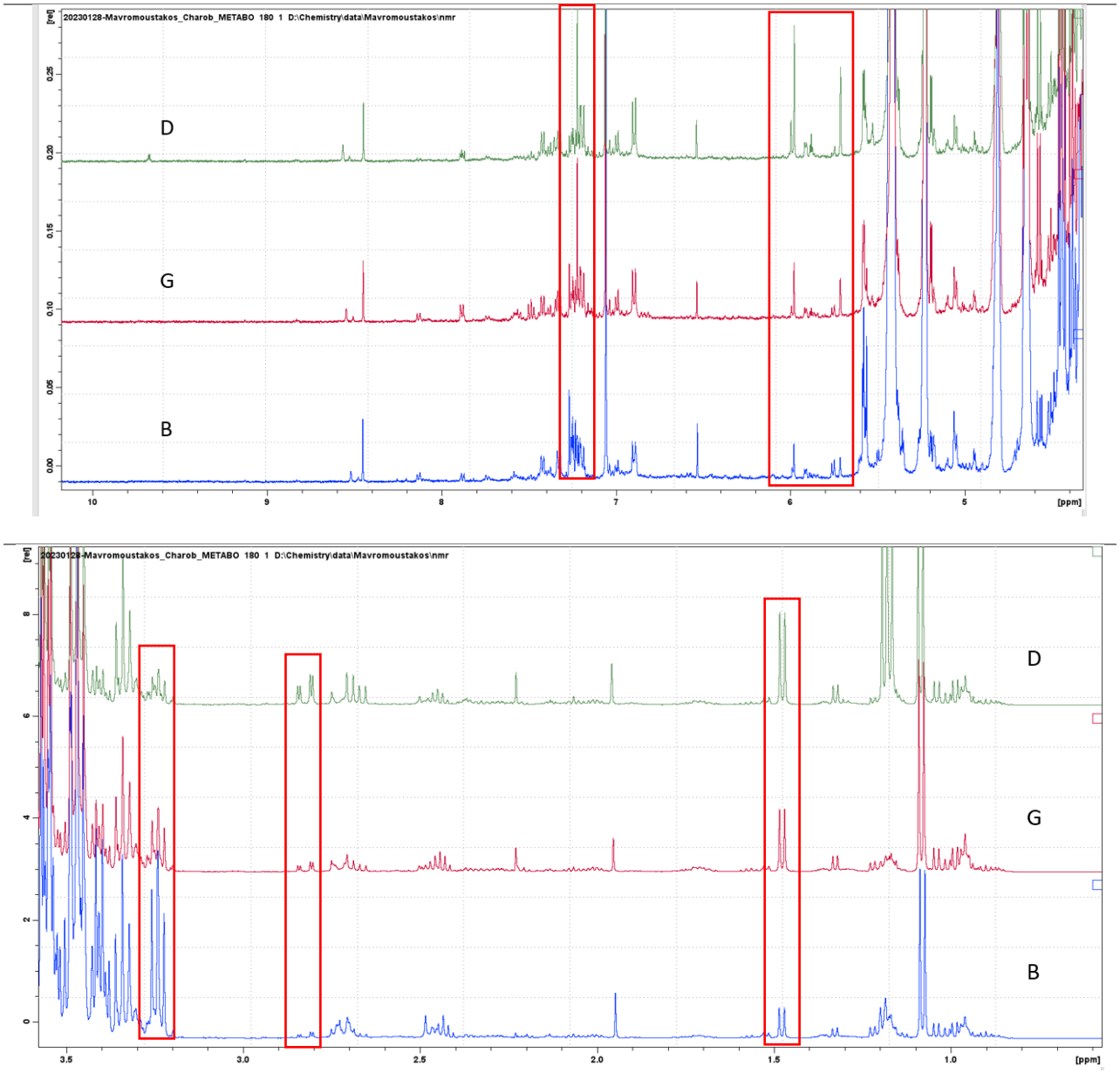
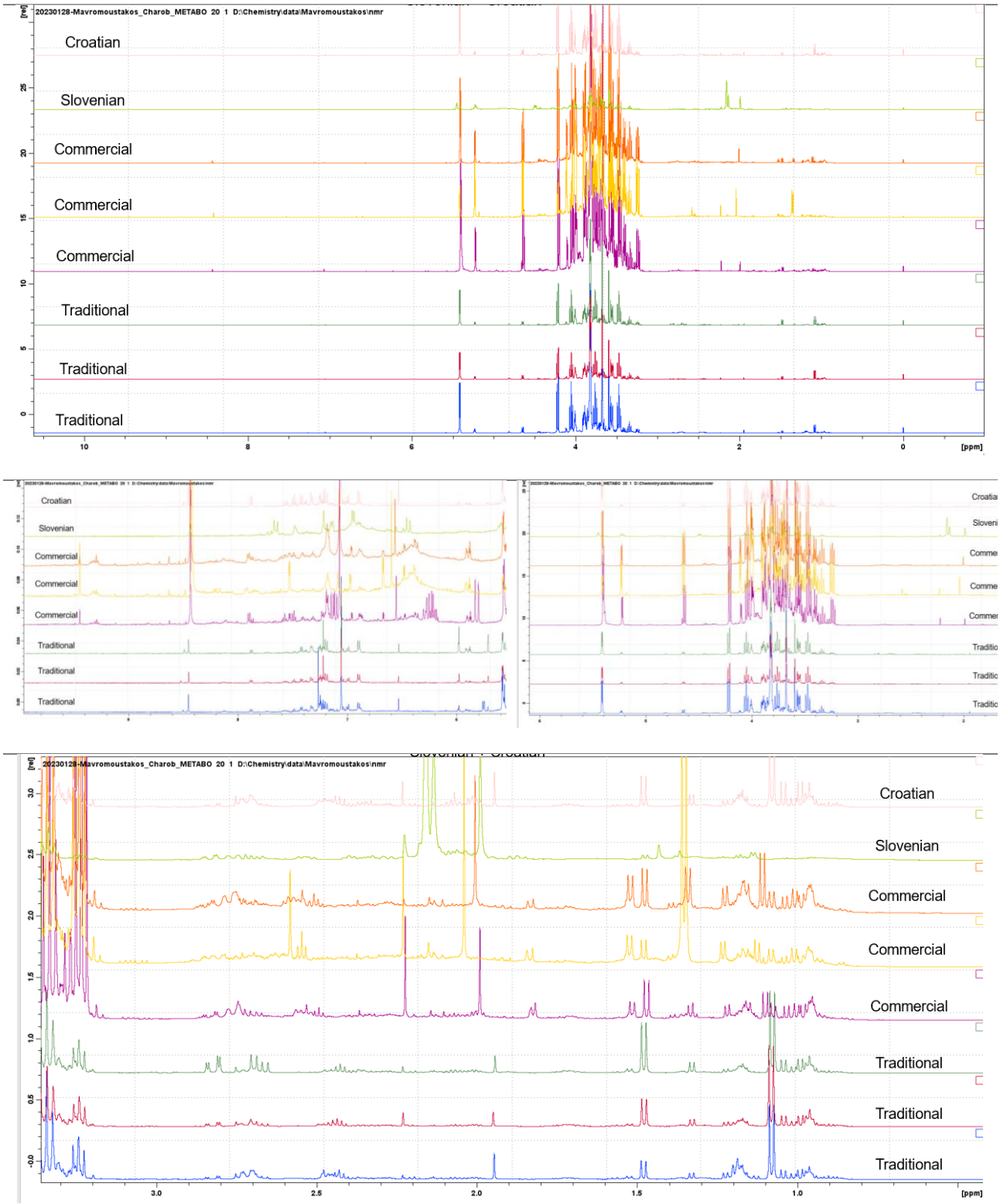


