Food Bioscience 50 (2022) 102092

Contents lists available at ScienceDirect

Food Bioscience

journal homepage: www.elsevier.com/locate/fbio

Differentiation of bee pollen samples according to the apiary of origin and harvesting period based on their amino acid content

Ana M. Ares [a], Laura Toribio [a], Jesús A. Tapia [a,b], Amelia V. González-Porto [c], Mariano Higes [c], Raquel Martín-Hernández [c,d], José Bernal [a,*]

[a] I. U. CINQUIMA, Analytical Chemistry Group (TESEA), Faculty of Sciences, University of Valladolid, Valladolid, Spain
[b] Department of Statistics and Operations Research, Faculty of Sciences, University of Valladolid, Valladolid, Spain
[c] Instituto Regional de Investigación y Desarrollo Agroalimentario y Forestal de Castilla La Mancha (IRIAF), Centro de Investigación Apícola y Agroambiental (CIAPA) de Marchamalo, Camino de San Martín, s/n, 19180 Marchamalo, Guadalajara, Spain
[d] Instituto de Recursos Humanos para la Ciencia y la Tecnología (INCRECYT-FEDER), Fundación Parque Científico y Tecnológico de Castilla – La Mancha, Guadalajara, Spain

ARTICLE INFO

Keywords:
Bee pollen
Biomarkers
Canonical discriminant analysis
Food authenticity
Free amino acids
HPLC-FLD

ABSTRACT

Bee pollen is currently one of the most widely consumed dietary supplements due to its high nutritional value and its potentially beneficial effects on health. Unfortunately, in recent years an increase in the fraudulent marketing of this product has been detected, mainly in terms of adulteration with pollen from other sources. This has made it necessary to seek new tools to ensure its authentication. Therefore, this study investigates the use of free amino acids as markers of the geographical origin and harvesting period of bee pollen. To demonstrate their potential as biomarkers, 72 samples from four apiaries (Pistacho, Tío Natalio, Monte and Fuentelahiguera), located in the same geographical area (Marchamalo, Guadalajara, Spain), were analyzed by liquid chromatography-fluorescence detection, with the data obtained undergoing canonical discriminant analysis. Variable amounts and numbers of free amino acids were found in the samples analyzed; proline predominated in all of them, in a concentration range of 298–569989 mg/kg. The differences observed in amino acid composition could be attributed to the flowering plants from which the bee pollen samples originated. In addition, it was possible to statistically assign over 75% of the samples to the corresponding apiary of origin, the best results being obtained for the Fuentelahiguera and Tío Natalio apiaries (100%); this classification was even superior in the case of the harvesting periods, as more than 90% of the samples were correctly assigned, and in one period (June) a 100% rate was obtained.

1. Introduction

The consumption of bee pollen has increased considerably in recent years due to both its beneficial effects on health and its nutritional properties (Conte et al., 2017). It is produced by the mixing of flower pollens with nectar (and/or honey) and salivary substances, which include bee enzymes such as catalase or amylase (Kaškonienė et al., 2020; Tutun et al., 2022). It has already been reported that bee pollen has antioxidant, anti-inflammatory, anti-cancer, anti-bacterial, anti--fungal, hepatoprotective and anti-atherosclerotic activities, as well as protecting the digestive and nervous systems (Araújo et al., 2017; Kaškonienė et al., 2020). These health promoting effects are directly related to the components of bee pollen, which comprise many different substances, including a high protein content, sugars, lipids, amino acids, lipid vitamins, minerals such as sodium, potassium and calcium, and organic and phenolic compounds (Ares et al., 2018; Kafantaris et al., 2021).

Both the composition and quality of bee pollen depend to a large extent on the type of plant and the geographical origin (Prđun et al., 2021), along with other factors such as climatic conditions and the type

Abbreviations: **AM**, April–May; **ANOVA**, analysis of variance; **CDA**, canonical discriminant analysis; **CIAPA**, *Centro de Investigación Apícola y Agroambiental*; **FH**, *Fuentelahiguera*; **FMOC-Cl**, 9-fluorenylmethyl chloroformate; **JA**, July–August; **JN**, June; **LC-FLD**, liquid chromatography with fluorescence detection; **LOD**, limit of detection; **LOQ**, limit of quantification; **MO**, *Monte*; **OPA**, *o*-phthalaldehyde; **PI**, *Pistacho*; **TN**, *Tío Natalio*.

