Tutor/s

Dr. Oscar Núñez

Departament d'Enginyeria Química i Química Analítica

Dra. Sònia Sentellas

Departament d'Enginyeria Química i Química Analítica

Grau de Química

Treball Final de Grau

Classification and Authentication of Honey by Chromatographic Profiles and Chemometric Methods

Classificació i Autenticació de Mel mitjançant Perfils Cromatogràfics i Mètodes Quimiomètrics.

Clàudia Martínez i Alfaro

January 2022

Universitat de Barcelona

B:KC Barcelona Knowledge Campus
Campus d'Excel·lència Internacional

En primer lloc voldria agrair als meus tutors, l'Oscar i la Sònia, per haver-me donat el suport necessari durant aquests mesos i per haver compartit amb mi el seu temps i coneixements amb tanta confiança. També a tots els membres dels grups de Fluorescència i d'Equilibris en solució i Quimiometria pel seu acolliment, en especial als meus dos companys de laboratori, en Victor i en Xavi, el quals han fet d'aquest projecte una experiència molt maca.

Per altra banda, vull donar les gràcies a la meva família per haver fet possible la meva carrera com a química i per estar al meu costat sempre que ho necessito. També vull agrair-li als meus amics de sempre el haver sigut un pilar fonamental en el meu dia a dia durant tant de temps.

Per últim, agrair als meus companys de la universitat, sobretot a l'Ekiñe, la Montse i la Júlia, per haver fet d'aquests anys una etapa inoblidable plena de comprensió i bones paraules.

Gràcies.

REPORT

CONTENTS

1. SUMMARY

Nowadays, food industry is suffering a hard competitiveness that has led to a fraud increase in order to achieve greater economical profits. These irregular actions have raised the interest in the quality and authentication of commercialized food products, not only from the consumers and producers, but also from the scientific research. Honey can be found among the most vulnerable products to be adulterated and it is mostly affected by two fraud types: the mixing of monofloral honeys with multifloral honeys, which are considered of less quality, and the addition of cheaper sweeteners as syrups or industrial sugars. Even though these honeys are adulterated, they are labelled without all the information and they are sold as products of higher quality.

This project aims to stablish a simple method to characterize honeys from different botanical and geographical origins and to classify them considering the results. A non-targeted method and a simple sample pre-treatment has been chosen to obtain chromatographic fingerprints, using high-performance liquid chromatography with an ultraviolet detector (HPLC-UV). The chromatographic profiles are evaluated as chemical descriptors by using different chemometric methods such as principal component analysis (PCA), partial least-squares discriminant analysis (PLS-DA) and, in some cases, hierarchical cluster analysis (HCA).

The PCA results prove that the proposed method is robust and reproducible, and that separation of honey samples in different groups is possible. PLS-DA has shown a classification between different botanical origins of samples and this clustering has been supported by HCA, firstly between honeydew honeys and blossom honeys, and then between some floral varieties. On the other side, a good classification according to the region of production has not been accomplished, and it will be studied in further research.

Keywords: Honey, Food authentication, High-performance liquid chromatography, UV detection, Chromatographic fingerprints, Chemometrics.

2. RESUM

Actualment, la industria de l'alimentació pateix una situació de dura competitivitat que ha portat a un augment dels fraus per tal d'aconseguir un major benefici econòmic. Aquestes accions irregulars han incrementat l'interès per la qualitat i autenticació dels aliments que es comercialitzen, tant per part de clients i productors, com per part de la investigació científica. Entre els productes més vulnerables a ser adulterats es troba la mel, la qual està afectada per dos tipus de fraus principals: la mescla de mels monoflorals amb mels multiflorals que es consideren de pitjor qualitat, i l'addició d'altres edulcorants més barats com per exemple xarops o sucres industrials. Tot i estar adulterades, aquestes mels s'etiqueten sense tota la informació i es segueixen venent com a productes de major qualitat.

Aquest projecte té com a objectiu establir un mètode simple per caracteritzar mels de diferents orígens botànics i geogràfics i classificar-les posteriorment a partir dels resultats. S'ha escollit un mètode no dirigit i un pretractament de mostra senzill que permeti l'obtenció d'empremtes cromatogràfiques utilitzant la cromatografia de líquids d'alta eficàcia amb detector d'ultraviolat (HPLC-UV). Al llarg del treball s'avaluen aquests perfils cromatogràfics com a descriptors químics amb l'ús de diferents mètodes quimiomètrics com són l'anàlisi de components principals (PCA), l'anàlisi discriminant amb regressió de mínims quadrats (PLS-DA) i en algun cas, l'anàlisi d'agrupament jeràrquic (HCA).

Els resultats del PCA demostren que el mètode proposat és robust i reproduïble, així com que la separació de les mels en diferents grups és possible. El PLS-DA ha mostrat una classificació entre els diferents orígens botànics de les mostres i el HCA ha recolzat aquest agrupament, en primer lloc entre les mels de melada i les mels que provenen del nèctar de les flors, i en segon pla entre algunes varietats florals. Per altra banda, no s'ha aconseguit una bona classificació segons la regió de producció de la mel, camp en el que es continuarà treballant en estudis futurs.

Paraules clau: Mel, Autenticació alimentaria, Cromatografia líquida d'alta eficàcia, Detecció UV, Empremtes cromatogràfiques, Quimiometria.

3. INTRODUCTION

3.1. FOOD AUTHENTICITY

Nowadays, the food industry is threatened with food adulteration and fraud, mostly to gain greater benefits in a very competitive market. This has led to an increasing interest of the consumers to understand what they buy and consume. For this reason, it is important that the information given in the labels of food products is accurate and reliable, ensuring the quality that is being promised. There are some food products especially affected by the arising of different frauds forms and illicit practices, as for example honey [1].

The case of honey is particularly vulnerable to adulteration because of the variable composition of this product regarding different conditions (botanical origin, region of production or climate conditions), and the similarity between many adulterants and the natural components of honey. The differences between types of honeys have an important role when it comes to the composition and properties of the product. Depending on the variety and uniqueness of the honey, its price can change significantly. For example, monofloral honeys, which have one predominant botanical source, have higher prices in the market. Also, some regions can also increase the price of its products as they are protected, for example by means of protected designation of origin (PDO) attributions [2]. Hence, the possibility of committing fraud by giving uncomplete information of the product in its label (especially in multifloral honeys) or by mixing high-quality honeys with cheaper ones is always present [3].

Another way of committing fraud is through the addition of adulterants either in the final product or in the bees feeding. These honeys can contain syrups or other sugars that are cheaper sweeteners than natural honey, which needs a longer time and higher economical investment to be obtained. Typical adulterants are beet, cane, inversed, or industrial sugars and syrups [4].

The danger of frauds and the practice of adulterating food products in general has entailed the development of control techniques [1]. In fact, as honey is the third most adulterated product in the world, the European Union has developed plans and programmes to protect this particularly vulnerable product [5,6].

3.2. HONEY

Honey is a natural product that comes from the nectar and other non-floral secretions. Bees, in particular *Apic Mallifera* and stingless bee species, transport these plant substances and make them react with substances of their own until concentrated and ripped honey is obtained in hives [4]. Throughout the years, honey has taken an essential role in different societies because of its curative and nutritive properties, and it has become a very appreciated product [7].

Honey can be classified by its geographical or botanical origin. Considering its botanical origin, multifloral and monofloral honeys can be found depending on the pollen content. If more than 45% of the pollen content is from one plant or floral specie, the honey can be considered monofloral and it is named after the name of the main plant or flower source. By contrast, if there is no predominant source, the product is considered multifloral [3].