Figure 33. A comparison of NMR spectra of traditional samples for series 2 (B, D, G). Each series corresponds to the same region. 1st image: The whole NMR spectra, 2nd image: Aromatic region, 3rd image: Aliphatic region.

Next, a qualitative comparison was performed among 3 traditional and 3 commercial carob syrups, along with one sample from Slovenia and Croatia (Figure 34). Through the spectra of 3 reference standards (D-xylose, D-fructose, D-glucose), the corresponding compounds were identified, as well as some difference in their peak intensities. This confirmed the presence of the 3 isomeric sugars, which were previously identified using the MS, but also revealed the possible presence of additional (non-natural) sugars, especially in the case of commercial syrups, but further investigated is needed. In the NMR spectra we observed the characteristic peaks at 5.23 (*d, J*=3.8 Hz) and 4.64 (*d, J*=8 Hz) for D-glucose and D-lactose,

5.41 (*d, J*=3.8 Hz) and 4.22 (*d, J*=8.8 Hz) for sucrose, 5.75 (*d, J*=7.8 Hz) for sorbitol and 5.19 (*d, J*=4.2 Hz) for D-xylose, respectively.

3 Traditional + 3 Commercial Samples + Slovenian + Croatian



Traditional + Commercial + Slovenian + Croatian + Standards (D-xylose, D-Fructose, D- Glucose)

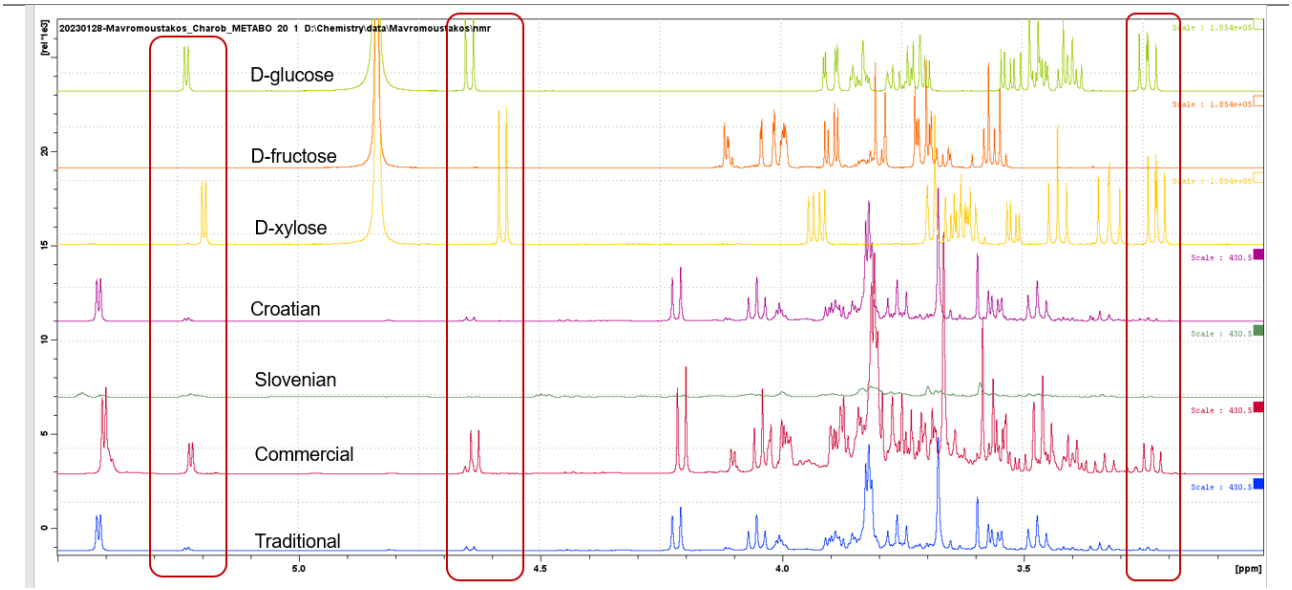


Figure 34. A comparison of NMR spectra of 3 traditional, 3 commercial samples, one Slovenian and one Croatian carob syrups. The last image also shows the spectra of the 3 standards.

Furthermore, samples were studied by employing multivariate statistical analysis. Prior to chemometrics the NMR spectra were automatically processed as described in the Experimental Part (Chapter 3). Specifically, NMR spectra were normalized to the total intensity and binned in equidistant size buckets (0.02 ppm). At the end, an NMR table was generated with 76 rows corresponding to the individual carob syrup samples and 504 columns corresponding to the different variables (spectral bins). The NMR data table was imported into SIMCA 14.0, and variables were centered and pareto scaled in order to make the features more comparable. Principal Component Analysis (PCA) is an unsupervised method used to examine variation in the dataset, visualise clustering of the samples' groups, and detect possible outliers. PCA scores plot of all samples (Figure 35) did not provide a clear classification between the samples (commercial vs traditional syrups). The model was built with 5 principal components with $R^2X(cum)$ was 0.904 and $Q^2(cum)$ was 0.778. The first two principal components (PCs) explained 70.6 % of the total variance.

The samples D1on, G1 and SL01 appeared as outliers. After a more careful study and observations of the spectra and the loading plots, the differentiation of the outliers of samples D1on and G1 is attributed to the characteristic alcohol peaks (added by the manufacturer), while the Slovenian sample shoed significant differences from all other spectra.

Scores Plot **Loadings Plot**

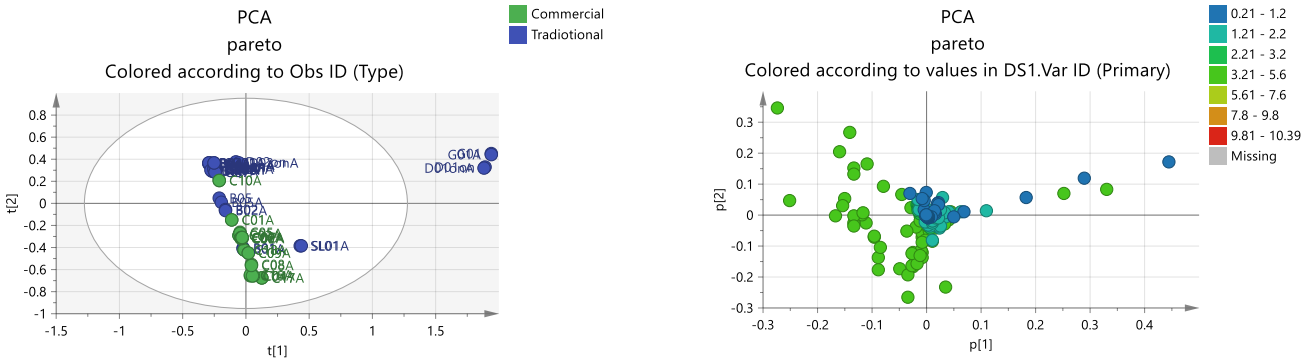


Figure 35. PCA-pareto (Scores and Loadings plots), coloured according to type (Green colour is for commercial and blue for traditional carob syrups).

To further study similarities of samples, the outliers were excluded, and clustering of the commercial and traditional carob syrup was achieved (Figure 36). PCA model was 5 principal components and the first 2 PCs explained 73.6 % of the total variance. Upon re-examination of the scores and loading plots, it appears that the clustering is attributed to the peaks of sugars. This observation led to the decision of removing the variables corresponding to the sugars for further investigation of their phenolic and aliphatic characteristics.

Scores Plot **Loadings Plot**

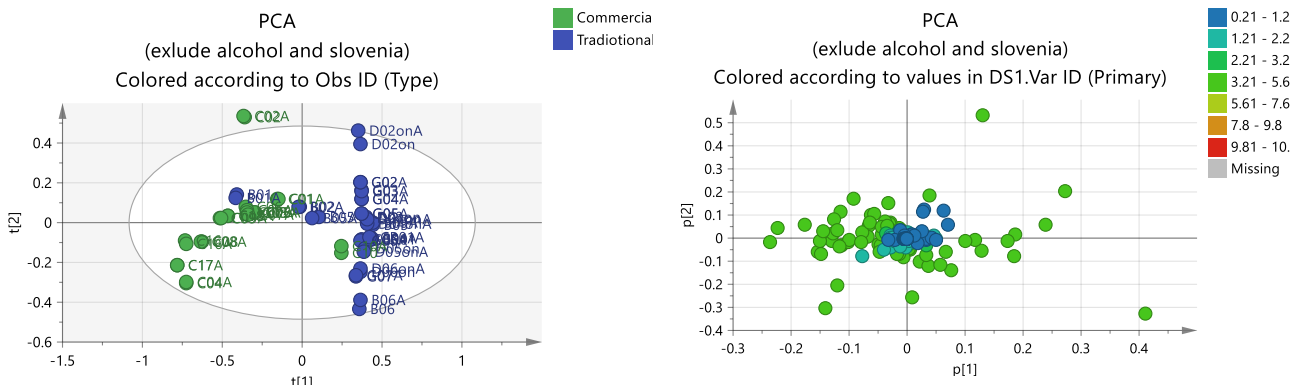


Figure 36. PCA analysis after excluding alcohol and Slovenia sample (Scores and Loadings plots).

The new PCA model was constructed again with 5 principal components and the first 2 PCs explained 56.9% of the total variance. In this model, after excluding the variables

corresponding to the sugar peaks, better clustering of the two group of samples was achieved on PC1, while intra group differences were observed along PC2. The respective scores and loadings plots are presented in Figure 37.

Scores Plot	Loadings Plot

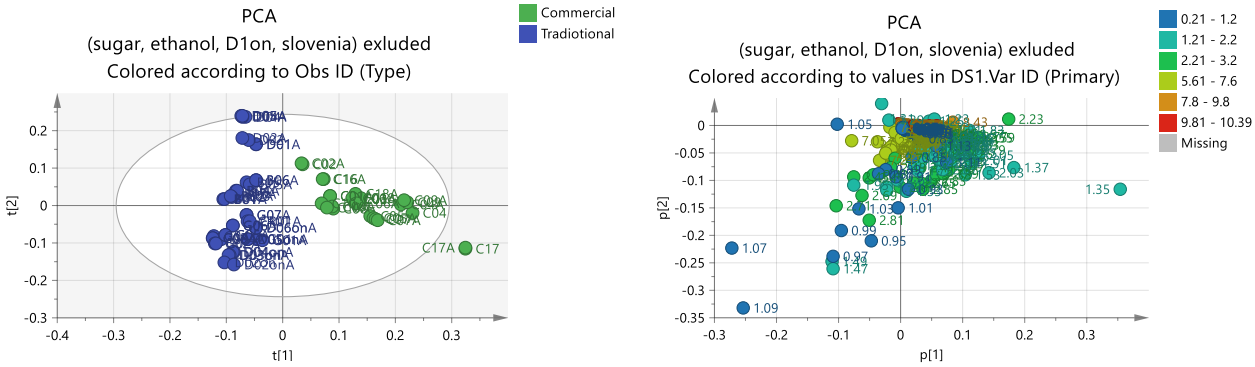


Figure 37. PCA analysis after excluding the variables of sugars, ethanol, D1, Slovenian syrups (Scores and Loadings plots).