Another way of classifying honeys is taking into account which part of the plant are the bees treating to produce the honey. Blossom honeys are the ones that come from the nectar of the flowers while honeydew honeys come from plant secretions or substances that plant-sucking insects can excrete. In general, holm oak, forest and pine honeys belong to this last type of honey. These two groups of honeys have significant sensory, physicochemical, and bioactive differences. Along many of them, honeydew honeys tend to be darker and they have higher electric conductivity, higher total polyphenol content (leading to higher antioxidant properties), and higher ash content (which is related to the mineral content). They also have lower fructose content than blossom honeys. Some exceptions in these comparisons have been reported in literature. For instance, chestnut, heather, strawberry trees and eucalyptus honeys do not always fit with the normal values of blossom honeys ever though they come from the flower nectar [8].

3.2.1. Legislation

Because of the need of protecting honey and its authenticity, CODEX (which elaborates international food standards to ensure the innocuousness and quality of food products) published their international standards for honey, making a clear differentiation between honeydew honey and blossom honey. In this standard, it is said that honey cannot have any additives except other honeys (these products should be labelled as blends) and its properties, composition and quality cannot be changed in any step of the treatment process.

The moisture content must be less than 20% except for heather honey, which moisture content cannot be higher than 23%. The total fructose and glucose content has to be over 60g for 100 g of honey except for honeydew honeys or blends of them with blossom honey, which content of fructose and glucose can be lower, but always over 45 g for 100 g of honey. The sucrose content is also defined by this standard and it cannot be more than the 5% of the honey weight apart from some specific varieties. Lastly, water insoluble solids content must be under the 0.1% (w/w) and under the 0.5% for pressed honey. Referring to the labels, their information may include botanical and geographical origin as well as if the product is considered honeydew honey, blossom honey, or a blend of both honeys.

Other requirements exist regarding hygiene, analyses, contaminants and additional composition and quality factors, and they are all specified in the CODEX standard [9].

3.2.2. Composition

The composition of honey is complex and can change depending on the botanical and geographic origin, but it can be defined as a supersaturated sugar solution with other components that can vary based on the honey type.

Carbohydrates are the main component followed by water which represents the 17% of the total content. Other components are proteins, enzymes, amino acids, organic acids, minerals, polyphenols, vitamins, solid particles, and volatile compounds [3].

3.2.2.1. Carbohydrates

The average percentage of sugars in honey is about a 60% (w/w), and it increases to a 95% if we refer to the weight of dry honey.

The predominant carbohydrates are glucose and fructose, and this last one is specially affected by the botanical source. The relation between these two sugars is also influenced by the variety of the honey. After honey intake, they are rapidly transported to blood and they act as an immediate energy provider.

On the other hand, these are not the only carbohydrates in honey, but other sugars as disaccharides, trisaccharides, and oligosaccharides have been identified in several proportions. The total amount of carbohydrates defines great part of the nutritional value of honey and confers its flavour [10].

3.2.2.2. Polyphenols

Even though polyphenols are present in small quantities compared to carbohydrates, they are especially relevant because they can be used as chemical markers and they have beneficial effects on health because of their antioxidant activity. The prevailing polyphenols are phenolic acids and flavonoids, which come from the flower nectar, pollen and propolis. In many cases, they are present in glycosylated forms.

Flavonoids are oxygen containing heterocyclic compounds and they can be found in plants in their free form or linked to a sugar. These can be classified by the oxidation level or the position of the oxygen atoms in the second aromatic ring, and they conform a big group of compounds that allow to characterize honey. The total weight of flavonoids is estimated to 6 mg/kg, but it depends on the botanical origin. The most common flavonoids in honey are pinocembrin, apigenin, luteolin, myricetin, galangin, herperetin, quercetin, kaempferol, pinobanskin, and isohamnetin.

Phenolic acids are constituted by a phenolic ring and a carboxylic group. The content of these compounds changes due to environmental conditions, botanical origin and the region where the plant of origin grew. In consequence, they are potentially good chemical markers for classifying honey and differentiating between different varieties. In fact, many studies have been done with the aim of using phenolic acids for authentication. These compounds have also an important contribution on the pH. The most abundant phenolic acids in honey are usually *p*-hydroxybenzoic, caffeic, *p*-coumaric, gallic, vallinic, syringic, ferulic, protocatechuic, and sinapic acids [11].

3.2.2.3. Other components

There are other substances in honey that have an influence in their physical and sensorial characteristics even though they are not very important from a nutritional point of view.

The aroma of honey is determined by their volatile compounds, in which amino acids and organic acids stand out. The most abundant amino acid is proline, but others can be detected in lower quantities. As amino acids come from pollen, the proportion between them are highly influenced by the botanical characteristics of the honey.

Colour is conferred by minerals and pollen which are also detected in honey samples. Darker honeys are related with a higher content of polyphenols too. Moreover, despite that the content

of mineral and vitamins is remarkably low, they play an important role from a biomedical point of view [3,10].

3.2.3. Properties

Therapeutic uses of honey come from its biomedical properties conferred by the physical properties and the different components.

One of the medical properties of honey is the antibacterial activity, which can be attributed to its high osmolality and acidity, and to the peroxide acid formed from the reactions of glucose with the enzyme glucose oxidase from the honeybees.

Polyphenols give to honey its antioxidant activity because they prevent the formation of free radicals from oxidation processes. There are also clinic studies that sustain the idea of phenolic derivates having antitumoral and anti-inflammatory activity. That means they can be involved in the recovery of wounds and regeneration of tissues. In addition, it is known that honey can prevent heart diseases, avoid oxidation of LDL, and decrease the possibility of platelet clotting [11].

Finally, honey has also bactericide activity against those agents that can cause ulcers, gastritis and other gastrointestinal conditions because of the sugar content [10].

3.3. ANALYTICAL METHODOLOGIES FOR HONEY CHARACTERIZATION

There are many methods that allow to study honey samples from different perspectives. Classic methods are able to give information about physicochemical parameters such as sugar content, enzymatic activity, pH, electrical conductivity and moisture content. The limitations of these techniques make modern methods more suitable to the analysis of the botanical and geographical origin of honey samples.

Using modern techniques, profiles of various analytes that are of great interest can be obtained. Chromatographic techniques have been useful to identify honeys for their sugar, amino acid and polyphenolic profile. This technique is the most common for authentication and classification of food due its selectivity, sensibility and resolution, and because it can efficiently separate many components of a food sample. It is also fast and easily automatable, and allows to obtain many information that can be treated afterwards. Moreover, HPLC-UV and HPLC-MS are two of the most powerful techniques for the analysis of food and feed and for fraud prevention [12]. Depending on the methodology that is going to be used, some previous extraction

techniques such as liquid-liquid extraction (LLE), solid-phase extraction (SPE) or others, may be convenient to perform a successful analysis [11,13].

Other techniques have been used in order to characterize honey such as stable isotope ratio mass spectrometry (IRSM), that detect adulterated honeys based on several isotope ratios. Infrared spectroscopy (IR) is useful to determinate sugar, amino acids and adulterants. For the authentication of honey, molecular techniques that involve protein and DNA determinations have also been used. If the aim is to obtain chemical and structural information, nuclear magnetic resonance (NMR) is a good option and it also allows to classify honey samples. Other techniques as gravimetric methods, differential scanning calorimetry (DSC), atomic absorption spectrophotometry (AAS), inductively coupling plasma mass spectrometry (ICP-MS), ultraviolet spectroscopy (UV) or the study of dielectric properties have been used for the analysis of pure and adulterated honeys [3].