These preliminary results strengthen our hypothesis for the removal of sugars and the in-depth study of the phenolic compounds in the future after the appropriate pre-treatment of samples (e.g., use of adsorption resin technology). Finally, more samples and replicates thereof are needed to build validated chemometric models, such as Orthogonal and/or Partial Least Squares Discriminant Analysis (PLS-DA) approached.

5. Chapter 5
Conclusions

Carob syrup appears to be a functional food with advanced nutritional value. In total, more than 48 compounds have been detected, many of which are characterized as nutraceuticals. The identified metabolites belong to 7 major categories. The richest chemical category is that of amino acids, accounting for 27%, followed by organooxygen compounds/carbohydrates + carbohydrate isomers and organic acids with a percentage of 19%, confirming thus its characterization as a natural chocolate. More specifically, commercial carob syrups are abundant in alanine (mean concentration = 16.4 mg kg^{-1}) and D-glucose (mean concentration = 12.4 mg kg^{-1}), while the corresponding traditional samples in alanine (mean concentration = 14.7 mg kg^{-1}) and L-serine (mean concentration = 10.8 mg kg^{-1}), respectively.

In commercial samples, the gallic acid, glutamic acid, L-serine, and aspartic acid concentrations varied significantly, showing variations in the production procedure. This highlights the necessity for process standardization. The results revealed that the overall profiles of commercial and traditional carob syrups are comparable, with the exemption of the values of some metabolites (pyroglutamic acid, glucose, and theobromine) that need more research. The values of organic acids (40.87%), glucose (88.50%), and theobromine (85.94%) all showed a substantial difference, with commercial carob syrups showing noticeably higher values. Traditional carob syrups were 10% richer in amino acids and flavonoids, although their average concentrations did not differ noticeably. In the house-made traditional samples the parameters of temperature and extraction time seem to play an important role in the extraction of certain metabolites, either positively or negatively. The concentrations of amino acids, organic acids, and theobromine rose with increasing temperature (20-100°C) and extraction time (5'-35'), whereas the concentration of glucose decreased at temperatures above 80°C. The greatest increase in concentration was observed after 25 minutes of extraction. The observed differences in the concentrations of specific metabolites in commercial and traditional carob syrups from other countries cannot lead to a solid conclusion due to the limited number of samples studied. However, this was expected due to the different production methods and growing environment.

NMR complemented the MS results and primarily highlighted the uniqueness of the aromatic and aliphatic regions, as well as the abundance of sugars (sucrose, D-xylose, D-glucose, D-lactose, sorbitol). Furthermore, there was an indication of possible presence of the

additional (non-natural) sugars, especially in the specific case of commercial syrups. The increase in temperature (20-100 °C) during the preparation of traditional carob syrups led to the increase of specific metabolites, revealing the importance of temperature during the production stage. The differences in spectra between different regions of Cyprus were not significant, as the obtained spectra were nearly identical, with some exceptions in certain peaks. This leads to the conclusion that the selected region does not affect the presence or absence of a metabolite, but only its concentration. This may be attributed to the microenvironment, the geochemical elements of the soil, and the small distances between regions within Cyprus. Further chemometric processing of NMR results, after the removal of sugars, enabled in the PCA the differentiation of commercial from traditional carob syrups. This led to the conclusion that commercial carob syrups have a similar composition, and that the aromatic and aliphatic regions can provide significant information regarding the authenticity and origin of the product. Although all the carob syrups were prepared under the same conditions, the Slovenian product seems to present more differences compared to the others (Croatian, Australian, Greek). However, this needs further investigation due to the limited number of samples.

6. Chapter 6
Future Work

a) The aromatic and aliphatic region can provide significant information regarding the authenticity of the product, after appropriated pre-treatment of samples, with extraction techniques for sugar removal (e.g., use of adsorption resin technology). These techniques should be implemented to strengthen the secondary variables. This could be combined with authenticity checks in commercial carob syrups.

b) Additionally, there is potential for new experiments to discover the food and medical effects of carob syrup. With the knowledge of syrup ingredients, research focus could be performed on a specific category of compounds for better understanding its pharmaceutical and nutraceutical effects.

c) Lastly, a more in-depth examination of the production parameters (e.g., extraction time, region, etc.), is suggested while keeping constant the other factors to avoid influencing of the results (one variable at a time approach).

7. Abbreviations

ΕΚΠΑ	Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
1D	One Dimension
2D	Two Dimensions
3D	Three Dimensions
APCI	Atmospheric Pressure Chemical Ionization
AQ	Acquisition Time
bbCID	Broadband Collision-Induced Dissociation
CID	Collision Induced Dissociation
COSY	Correlated Spectroscopy
D_1	Relaxation Delay
DC	Direct current
DDA	Data Dependent Acquisition
DESI	Desorption Electrospray Ionization
DIA	Data Independent Acquisition
DOEs	Design of experiments
DSS	Dimethyl-4-Silapentane-1-Sulfonic Acid
ECs	Emerging Contaminants
EMA	Economically Motivated Adulteration
ESI	Electrospray Ionization
FAO	Food and Agriculture Organization of the United Nations
FDA	Food and Drug Administration
FID	Free Induction Decay
FT	Fourier transformation
FTICR	Fourier Transform Ion Cyclotron Resonance
FWHM	Full Width at Half Maximum
GC	Gas Chromatography
GC-MS	Gas Chromatography- Mass Spectrometry
HCD	High-Energy Collision-Induced Dissociation
HMF	Hydroxymethylfurfural
HR	High Resolution
HRMS	High Resolution Mass Spectrometry
IS	Internal Standard
ISO	International Organization for Standardization

LC	Liquid Chromatography
LC-MS	Liquid Chromatography- Mass Spectrometry
LOD	Limit of Detection
LOQ	Limit of Quantitation
m/z	Mass-to-charge Ratio
MALDI	Matrix-Assisted Laser Desorption/Ionization
MF	Molecular Formula
MoNA	Mass Bank of North America
MS	Mass Spectrometry
MSI	Mass Spectrometry Imaging
MW	Molecular Weight
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear OverHauser Effect Spectroscopy
NS	Number of Scans
OT	Orbitrap
PCA	Principal Components Analysis
PLS	Partial Least Squares
Q	Quadrupole
QSPR	Quantitative Structure-Property Relationships
R^2	Explained Initial Variance
RF	Radio Frequency
RP	Reverse Phase
RTI	Retention Time Prediction
S/N	Signal to Noise Ration
S_{bl}	Standard Deviation of the blank
$\overline{S}_{bl}$	Sum of the Mean Blank Signal
SD	Standard Deviation
SDGs	Sustainable Development Goals
S_m	Minimum Analytical Signal
SW	Spectral Width
TMSP	Trimethylsilylpropanoic-2,2,3,3-d4 Acid Sodium Salt
TOCSY	Total Correlated Spectroscopy
TOF	Time-Of Flight
t_R	Retention Time

TSP	Trimethylsilyl Propanoic Acid
U-HPLC	Ultra-High Performance Liquid Chromatography
UHPLC-QTOF-ESI-MS	Ultra-High Performance Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry
VOCs	Volatile Organic Compounds
YSTAT	National Institute of Statistics

8. Chapter 8
References

1. María Emilia Brassesco, Teresa R.S. Brandão, Cristina L.M. Silva, Manuela Pintado, “Carob bean (Ceratonia siliqua L.): A new perspective for functional food”, Trends in Food Science & Technology, Vol. 114, 2021, p. 310-322.
2. Aristea Gioxari, Charalampia Amerikanou, Irini Nestoridi, Eleni Gourgari, Harris Pratsinis, Nick Kalogeropoulos, Nikolaos K. Andrikopoulos, Andriana C. Kaliora, “Carob: A Sustainable Opportunity for Metabolic Health”, Foods, Vol. 11, 2022, p. 2154.
3. Panagiota Fella, Kyriaki Kaikiti, Marinos Stylianou, Agapios Agapiou, “HS-SPME-GC/MS Analysis for Revealing Carob's Ripening”, Metabolites, Vol. 15;12(7), 2022, p. 656.
4. Meijun Xing, Shenghao Wang, Jianzhong Lin, Feng Xia, Jianghua Feng, Guiping Shen, “Composition Profiling and Authenticity Assessment of Camellia Oil Using High Field and Low Field 1H NMR”, Molecules, Vol. 26, 2021, p. 4738.
5. Chrystalleni Lazarou, Της χαρουπιάς μελώματα-Πόπαστα καμωμένα με το τερατσόμελον, Museum of Cypriot Food and Nutrition, 2023, p.14.
6. Papaefstathiou, E., Agapiou, A., Giannopoulos, S., & Kokkinofta, R., “Nutritional characterization of carobs and traditional carob products”, Food Science and Nutrition, Vol. 6(8), 2018, p. 2151–2161.
7. Zakaria, Z. A., Boopathy, R., & Dib, J. R., “Valorisation of agro-industrial residues”, S. Nature, Republic of Cyprus, Statistical service, Agricilture Volume I: Biological approaches, YSTAT 2021.
8. Azadeh Vafaei, Shabnam Mohammadi, Alireza Fazel, Mohammad Soukhtanloo, Abbas Mohammadi pour, Farimah Beheshti, “Effects of Carob (Ceratonia siliqua) on Sperm Quality, Testicular Structure, Testosterone Level and Oxidative Stress in Busulfan-Induced Infertile Mice”, Pharmaceutical Sciences, Vol. 24, 2018, p. 104-111.
9. Cited April 2023, Available from: https://ucyweb.ucy.ac.cy/carob/el/the-programme/carob-history.
10. Georgia D. Ioannou, Ioanna K. Savva, Atalanti Christou, Ioannis J. Stavrou, Constantina P. Kapnissi-Christodoulou, “Phenolic Profile, Antioxidant Activity, and Chemometric Classification of Carob Pulp and Products”, Molecules, Vol 28, 2023, p. 2269.

11. Vlasios Goulas, Evgenios Stylos, Maria V. Chatziathanasiadou, Thomas Mavromoustakos, Andreas G. Tzakos, “Functional Components of Carob Fruit: Linking the Chemical and Biological Space”, Int. J. Mol. Sci., Vol 17, 2016, p. 1875.
12. Cited April 2023, Available from: http://www.unesco.org.cy/Programmes-Kypriaki_charoypia_paradosiakes_technikes_kalliergeias_sygkomidis_epeksergasias,GR-PROGRAMMES-04-02-03-28,GR.
13. Stavroula A. Vekiari, Georgia Ouzounidou, Münir Ozturk, Güven Görk, “Variation of quality characteristics in Greek and Turkish carob pods during fruit development”, Procedia - Social and Behavioral Sciences, Volume 19, 2011, p. 750-755.
14. Mohamed Fawzy Ramadan, Mohamed A. Farag, “Mediterranean Fruits Bio-wastes”, Chemistry, Functionality and Technological Applications, p. 765-798.
15. Ummay Mahfuza Shapla, Md. Solayman, Nadia Alam, Md. Ibrahim Khalil, Siew Hua Gan, “5-Hydroxymethylfurfural (HMF) levels in honey and other food products: effects on bees and human health”, Chemistry Central Journal, Vol. 12:35, 2018, p. 1-18.
16. Chrysanthi Christou, Evdokia Poulli, Stelios Yiannopoulos, Agapios Agapiou, “GC–MS analysis of D-pinitol in carob: Syrup and fruit (flesh and seed)”, Vol. 1116, 2019, p. 60-64.
17. Chunxiu Hu, Guowang Xu, “Mass-spectrometry-based metabolomics analysis for foodomics”, TrAC Trends in Analytical Chemistry, Vol. 52, 2013, p. 36-46.
18. Yong-Jiang Xu, Xi Wu, “Foodomics in microbiological investigations”, Current Opinion in Food Science, Vol. 4, 2015, p. 51-55.
19. Alberto Valdés, Alejandro Cifuentes, Carlos León, “Foodomics evaluation of bioactive compounds in foods”, TrAC Trends in Analytical Chemistry, Vol. 96, 2017, p. 2-13.
20. Cited May 2023, Available from: http://foodomics.gr/.
21. Alberto Valdés, Gerardo Á lvarez-Rivera, Bárbara Socas-Rodríguez, Miguel Herrero, Elena Ibáñez, Alejandro Cifuentes, “Foodomics: Analytical Opportunities and Challenges”, Anal. Chem., Vol. 94, 2022, p. 366−381.
22. Cited May 2023, Available from: https://masspec.scripps.edu/learn/ms/.
23. Sagar Aryal, “Mass Spectrometry (MS) - Principle, Working, Parts, Steps, Uses”, 2022.
24. Cited April 2023, Available from: https://eclass.uoa.gr/modules/document/file.php/CHEM165/07-SAT_02_Mass_Spectrometry_pt1.pdf.
25. CS Ho CL, MHM Chan, RCK Cheung, LK Law, LCW Lit, KF Ng, MWM Suen, and HL Tai. “Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications”, Clin. Biochem., 2003, p. 24.