Once the required information is obtained from any of the previous methods, the characterization of honey samples can be done. There are two analytical strategies that can be employed for that: targeted and non-targeted approaches, which are explained in the following section.

3.3.1. Targeted analysis

Targeted analysis has the objective of detecting or quantifying specific chemical markers that are important for the authentication or characterization of a sample. Chemical makers are primary if they give direct information about the authentication matter, and they are secondary if they give information about the adulteration of the samples in an indirect manner.

This analytical strategy is usually quantitative and it often has better selectivity and sensibility than non-targeted methods. However, it can be limited because of the high number of modifications and adulterants that can be added to a food product, and the requirement of analytical standards for every employed marker [14].

3.3.2. Non-targeted analysis

Non-targeted or non-supervised analysis uses unspecific chemical descriptors to define samples, such as signals or peaks, but without the requirement of having previous information about the identity of the chemical compounds being monitored. The obtaining of these chemical descriptors is known as fingerprinting, and it usually gives qualitative results.

In this type of analyses the whole sample is studied, and they offer the opportunity of finding unexpected chemical compounds or detecting small changes between similar samples. This can be beneficial because of the complex forms of adulteration that are being done currently. To achieve the objective of avoiding fraud and adulteration, it is important that non-targeted methods give as many information as possible so that small differences between fingerprints are enough to discriminate between samples according to their botanical origin, presence of adulterants or region of production.

Due to the big amount of chemical data that can be obtained from the previously named modern instrumental techniques and non-targeted strategies, statistical analyses are needed in order to find patterns or differences among the fingerprints. One option for these statistical treatments are chemometric methods, which will be used in this project to classify honey samples [14,15].

3.4. CHEMOMETRICS

Chemometrics is a discipline of chemistry that uses mathematics and statistics to select measurement procedures and experiments that are optimal to provide as much important chemical information as possible by the analysis of chemical data, and to obtain knowledge about chemical systems.

Before using any chemometric method, pre-treatments of the information collected in data matrices can be made to adapt them to the pertinent analysis [16].

3.4.1. Principal component analysis (PCA)

Principal component analysis is a multivariate decomposition chemometric method that reduces the obtained data matrix (X-matrix) into principal components (PCs), which are uncorrelated and ordered variables. These components retain the variation of the original data, and the first principal component (PC1) is the one that represents most of the data variance followed by the subsequent (PC2, PC3…). If the original data matrix contains correlated variables, PCA can describe it with less variables (PCs) so its dimensionality is reduced [17].

PCA is an exploratory technique that can be useful to study food samples because it depicts the similarity of the analysed samples. Samples can be depicted in score plots, where the PCs are in the axis, and their distribution can show patterns that can help to distinguish variety, region

of production or other food product features. Similar samples appear clustered in the plot and can be separated from other sample groups [16].

3.4.2 Partial least square regression-discriminant analysis (PLS-DA)

Partial least square regression-discriminant analysis (PLS-DA) is a supervised classification method. The aim of this method is to define a model that can classify the studied samples into predefined classes that are contained in the Y-matrix. The X-matrix still contains all the data from the analysis. The correlation between these matrices is conducted through other variables, named latent variables (LVs), which also reduce the dimension of the data matrix as previously seen in PCA.

In this case, LVs can also be used as axis to depict the analysed samples into a score plot. The separation of the samples in clusters indicates an existing relation between samples previously grouped in the same sample class.

3.4.3 Hierarchical cluster analysis (HCA)

Hierarchical cluster analysis (HCA) is a non-supervised method where the studied objects are not considered of any class. This method calculates the distances between samples or objects using different algorithms that differ from the linkage function. After that, the samples that have the smallest distance between them are grouped together into clusters, and the same process is made repeatedly with the resulting clusters.

There are two types of methods to perform HCA. The first one is an agglomerative method, which builds clusters from the individual samples until one big group is obtained. Secondly, divisive methods start from one big cluster with all the samples in it and then it is divided into smaller clusters. Regardless of the chosen methods, the results can be depicted graphically in a dendrogram.

4. OBJECTIVES

The aim of this project is to develop a simple non-targeted high-performance liquid chromatography with ultraviolet detection (HPLC-UV) fingerprinting method to characterize and classify honey samples considering different variables such as their region of production or their botanical origin. To achieve this aim, the next steps will be performed:

1. Honey samples from several botanical and geographical origins will be submitted to a simple sample treatment to extract as many bioactive components as possible.
2. The resulting extracts will be analysed using HPLC-UV to obtain a fingerprint for each sample. These HPLC-UV fingerprints will then be evaluated as sample chemical descriptors to achieve sample characterization and classification by chemometrics.
3. PCA will be employed as an exploratory discrimination method to evaluate the reproducibility and robustness of the method and the chemometric results.
4. PLS-DA will be used to study the classification of honey samples by their botanical and geographical origin.
5. HCA will be applied to support the sample discrimination and classification accomplished by PLS-DA results.

5. EXPERIMENTAL SECTION

5.1. REAGENTS AND SOLUTIONS

For the mobile phase of the employed chromatographic method and the sample treatment, the used reagents were:

- Acetonitrile for UHPLC (Supergradient ACS quality, from Panreac, Barcelona, Spain)
- Formic acid (≥96% from Sigma-Aldrich, MO, USA)
- Methanol (HPLC gradient grade from Fischer Chemical, NH, USA)
- Milli-Q water. This purified water was obtained with a 0.22 μm nylon filter coupled to the Milli-Q Reference A+ system (from Merck, Darmstadt, Germany).

5.2. SAMPLES AND SAMPLE TREATMENT

The samples consisted of different honeys, including several botanical origins (varieties) and regions of production, that were purchased in supermarkets and markets in Spain. Two heather honeys were provided by Miel de Braña (León, Spain). A total number of 137 honey samples were analysed. The studied varieties were blossom (BL), eucalyptus (EU), forest (FO), heather (HE), holm oak (HO), mountain (MO), multifloral (MF), rosemary (RO) and thyme (TH) honeys. For these honey varieties, the samples were also classified for geographical origin which were Aragon (AR), Asturias (AS), Basque Country (BC), Cantabria (CN), Castile La Mancha (CM), Castile and Leon (CL), Catalonia (CT), Extremadura (E), Balearic Islands (BI), Navarre (N), France (F), Spain (S) and Spain and others (SO). Table 5.1 shows the number of honey samples for each botanical and geographical origin.

Table 5.1. Number of analysed honey samples considering botanical and geographical origin.

Botanical origin	Number of samples	Geographical origin	Number of samples	Geographical origin	Number of samples
Heather (HE)	18	Asturias	2	Castile and Leon	5
		Basque Country	1	Catalonia	2
		Cantabria	3	Extremadura	5

Mountain (MO)	6	Asturias	2	Castile La Mancha	1
		Castile and Leon	1	Catalonia	2
Blossom (BL)	12	Aragon	1	Catalonia	3
		Balearic Islands	2	Extremadura	1
		Cantabria	1	Spain	2
		Castile and Leon	2		
Eucalyptus (EU)	13	Andalusia	1	Catalonia	1
		Asturias	2	Extremadura	4
		Cantabria	1	Navarre	1
		Castile and Leon	2	Spain and others	1
Forest (FO)	10	Balearic Islands	2	Catalonia	1
		Cantabria	2	Spain	4
		Castile and Leon	1		
Holm Oak (HO)	10	Aragon	2	Extremadura	6
		Castile and Leon	2		
Multifloral (MF)	35	Asturias	1	Extremadura	4
		Balearic Islands	4	France	1
		Cantabria	1	Navarre	4
		Castile and Leon	7	Spain	1
		Castile La Mancha	2	Spain and others	2
		Catalonia	8		
Rosemary (RO)	26	Andalusia	1	Cantabria	2
		Aragon	4	Catalonia	6
		Asturias	1	Extremadura	4
		Balearic Islands	4	Navarre	1
		Castile and Leon	1	Spain	2
Thyme (TH)	7	Castile and Leon	1	Extremadura	2
		Castile La Mancha	2	Spain	1
		Catalonia	1		

The first step to analyse the honeys was to weight 1 g approximately of each liquid honey in a falcon tube. In the case of crystallised honeys (which is a normal state of natural honey), they were introduced in a water bath at 45 °C until they melted. After their homogenization and cooling, they were weighted in the same way. Then, 10 mL of Milli-Q water was added to the falcon tubes and the content was mixed using a VibraMix Vortex from OVAN (Barcelona, Spain) until dissolution of the honey. After that, the mixture was centrifuged with the Rotina 420 Centrifugue from Hettich (Tuttlingen, Germany) at 3500 rpm for 5 minutes. Then the extract was diluted with methanol for a better conservation of the organic honey samples in a 1:1 ratio into chromatographic injection vials and they were kept at 4 °C until the analysis. Honey samples were randomly analysed with the proposed HPLC-UV method.

For control matters, acetonitrile and a quality control (QC), which was prepared by mixing 50 µL of each sample extract, were also injected. The QC was used to detect any possible instrumental drifts and to assess robustness and reproducibility of the method and the chemometric results. To accomplish this, both vials were injected after every 10 samples.

5.3. INSTRUMENTATION AND CHROMATOGRAPHIC CONDITIONS

The honey chromatographic fingerprints were obtained with an Agilent HPLC 1200 Series instrument (CA, USA) that has a degasser (G1322A), a quaternary pump (G1311A) and an autosampler (G1329A). The detector used was a diode array detector (G4212B) which information was registered in a linked computer equipped with the Agilent ChemStation software.

The HPLC separation was performed with a Kinetex C-18 core-shell column (particle size of 2.6 µm; 100 x 4.6 mm I.D.) from Phenomenex (CA, USA). The elution was performed using solvent A (formic acid 0.1% (v/v) aqueous solution) and solvent B (acetonitrile) as mobile phase components. The initial conditions for the mobile phase were 3% of solvent B and 97% of solvent A. After 5 min, the acetonitrile contribution was increased to 95% until minute 13, and maintained at this percentage until minute 15. Then, a return to the initial conditions was programmed during 30 seconds to ensure the column re-equilibration, and isocratic conditions were maintained until the end of the programme. The elution gradient is summarized in Table 5.2. The elution programme lasted 20 min for each injection, and it had a flow rate of 0.4 $mL \cdot min^{-1}$. Besides, the injection volume was of 5 µL for each vial.

Table 5.2. Elution gradient used for HPLC-UV separation

Time [min]	**Solvent B** [%]	**Elution mode**
0-5	3	Isocratic
5-13	3-95	Lineal
13-15	95	Isocratic
15-15.50	95-3	Lineal
15.50-20	3	Isocratic

HPLC-UV fingerprints were registered at 250 nm, 280 nm, 310 nm, 370 nm and 550 nm absorbance wavelengths. Regardless of that, the only wavelength used for the posterior chemometric analysis was 280 nm because of the number of peaks observed.

5.4. DATA ANALYSIS

The HPLC-UV analysis of all the samples resulted in a fingerprint for each wavelength. The fingerprints at 280 nm wavelength had richer information, so they were the chosen ones to develop the analysis. Once the raw data was exported to a spreadsheet using Unichrom from New Analytical Systems (Minsk, Belarus), the needed matrices were built. At this point, chemometric calculations could be performed through a stand-alone chemometrics software (SOLO) from Eigenvector Research (WA, USA).

Before building the model for the chemometric analyses, the matrices with the data were pre-treated to obtain reliable results. The pre-treatment consisted of smoothing, base-line correction, aligning and autoscaling.

The X-data matrix employed for PCA consisted of the HPLC-UV fingerprints at 280 nm (absorbance signal vs time) of each sample and QCs. On the other hand, PLS-DA uses the previously named X-matrix and a second one, called Y-matrix, which distinguishes the studied samples into predefined classes (honey botanical variety or honey geographical origin). The obtained score plot of PLS-DA can reflect the similarities between samples of the same class by the formation of clusters or separated sample groups. The number of LVs used to build the model was stablished by the first relevant minimum of the cross-validation (CV) error from the Venetian blind approach when the matrix included more than 20 samples, and the Leave one out approach for studies with less than 20 samples. A hierarchical cluster analysis (HCA) was also used and

the Ward's method was employed for the chemometric calculations, which is an agglomerative method that groups the samples included in the X-matrix into clusters considering the distance between them, creating bigger clusters each time. In this case, the X-matrix contained the HPLC-UV fingerprints at 280 nm of each sample without QCs.

6. RESULTS AND DISCUSSION

6.1. HPLC-UV FINGERPRINTS

As previously mentioned, all the honey samples were analysed with the proposed non-targeted HPLC-UV method and, as a result, chromatographic fingerprints registered at 280 nm were obtained. Figure 6.1 shows the corresponding chromatographic fingerprints of the selected honey samples for each botanical variety.

At first sight, important differences and similarities between the different botanical varieties that are under study can be observed. These differences are related to both the number and the intensity of the detected signals. Three important fingerprint segments can be observed in all the obtained fingerprints: from (i) 2-7 min, (ii) 8-10 min, and (iii) 11-15 min. As can be seen, blossom and rosemary honeys show similar fingerprinting profiles, with more or less the same peaks and signal intensities. The same behaviour is observed for thyme and eucalyptus honeys. On the other hand, holm oak, forest and mountain fingerprints also look alike but they have higher peak intensities in the first and third fingerprint segments when compared with the other previously commented honey varieties. The heather honey fingerprint shows a similar profile to holm oak, forest and mountain fingerprints, but also depicts a relatively important peak signal in the second fingerprint segment, although less intense in comparison to blossom, rosemary, thyme and eucalyptus samples. Lastly, multifloral honeys show very rich fingerprints regarding both peak signals and intensities in comparison to the monofloral varieties, as expected.

It should be mentioned that, in general, the obtained chromatographic fingerprints within a specific floral variety are quite similar, although with small differences that may be attributed to the composition of the other floral varieties involved in the final honey (as commented in the introduction section, monofloral honeys are named according to the most abundant floral variety if its content is higher than 45%), or to the geographical origin. Thus, the obtained fingerprints were used as honey chemical descriptors to characterize and classify the analysed honey samples with exploratory and classificatory chemometric methods.