26. Yuling Li, Junbo Zhao, Yinlong Guo, “The Principles and Applications of Electrospray-Based Ambient Ionization”, Vol. 31Issue (1), 2019, p. 94-109.
27. Louie KB, Kosina SM, Hu Y, Otani H, de Raad M, Kuftin AN, et al. “Mass Spectrometry for Natural Product Discovery”, Comprehensive Natural Products III, 2020, p. 263-306.
28. Hermans J, Ongay S, Markov V, Bischoff R., “Physicochemical Parameters Affecting the Electrospray Ionization Efficiency of Amino Acids after Acylation”, Analytical Chemistry, Vol. 89(17), 2017, p. 9159-66.
29. Amina Bouslimani, Laura M Sanchez, Neha Garg, and Pieter C Dorrestein, “Mass spectrometry of Natural Products: Current, Emerging and Future Technologies”, Nat. Prod. Rep., Vol. 31(6), 2014, p. 718–729.
30. Rebane R, Kruve A, Liigand P, Liigand J, Herodes K, Leito I. “Establishing Atmospheric Pressure Chemical Ionization Efficiency Scale”, Analytical Chemistry, Vol 88(7), 2016, p. 435-9.
31. Olmo-Garcia L, Kessler N, Neuweger H, Wendt K, Olmo-Peinado JM, Fernandez-Gutierrez A, et al. “Unravelling the Distribution of Secondary Metabolites in Olea europaea L.: Exhaustive Characterization of Eight Olive-Tree Derived Matrices by Complementary Platforms (LC-ESI/APCI-MS and GC-APCI-MS) ”, Molecules, Vol. 23(10), 2018.
32. Allen DR, McWhinney BC. “Quadrupole Time-of-Flight Mass Spectrometry: A Paradigm Shift in Toxicology Screening Applications”, Clin. Biochem. Rev., Vol 40(3), 2019, p. 135-46.
33. Eric J. Reiner, Karl J. Jobst, David Megson, Frank L. Dorman, Jean-François Focant, “Chapter 3-Analytical Methodology of POPs”, Environmental Forensics for Persistent Organic Pollutants, 2014, p. 59-139.
34. Cited April 2023, Available from: https://webmail.life.nthu.edu.tw/~labcjw/BioPhyChem/Mass/quadrupo.htm.
35. Edmond de Hoffmann VS, “Mass Spectrometry, Principles and Applications”, 2007.
36. Karoly Heberger, “Chapter 7–Chemoinformatics-multivariate mathematical–statistical methods for data evaluation”, Medical Applications of Mass Spectrometry, 2008, p. 141-169.
37. M. Guilhaus, “Time of flight mass spectrometry”, Trends in Analytical Chemistry, 2021, p. 4.
38. Jabbour RE, Snyder AP., “14 - Mass spectrometry-based proteomics techniques for biological identification”, **This chapter has been written by two employees of the US Army Research, Development and Engineering Command (RDECOM), operated by

the Edgewood Chemical Biological Center for the US Department of Defense-Army. In: Schaudies RP, editor. Biological Identification: Woodhead Publishing, 2014. p. 370-430.

39. Ruedi Aebersold, Matthias Mann, “Mass spectrometry-based proteomics”, Nature Volume, Vol. 422, 2003, p. 198–207.
40. Stewart D, Dhungana S, Clark R, Pathmasiri W, McRitchie S, Sumner S. “Chapter 4 - Omics Technologies Used in Systems Biology”, Systems Biology in Toxicology and Environmental Health. Boston: Academic Press, 2015. p. 57-83.
41. Allen DR, McWhinney BC. “Quadrupole Time-of-Flight Mass Spectrometry: A Paradigm Shift in Toxicology Screening Applications”, Clin. Biochem. Rev., Vol. 40(3), 2019, p. 135-46.
42. Nika, M.C., PhD thesis: “Identification of transformation products of emerging contaminants during tertiary treatment processes and their disposal in the envirinment by mass spectrometric techniques”, 2017.
43. Dasenaki ME, Thomaidis NS, “Multi-residue determination of 115 veterinary drugs and pharmaceutical residues in milk powder, butter, fish tissue and eggs using liquid chromatography-tandem mass spectrometry”, Anal. Chim. Acta., Vol. 880, 2015, p. 103-21.
44. Panara A, Katsa M, Kostakis M, Bizani E, Thomaidis NS., “Monitoring of Aflatoxin M1 in Various Origins Greek Milk Samples Using Liquid Chromatography Tandem Mass Spectrometry”, Separations, Vol. 9, 2022.
45. Hajrulai-Musliu Z, Uzunov R, Jovanov S, Jankuloski D, Stojkovski V, Pendovski L, et al., “A new LC-MS/MS method for multiple residues/contaminants in bovine meat”, BMC Chem, Vol. 15(1), 2021, p.62.
46. Xinjing Dou LZ, Ruinan Yang, Xiao Wang,Li Yu, Xiaofeng Yue, Fei Ma, Jin Mao, Xiupin Wang, Wen Zhang, Peiwu Li., “Mass spectrometry in food authentication and origin traceability”, Mass Spectrometry Reviews, 2022.
47. Yukihiro Y, Zaima N., “Application of Mass Spectrometry Imaging for Visualizing Food Components”, Foods, Vol. 9(5), 2020.
48. Petrie B, Youdan J, Barden R, Kasprzyk-Hordern B. “Multi-residue analysis of 90 emerging contaminants in liquid and solid environmental matrices by ultra-high-performance liquid chromatography tandem mass spectrometry”, J. Chromatogr. A., Vol. 1431, 2016, p. 64-78.

49. Hernandez F, Sancho JV, Ibanez M, Abad E, Portoles T, Mattioli L., “Current use of high-resolution mass spectrometry in the environmental sciences”, Analytical and Bioanalytical Chemistry, Vol. 403(5), 2012, p. 1251-64.
50. Gramowska H, Krzyzaniak I, Baralkiewicz D, Goldyn R. “Environmental applications of ICP-MS for simultaneous determination of trace elements and statistical data analysis”, Environ. Monit. Assess., Vol. 160(1-4), 2010, p. 479-90.
51. Van Wijk XMR, Goodnough R, Colby JM, “Mass spectrometry in emergency toxicology: Current state and future applications”, Crit. Rev. Clin. Lab Sci., Vol. 56(4), 2019, p. 225-38.
52. Wen B, Zhu M. “Applications of mass spectrometry in drug metabolism: 50 years of progress”, Drug Metab. Rev., Vol. 47(1), 2015, p. 71-87.
53. Géhin C, Holman SW, “Advances in high-resolution mass spectrometry applied to pharmaceuticals in 2020: A whole new age of information”, Analytical Science Advances, Vol. 2 (3-4), 2021, p. 142-56.
54. Hollender J, van Bavel B, Dulio V, Farmen E, Furtmann K, Koschorreck J, et al., “High resolution mass spectrometry-based non-target screening can support regulatory environmental monitoring and chemicals management”, Environmental Sciences Europe, Vol. 31(1), 2019.
55. Christine M. Fisher, Katherine T. Peter, Seth R. Newton, Andrew J. Schaub, Jon R. Sobus, “Approaches for assessing performance of high-resolution mass spectrometry–based non-targeted analysis methods”, Analytical and Bioanalytical Chemistry, Vol. 414, 2022, p. 6455–6471.
56. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. “Identifying small molecules via high resolution mass spectrometry: communicating confidence”, Environmental Science & Technology, Vol. 48(4), 2014, p. 2097-8.
57. Murray KK, Boyd RK, Eberlin MN, Langley GJ, Li L, Naito Y. “Definitions of terms relating to mass spectrometry”, Pure and Applied Chemistry, Vol. 85(7), 2013, p. 1515-609.
58. Li H, Yu Y, Wang Z, Geng J, Dai Y, Xiao W, et al. “Chemical profiling of Re-Du-Ning injection by ultra-performance liquid chromatography coupled with electrospray ionization tandem quadrupole time-of-flight mass spectrometry through the screening of diagnostic Ions in MS(E) mode”, PloS one, Vol. 10 (4), 2015, p. 121-31.
59. Zheng C, Hao H, Wang X, Wu X, Wang G, Sang G, et al., “Diagnostic fragment-ion-based extension strategy for rapid screening and identification of serial components of homologous families contained in traditional Chinese medicine prescription using high-

resolution LC-ESI- IT-TOF/MS: Shengmai injection as an example", J. Mass Spectrom, Vol. 44(2), 2009, p. 230-44.

60. Franski R, Gierczyk B, Kozik T, Popenda L, Beszterda M., "Signals of diagnostic ions in the product ion spectra of [M - H](-) ions of methoxylated flavonoids", Rapid communications in mass spectrometry: RCM, Vol. 33(1), 2019, p. 125-32.
61. Cited April 2023, Available from: https://msbi.ipb-halle.de/MetFrag/.
62. Djoumbou-Feunang Y, Pon A, Karu N, Zheng J, Li C, Arndt D, et al., "CFM-ID 3.0: Significantly Improved ESI-MS/MS Prediction and Compound Identification", Metabolites, Vol 9(4), 2019.
63. Aalizadeh R, Nika MC, Thomaidis NS. "Development and application of retention time prediction models in the suspect and non-target screening of emerging contaminants", Journal of Hazardous Materials, Vol. 363, 2019, p. 277-85.
64. Krauss M, Singer H, Hollender J., "LC-high resolution MS in environmental analysis: from target screening to the identification of unknowns", Analytical and Bioanalytical Chemistry, Vol. 397(3), 2010; p. 943-51.
65. Hollender J, van Bavel B, Dulio V, Farmen E, Furtmann K, Koschorreck J, et al. "High resolution mass spectrometry-based non-target screening can support regulatory environmental monitoring and chemicals management", Environmental Sciences Europe, Vol. 31, 2019.
66. Friebolin H & Becconsall JK., "Basic one-and two-dimensional NMR spectroscopy", Weinheim: Wiley-vch, Vol. 7, 2005.
67. Antigoni Cheilari, PhD thesis: "NMR and LC-MS -based Methodologies for the Characterisation of Medicinal Plants", 2016.
68. Jacobsen NE., "NMR data interpretation explained: understanding 1D and 2D NMR spectra of organic compounds and natural products", John Wiley & Sons, 2016.
69. Derome AE. "Modern NMR techniques for chemistry research", Elsevier, 2013.
70. Cited April 2023, Available from: https://www.technologynetworks.com/analysis/articles/nmr-spectroscopy-principles-interpreting-an-nmr-spectrum-and-common-problems-355891#:~:text=NMR%20spectroscopy%20is%20a%20physicochemical,atomic%20nuclei%20called%20nuclear%20spin.
71. "Nuclear Magnetic Resonance Spectroscopy - AQA Chemistry A-level", Vol. 3.15, p.1-5.