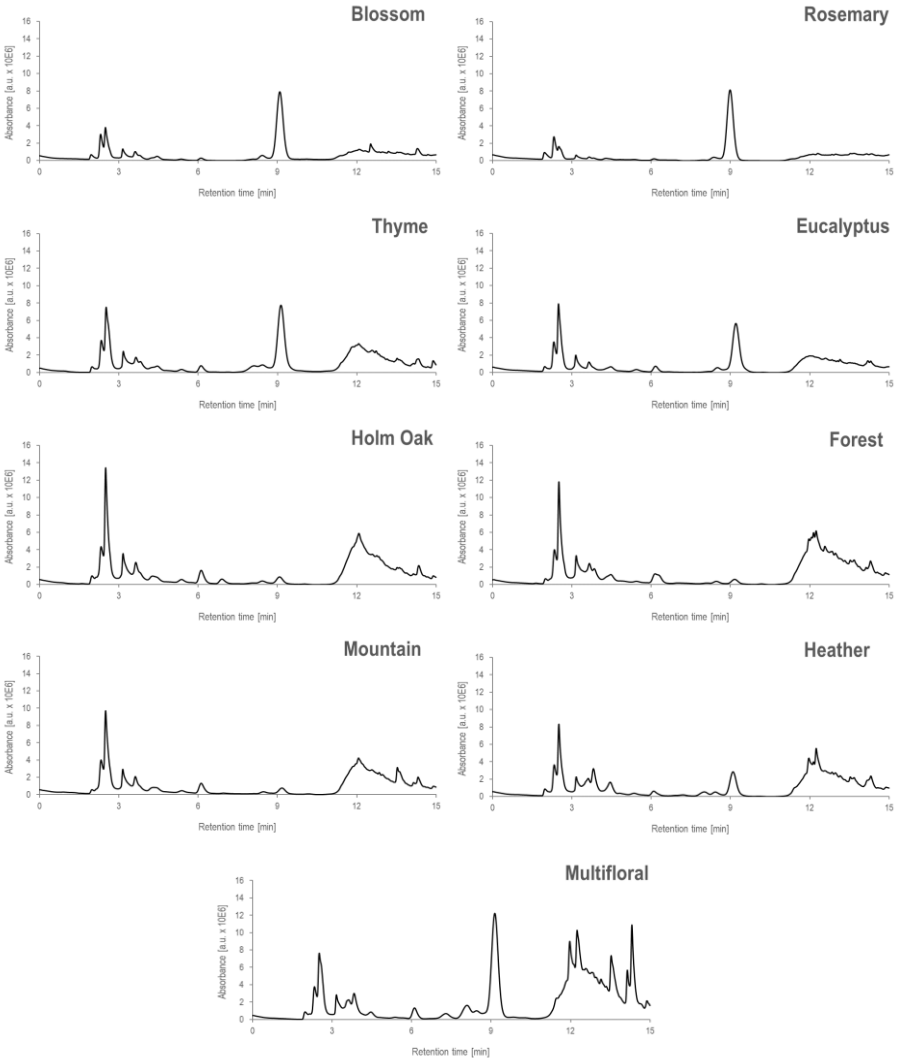

Figure 6.1. HPLC-UV chromatographic fingerprints at 280 nm of a selected honey sample for each floral variety.

6.2. CHARACTERIZATION AND CLASSIFICATION OF HONEY SAMPLES

Once the HPLC-UV fingerprints of all the honey samples were obtained with the proposed method, the robustness and reproducibility of the method, and the characterization and classification of the analysed honey samples when using the fingerprints as sample chemical descriptors were assessed by PCA, PLS-DA and HCA. PCA was employed to evaluate the reproducibility and robustness of the method by the depiction of all the QC injections in the score plot. This chemometric analysis can also be helpful to detect separation between different groups of samples. On the other hand, PLS-DA and HCA were performed for classificatory purposes. Classification of samples were evaluated taking into account their botanical origin and their production region.

6.2.1. Characterization and classification of honey by botanical origin

Firstly, with all the collected data from the HPLC-UV fingerprints, a non-supervised PCA study was conducted and the obtained score plot of PC1 vs PC2 is depicted in Figure 6.2. First, reproducibility of the proposed HPLC-UV method needs to be assessed by studying the QCs distribution. As can be seen in the score plot, all QCs are clustered in the middle of the plot, verifying that the method is both reproducible and robust, and that no instrumental variations could lead to unreliable results.

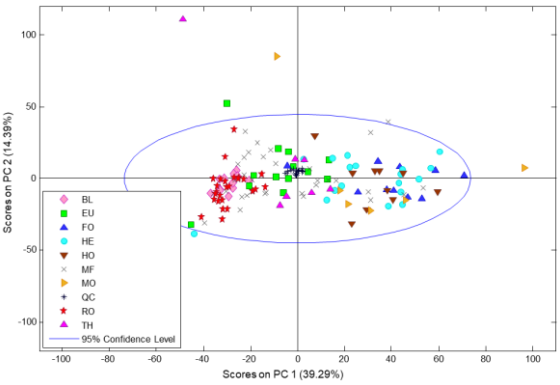

Figure 6.2. PCA score plot of PC1 vs PC2 for honey samples classified by botanical origin when using HPLC-UV fingerprints at 280 nm as sample chemical descriptors. A total of 8 PCs were used to build the model.

Furthermore, it can be observed that the analysed samples tend to be distributed and more or less grouped in different zones of the plot depending on their botanical origin, and their distribution is strongly influenced by PC1. For example, it can be easily seen that blossom (BL) and rosemary (RO) honey varieties are well grouped in the left part of the plot (negative PC1 values). In contrast, mountain (MO), holm oak (HO), heather (HE) and forest (FO) honey varieties are more widely distributed but in the right area of the score plot (presenting positive PC1 values). In the middle area, thyme (TH) and eucalyptus (EU) honey varieties are found, even though the last one seems to be more dispersed. Finally, it can be observed that honeys labelled as multifloral (MF) are spread along the plot which can be justified by the great number of plants in different proportions that a honey of this variety can present.

In order to simplify the obtained results, a PCA model was rebuilt without considering MF honey samples, and the obtained score plot (PC1 vs PC2) is shown in Figure 6.3. As can be seen, the distribution of each honey variety group in the plot remains more or less the same, with three clear sample groups separated along the horizontal axis represented by PC1: from left to right, group 1 (BL and RO), group 2 (TH and EU), and group 3 (MO, HO, HE, and FO).

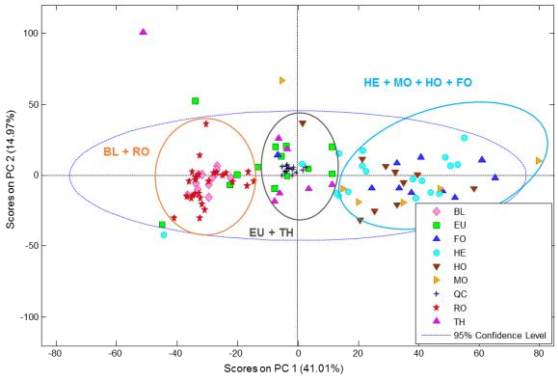

Figure 6.3. PCA score plot of PC1 vs PC2 for honey samples (except multifloral honeys) classified by botanical origin when using HPLC-UV fingerprints at 280 nm as sample chemical descriptors. A total of 4 PCs were used to build the model.

After the non-supervised PCA study, the obtained HPLC-UV fingerprints were also submitted to a supervised classificatory PLS-DA chemometric method to classify the samples according to their botanical origin. In this case, Figure 6.4a shows the obtained PLS-DA score plot of LV1 vs

LV2. As can be seen, a clear separation and discrimination between some varieties is again observed, although a different sample distribution than the one observed by PCA is obtained. Now, in the left area of the plot, MO, HO, HE and FO honey varieties can be found. The TH honeys are located in the centre of the plot while EU samples can be found slightly displaced to the centre-right area. In contrast, in the right half of the plot we can find BL and RO honey samples, although it seems that BL honeys tend to be located at higher values of LV2 in comparison to RO honeys. Again, MF honey samples are distributed along the whole plot as expected due to their more complex floral composition. It is important to mention that with PLS-DA, not only LV1 is clearly describing the distribution of the samples.

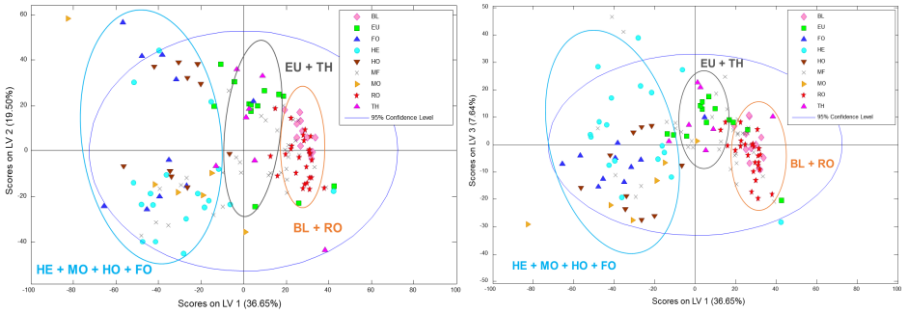

Figure 6.4. PLS-DA score plot of (a) LV1 vs LV2 and (b) LV1 vs LV3 for the classification honey samples according to botanical origin when using HPLC-UV fingerprints at 280 nm as sample chemical descriptors. A total of 3 LVs were used to build the model.

During the analysis of PLS-DA plots, it was noticed that representing other latent variables benefited the grouping of several varieties. Figure 6.4b shows the PLS-DA score plot of LV1 vs LV3 and, comparing both plots, it can be observed that LV3 helped in the grouping of some samples such as FO and HO honeys, which are displaced down the plot, and TH and EU honeys, that shifted to the upper half area. The fact that the named varieties are now better assembled is also remarkable. In both representations, MF samples do not separate from the rest nor form a compact group regardless of the chosen latent variables.

If MF honey samples are removed from the study again, the sample distribution does not change significantly, as can be seen from the obtained PLS-DA score plot of LV1 vs LV2 depicted in Figure 6.5a. Moving through the plot from left to right, varieties appear more or less in the same

order, but it cannot be observed a discrimination of varieties along the vertical axis which corresponds to LV2.

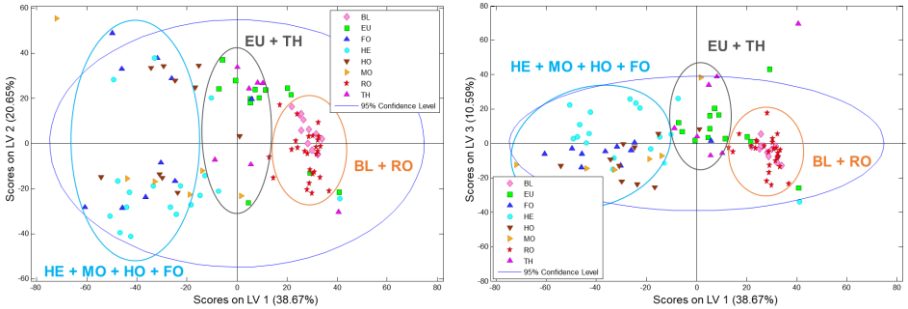

Figure 6.5. PLS-DA score plot of (a) LV1 vs LV2 and (b) LV1 vs LV3 for the classification of honey samples (except multifloral honeys) when using HPLC-UV fingerprints at 280 nm as sample chemical descriptors. A total of 3 LVs were used to build the model.

Similarly, other plots were studied, and the PLS-DA of LV1 vs LV3 (Figure 6.5b) was able to distinguish between samples along the upper and lower part of the plot. In addition, an improvement of the sample grouping was again achieved. Irrespective the PCA and PLS-DA study, it can be observed that the analysed honey samples tend to be distributed within three main groups showing similarities within the obtained HPLC-UV fingerprints: a group including BL and RO samples; a group including TH and EU samples; and a bigger group including MO, HO, HE, and FO samples.

Although a non-targeted HPLC-UV fingerprinting method was employed, it can be assumed that this differentiation is due to the chemicals present in these fingerprints. In fact, it has been reported in the literature that honeydew honeys (which come from plant secretions) have different polyphenol content than blossom honeys (which come from flower nectar) [8]. Honeydew honeys include holm oak, pine and forest honeys and they have more phenolic acids than blossom honeys, but less flavonoids. Besides, the total phenolic content (TPC) is higher in honeydew honeys. This can justify the fact that HO and FO varieties appear grouped in the plots. Despite that HE honey is considered a blossom honey, the literature remarks some similarities of this variety with honeydew honeys, as for example their dark colour or the TPC, which is particularly high. This could also explain the position of HE honeys in the plot grouped with the honeydew

honeys. In addition, the studied MO honeys are mostly made from holm oak, pine, oak, heather and chestnut. Considering that the chestnut variety is a blossom honey with the same behaviour as the heather ones, it could be reasonable that MO samples are also grouped with HO, FO and HE varieties. Finally, EU and TH have a similar flavonoids composition which can be the reason why they appear grouped in the plot [18].

To endorse the obtained classification of honeys, a hierarchical cluster analysis (HCA) was used. The resulting dendrogram in shown in Figure 6.6.

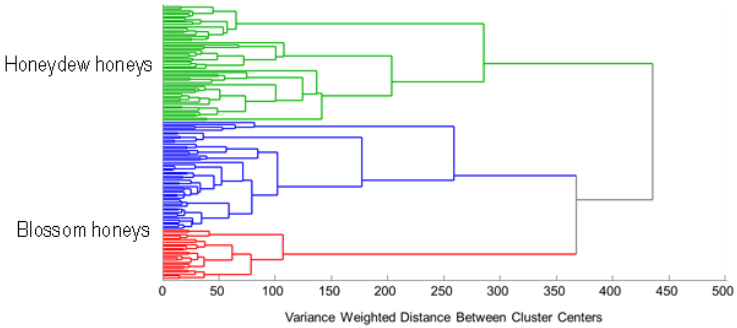

Figure 6.6. Dendrogram for all the honey samples using Ward's method.

A first division of the honey samples into two big groups can be observed, which correspond to blossom honeys (blue and red) and honeydew honeys (green). In the green coloured samples, heather and mountain honeys have been also detected, but their presence with the honeydew honeys has been justified before. Blossom honeys are also separated into two parts: the red one, which includes mainly eucalyptus and thyme honey varieties, and the blue one, which is composed mainly with blossom and rosemary honey varieties. These results support the sample discrimination previously accomplished by PCA and PLS-DA chemometric methods.

Further studies of the polyphenols content of each sample will be needed to confirm this explanation.

6.2.2. Characterization and classification of honey varieties within each production region

Another way to approach the honey analysis is to look for the differences between varieties considering honeys from a single region. In this occasion, more limited PLS-DA chemometric

methods are presented because there was not a considerable number of available samples for all the studied production regions. The studied national regions were Aragon (AR), Asturias (AS), Cantabria (CN), Castile La Mancha (CM), Castile and Leon (CL), Catalonia (CT), Extremadura (E), Balearic Islands (BI), Navarre (N), and Spain (S) (no other information was available). Some examples of regions that revealed interesting tendencies are shown.