72. Mukesh Kumar Singh, Annika Singh, “Chapter 14-Nuclear magnetic resonance spectroscopy: Characterization of Polymers and Fibres”, The Textile Institute Book Series, 2022, p. 321-339
73. Bible RH. “Interpretation of NMR spectra: an empirical approach”, Springer Science & Business Media, 2013.
74. Maciel GE. “NMR in industrial process control and quality control”, In Nuclear Magnetic Resonance in Modern Technology, p. 225-275.
75. Diercks T, Coles M, & Kessler H., “Applications of NMR in drug discovery”, Curr Opin Chem Biol., Vol. 5(3), 2001, p. 285-291.
76. Olaide AJ, Olugbenga E & Abimbola D. “A review of the application of nuclear magnetic resonance in petroleum industry”, Int. J. Geosci., Vol. 11(4), 2020, p. 145.
77. Brown SP. “Applications of high-resolution 1H solid-state NMR. Solid State”, Nucl. Mag., Vol. 41, 2012, p. 1-27.
78. Ishima R, & Torchia DA., “Protein dynamics from NMR”, Nat Struct Biol., Vol. 7(9), 2000, p. 740-743.
79. Takeuchi K, & Wagner G., “NMR studies of protein interactions”, Curr Opin Struct Biol., Vol. 16(1), 2006, p. 109-117.
80. Markley JL, Brüschweiler R, Edison AS et al., “The future of NMR-based metabolomics”, Curr. Opin. Biotech., Vol. 43, 2017, p. 34-40.
81. Vlaardingerbroek MT & Boer JA. “Magnetic resonance imaging: theory and practice”, Springer Science & Business Media, 2013.
82. Hatzakis E, “Nuclear magnetic resonance (NMR) spectroscopy in food science: A comprehensive review”, Compr. Rev. Food Sci. F., Vol. 18(1), 2019. p. 189-220.
83. G. A. Nagana Gowda, Daniel Raftery, “NMR Based Metabolomics”, Adv. Exp. Med. Biol., Vol. 1280, 2021, p. 19–37.
84. Lorraine Brennan, “NMR-based metabolomics: From sample preparation to applications in nutrition research”, Progress in Nuclear Magnetic Resonance Spectroscopy, Vol. 83, 2014, p. 42-49.
85. Zikuan Song, Haoyu Wang, Xiaotong Yin, Pengchi Deng, Wei Jiang, “Application of NMR metabolomics to search for human disease biomarkers in blood”, Clin. Chem. Lab Med., Vol. 57(4), 2019, p. 417–441.
86. Cited April 2023, Available from: https://docs.tibco.com/pub/stat/14.0.0/doc/html/UsersGuide/GUID-F02AD2C4-E4C1-4E55-95A9-4E18F3D4B72D.html.

87. Cited March 2023, Available from: https://blog.minitab.com/en/applying-statistics-in-quality-projects/design-of-experiments-fractionating-and-folding-a-doe.
88. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S., “MetFrag relaunched: incorporating strategies beyond in silico fragmentation”, Journal of Cheminformatics, Vol.8, 2016, p. 3.
89. Fei Wang, Jaanus Liigand, Siyang Tian, David Arndt, Russell Greiner, David S. Wishart, “CFM-ID 4.0: More Accurate ESI-MS/MS Spectral Prediction and Compound Identification”, Anal. Chem., Vol. 93, 2021, p. 11692−11700.
90. Kind T, Fiehn O. “Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry”, BMC Bioinformatics, Vol.8, 2007, p. 105.
91. Cited May 2023, Available from: https://www.creative-proteomics.com/services/unknown-metabolites-identification.htm#:~:text=Defined%20by%20the%20Metabolomics%20Standards,unknown%20compounds%20(level%204).
92. Catherine Deborde, Jean-Xavier Fontaine, Daniel Jacob, Adolfo Botana, Valérie Nicaise, Florence Richard-Forget, Sylvain Lecomte, Cédric Decourtil, Kamar Hamade, François Mesnard, Annick Moing, Roland Molinié, “Optimizing 1D 1 H-NMR profiling of plant samples for high throughput analysis: extract preparation, standardization, automation and spectra processing”, Vol 15, Metabolomics (2019), p. 28.
93. Anthi Panara, Evagelos Gikas, Ilias Tzavellas and Nikolaos S. Thomaidis, “Comprehensive HRMS chemical characterization of pomegranate-based antioxidant drinks via a newly developed suspect and target screening workflow”, Molecules, 2023.
94. Sam Zhang LL, Ashok Kumar. Materials Characterization Techniques. Boca Raton2008.
95. Devices ITBaceom. ISO 10993-18:2020. Biological evaluation of medical devices — Part 18: Chemical characterization of medical device materials within a risk management process.
96. Géhin C, Holman SW. Advances in high-resolution mass spectrometry applied to pharmaceuticals in 2020: A whole new age of information, Analytical Science Advances, Vol. 2(3-4), 2021, p. 142-56.
97. Zhang C, Zuo T, Wang X, Wang H, Hu Y, Li Z, et al. Integration of Data-Dependent Acquisition (DDA) and Data-Independent High-Definition MS(E) (HDMS(E)) for the

Comprehensive Profiling and Characterization of Multicomponents from Panax japonicus by UHPLC/IM-QTOF-MS, Vol. 24(15), Molecules, 2019.

98. Menelaos Papagianopoulos, Hans Rainer Wollseiffen, Annett Mellethin, Bernd Haber, Rudolf Galensa, Identification and Quantification of Polyphenols in Carob Fruits (Ceratonia siliqua L.) and Derived Products by HPLC-UV-ESI/MS^n, J. Agric. Food Chem., Vol. 52, 2004, p. 3784−3791.
99. Esfanjani AF, Assadpour E, Jafari SM, Improving the bioavailability of phenolic compounds by loading them within lipid-based nanocarriers, Trends in Food Science & Technology, Vol. 76, 2018, p. 56-66.
100. Ioannis J. Stavrou, Atalanti Christou, Constantina P. Kapnissi-Christodoulou, Polyphenols in carobs: A review on their composition, antioxidant capacity and cytotoxic effects, and health impact, Food Chemistry, 2018.
101. Atalanti Christou, Ana B. Martinez-Piernas, Ioannis J. Stavrou, Juan F. Garcia-Reyes, Constantina P. Kapnissi-Christodoulou, HPLC-ESI-HRMS and chemometric analysis of carobs polyphenols – Technological and geographical parameters affecting their phenolic composition, Journal of Food Composition and Analysis, Vol.114, 2022, p. 104744.
102. Olga Deda, Olga Begou, Helen Gika, Georgios Theodoridis, Agapios Agapiou, Optimization of Carob Products Preparation for Targeted LC-MS/MS Metabolomics Analysis, Metabolites, Vol. 13, 2023, p. 645.