For example, the PLS-DA score plot of LV1 vs LV2 obtained for honey samples produced in Balearic Islands is depicted in Figure 6.7. As can be seen, despite the low number of samples available, a good discrimination of the analysed samples according to their botanical origin was obtained. As usual, MF does not form any clustered group and they are distributed within the plot. As expected, RO and BL samples tend to be closely grouped (although more or less separated) because of their compositional similarities [8].

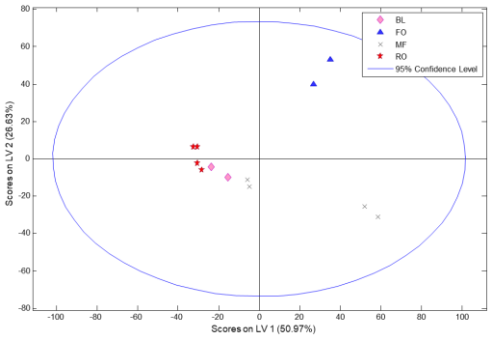

Figure 6.7. PLS-DA score plot of LV1 vs LV2 for honey samples produced in Balearic Islands when using HPLC-UV fingerprints at 280 nm as sample chemical descriptors for the classification of samples according to their botanical origin. A total of 4 LVs were used to build the model.

The PLS-DA score plot of LV1 vs LV2 for the analysed honey samples produced in Extremadura is shown in Figure 6.8. A similar sample distribution to the ones observed from previous models was observed. BL and RO varieties keep appearing closely grouped in the right part of the plot, EU and TH varieties in the middle area, and HE and HO varieties in the left part, although a certain discrimination between these two honeys is accomplished being more or less separated by LV2 (HE at the top and HO at the bottom of the plot). MF honeys exhibit the same behaviour as previously commented, being distributed through all the score plot.

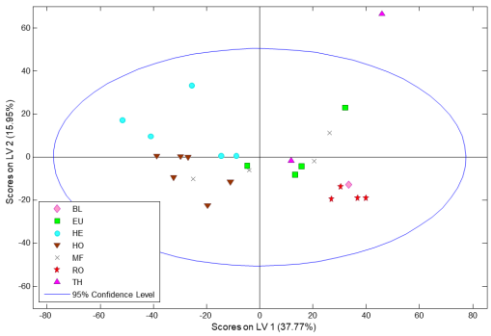

Figure 6.8. PLS-DA score plot of LV1 vs LV2 for honey samples produced in Extremadura when using HPLC-UV fingerprints at 280 nm as sample chemical descriptors for their classification according to their botanical origin. A total of 7 LVs were used to build the model.

Interesting results were also obtained from the honey samples produced in Catalonia (see PLS-DA score plot of LV1 vs LV2 in Figure 6.9), where three differentiated zones can be observed: on one side BL and RO; on the other side HE, MO and FO; and TH and EU in the central part. MF variety does not change its behaviour from the previously commented results.

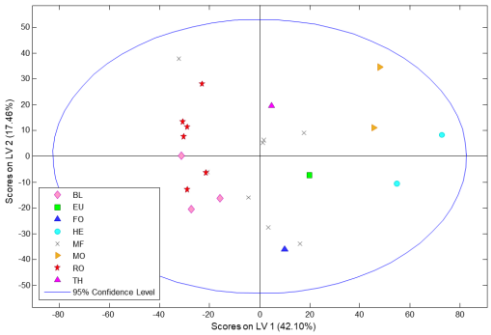

Figure 6.9. PLS-DA score plot of LV1 vs LV2 for honey samples produced in Catalonia when using HPLC-UV fingerprints at 280 nm are used as sample chemical descriptors for the classification according to the botanical origin. A total of 6 LVs were used to build the model.

Lastly, PLS-DA was also applied to the honey samples produced in the region of Castile and Leon, and the obtained results are shown in Figure 6.10. The three distinguished groups that

have repeatedly appeared continue to be identifiable in this model. However, in this case, EU samples shifted to the right area of the plot with BL and RO honey samples.

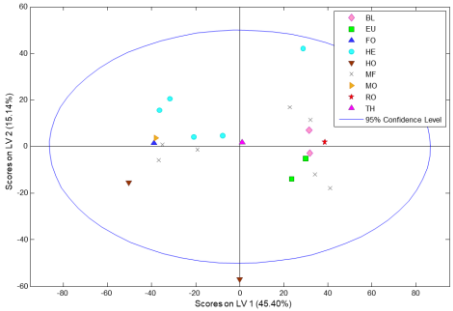

Figure 6.10. PLS-DA score plot of LV1 vs LV2 for honey samples produced in Castile and Leon when using HPLC-UV fingerprints at 280 nm as sample chemical descriptors for their classification according to the botanical origin. A total of 2 LVs were used to build the model.

Up to this point, it can be concluded that the obtained non-targeted HPLC-UV fingerprints at 280 nm are good sample chemical descriptors to address a certain classification of honey samples according to their botanical origin, independently of their production region. At least three main honey-groups can be differentiated, which are clearly related to their chemical composition on bioactive substances, but without the requirement of targeting any of these chemical compounds.

6.2.3. Characterization and classification of honey samples by the geographical production region

The obtained HPLC-UV fingerprints were also evaluated as sample chemical descriptors to address the classification of honey samples according to their geographical production region considering just national honeys. For that purpose, the non-supervised PCA was also studied from the point of view of the geographical production region of the analysed honey samples, and the obtained PCA score plot of PC1 vs PC2 is shown in Figure 6.11. This plot is exactly the same than the one depicted in Figure 6.1, but labelling the samples by production region instead of botanical origin.

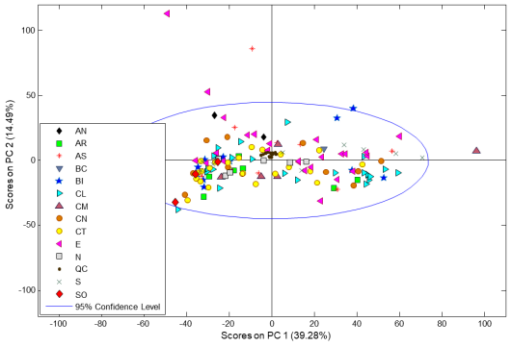

Figure 6.11. PCA score plot of PC1 vs PC2 for honey samples classified by geographical production region when using HPLC-UV fingerprints at 280 nm as sample chemical descriptors. A total of 8 PCs were used to build the model.

In contrast to the results obtained by botanical origin, it was not possible to make a clear distinction between classes of samples (no clear grouping can be detected). In addition to that, none of the principal component combinations could enhance the discrimination of honey samples according to their production region.

Supervised classificatory PLS-DA was also evaluated, and the score plot of LV1 vs LV2 is depicted in Figure 6.12. Similarly to PCA, samples are still not gathering in clusters or forming separated groups according to each production region, confirming that these non-targeted HPLC-UV fingerprints are not good enough to classify samples based on the production region.

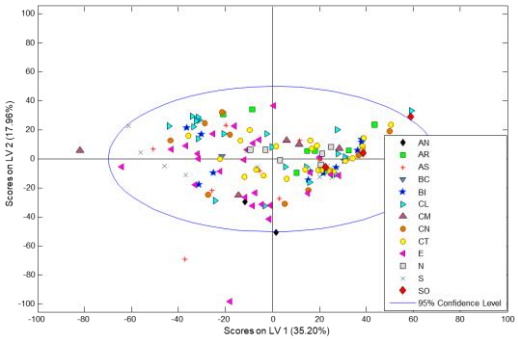

Figure 6.12. PLS-DA score plot of LV1 vs LV2 for honey samples classified by geographical production region when using HPLC-UV fingerprints at 280 nm as sample chemical descriptors. A total of 4 LVs were used to build the model.

7. CONCLUSIONS

In this project a non-targeted HPLC-UV fingerprinting method has been used for the characterization and classification of 137 honeys, with a previous simple sample treatment. The classification has been studied based on the botanical and geographical origin separately. The obtained fingerprints were employed as sample chemical descriptors and submitted to chemometric methods, and the results can be summarized in the next conclusions:

- The HPLC-UV fingerprinting method was reproducible and robust, as demonstrated by the good clustering of QCs when performing PCA.
- Important visual differences in the obtained HPLC-UV fingerprints of the analysed honey samples were observed among the different botanical origins. Thus, HPLC-UV fingerprints may be considered good chemical descriptors to address honey characterization and classification.
- Classification of honey samples from different locations according to their botanical origin was achieved to a certain level by PLS-DA. It was possible to distinguish between blossom honeys and honeydew honeys. Honeydew honeys clusters included also some blossom honeys such as heather and mountain honeys, probably due to their similarity in polyphenolic content. Similar results were obtained when PLS-DA was performed by studying honey samples from the same national region of production.
- PLS-DA results also showed a separation between blossom honeys varieties. One group contained eucalyptus and thyme honeys, and the other one had blossom and rosemary honeys.
- The clustering analysis performed with HCA supports the classification observed with the PLS-DA models.
- The proposed non-targeted HPLC-UV fingerprinting methodology was not able to distinguish honeys according to their region of production. A targeted polyphenolic HPLC-UV method will be evaluated in the future with this purpose.

8. REFERENCES

1. Abbas, O.; Zadravec, M.; Baeten, V.; Mikuš, T.; Lešić, T.; Vulić, A.; Prpić, J.; Jemeršić, L.; Pleadin, J. Analytical methods used for the authentication of food of animal origin. *Food Chem.* **2018**, *246*, 6–17, doi:10.1016/j.foodchem.2017.11.007.
2. Soares, S.; Amaral, J.S.; Oliveira, M.B.P.P.; Mafra, I. A Comprehensive Review on the Main Honey Authentication Issues: Production and Origin. *Compr. Rev. Food Sci. Food Saf.* **2017**, *16*, 1072–1100, doi:10.1111/1541-4337.12278.
3. Chin, N.L.; Sowndhararajan, K. A Review on Analytical Methods for Honey Classification, Identification and Authentication. *Intechopen.* **2020**, doi:10.5772/intechopen.90232.
4. Naila, A.; Flint, S.H.; Sulaiman, A.Z.; Ajit, A.; Weeds, Z. Classical and novel approaches to the analysis of honey and detection of adulterants. *Food Control.* **2018**, *90*, 152–165, doi:10.1016/j.foodcont.2018.02.027.
5. Rachele R. At a glance: The EU's beekeeping sector. European Parliamentary Research Service. **2017**, PE 608.786
6. Parlamento Europeo. Perspectivas y desafíos para el sector apícola de la Unión. **2018**, P8_TA(2018)0057.
7. Jones, R. Prologue: Honey and healing through the ages. *J. ApiProduct ApiMedical Sci.* **2009**, *1*, 2–5, doi:10.3896/ibra.4.01.1.02.
8. Pita-Calvo, C.; Vázquez, M. Differences between honeydew and blossom honeys: A review. *Trends Food Sci. Technol.* **2017**, *59*, 79–87, doi:10.1016/j.tifs.2016.11.015.
9. Food and Agriculture Organization of the United Nations; Word Health Organization. Codex Alimentarius International Food Standards. Standard for honey. **2019**, CXS12-1981.
10. Alvarez-Suarez, J.M.; Tulipani, S.; Romandini, S.; Bertoli, E.; Battino, M. Contribution of honey in nutrition and human health: A review. *Med. J. Nutrition Metab.* **2010**, *3*, 15–23, doi:10.1007/s12349-009-0051-6.
11. Gašić, U.M.; Milojković-Opsenica, D.M.; Tešić, Ž.L. Polyphenols as possible markers of botanical origin of honey. *J. AOAC Int.* **2017**, *100*, 852–861, doi:10.5740/jaoacint.17-0144.
12. Núñez, O.; Lucci, P. *Application of Liquid Chromatography in Food Analysis*; Núñez, O., Lucci, P., Eds.; MDPI: Basel, Switzerland, **2020**; ISBN 978-3-03943-362-9.
13. Esteki, M.; Shahsavari, Z.; Simal-Gandara, J. Food identification by high performance liquid chromatography fingerprinting and mathematical processing. *Food Res. Int.* **2019**,

122, 303–317, doi:10.1016/j.foodres.2019.04.025.

14. Ballin, N.Z.; Laursen, K.H. To target or not to target? Definitions and nomenclature for targeted versus non-targeted analytical food authentication. *Trends Food Sci. Technol.* **2019**, *86*, 537–543, doi:10.1016/j.tifs.2018.09.025.
15. Esslinger, S.; Riedl, J.; Fauhl-Hassek, C. Potential and limitations of non-targeted fingerprinting for authentication of food in official control. *Food Res. Int.* **2014**, *60*, 189–204, doi:10.1016/j.foodres.2013.10.015.
16. Granato, D.; Putnik, P.; Kovačević, D.B.; Santos, J.S.; Calado, V.; Rocha, R.S.; Cruz, A.G. Da; Jarvis, B.; Rodionova, O.Y.; Pomerantsev, A. Trends in Chemometrics: Food Authentication, Microbiology, and Effects of Processing. *Compr. Rev. Food Sci. Food Saf.* **2018**, *17*, 663–677, doi:10.1111/1541-4337.12341.
17. Kumar, N.; Bansal, A.; Sarma, G.S.; Rawal, R.K. Chemometrics tools used in analytical chemistry: An overview. *Talanta.* **2014**, *23*,186–199,doi:10.1016/j.talanta.2014.02.003.
18. Martos, I.; Cossentini, M.; Ferreres, F.; Tomá S-Barberán, F.A. Flavonoid Composition of Tunisian Honeys and Propolis. *J. Agric. Food Chem.* **1997,** *45,2824−2829*, doi:10.1021/jf9609284.

12. ACRONYMS

AR: Aragon

AS: Asturias

BC: Basque Country

BI: Balearic Islands

BL: Blossom

CL: Castile and Leon

CM: Castile La Mancha

CN: Cantabria

CT: Catalonia

CV: Cross-Validation

E: Extremadura

EU: Eucalyptus

F: France

FO: Forest

HCA: Hierarchical cluster analysis

HE: Heather

HO: Holm oak

HPLC-UV: High-Performance Liquid Chromatography with Ultraviolet Detection

I.D.: Internal diameter

LV: Latent variable

MF: Multifloral

MO: Mountain

N: Navarre

PC: Principal component

PCA: Principal component analysis

PLS-DA: Partial least squares-discriminant analysis

QC: Quality control

RO: Rosemary

S: Spain

SO: Spain and others

TH: Thyme

UHPLC: Ultra High-Performance Liquid Chromatography