* Corresponding author.
E-mail address: jose.bernal@uva.es (J. Bernal).
URL: http://tesea.uva.es (J. Bernal).

https://doi.org/10.1016/j.fbio.2022.102092
Received 11 May 2022; Received in revised form 22 September 2022; Accepted 2 October 2022
Available online 12 October 2022

of soil (Gardana et al., 2018). Thus, the different source-dependent composition of bee pollen can be very useful for combatting the fraudulent practice that is currently affecting the beekeeping industry, as is adulteration with pollen from non-declared origins, (Wang et al., 2021, 2022). Consequently, authentication of bee products, especially honey and pollen, is essential to avoid illegal competition (Taha et al., 2019). Different strategies have been proposed to determine the origin of bee pollen, including morphological feature inspection, computer vision, and microscopic examination, all of which are directly related to palynological analysis, in addition to an investigation of their main/specific constituents (Chica & Campoy, 2012; Wang et al., 2021, 2022). The aim here is to evaluate not only their potential as quality indicators, but also as markers of the source of the pollen (Ares, Tapia, et al., 2022; Bleha et al., 2021; De-Melo et al., 2018; Gardana et al., 2018; Gonçalves et al., 2018; Isopescu et al., 2020; Milla et al., 2022; Restrepo et al., 2021; Taha et al., 2019; Themelis et al., 2019; Zeghoud et al., 2021). The main goal of most of these studies was to determine the botanical origin with the highest nutritional value or antioxidant effects in terms of its composition, scant attention being paid to investigating the geographical origin or detecting potential adulteration. Nevertheless, we have recently demonstrated the potential of glucosinolates as biomarkers of bee pollen origin, as it was possible to distinguish the apiary of origin and harvesting period on account of their specific content (Ares, Tapia, et al., 2022). Thus, we consider it appropriate from a scientific point of view to demonstrate whether other compounds, like amino acids, could also be employed to perform this task.

Amino acids, which are among the main components of bee pollen, have already been investigated in this substance (Al-Kahtani et al., 2020; Ares, Martín, et al., 2022; Bayram, 2021; Bayram et al., 2021; Gardana et al., 2018; Ghosh & Jung, 2022; Sommano et al., 2020; Stabler et al., 2018; Taha et al., 2019; Thakur & Nanda, 2018; Themelis et al., 2019) in order to characterize and determine botanical and/or geographical origins that may be of greater interest from a nutritional point of view. However, to the best of our knowledge, the specific potential of amino acids as markers of the apiary of origin has not yet been proven. These have generally been determined in bee pollen by liquid chromatography with fluorescence detection (LC-FLD), with a requisite derivatization step. In some cases, total amino acid content was studied; this implies not only free amino acids but also those bound to proteins, which were released by means of acid hydrolysis (Stabler et al., 2018; Taha et al., 2019; Themelis et al., 2019).

Therefore, the main aim of this paper is to investigate the potential of free amino acids as markers of geographical origins, by employing LC-FLD to determine their content in 72 bee pollen samples, from four different apiaries located within the same area (Marchamalo, Guadalajara, Spain). This represents the first ever study of whether bee pollen samples may be classified, by means of canonical discriminant analysis (CDA) of their free amino acid content, in terms of the apiary of origin and the harvesting period, which took place during three consecutive foraging periods in 2018 (April–May; June; July–August).

2. Materials and methods

2.1. Chemicals and standards

9-fluorenylmethyl chloroformate (FMOC-Cl), *o*-phthalaldehyde (OPA) and mercaptopropionic acid were supplied by Sigma–Aldrich Chemie Gbmh (Steinheim, Germany). LC grade methanol and acetonitrile were both obtained from Lab-Scan Ltd. (Dublin, Ireland), whilst sodium acetate was purchased from Carlo Erba (Barcelona, Spain). Boric acid and sodium hydroxide were provided by Merck (Darmstadt, Germany), and ultrapure water was obtained from Millipore Milli-RO plus and Milli-Q systems (Bedford, MA, USA). An Eppendorf Centrifuge 5810R (Hamburg, Germany), a Moulinette chopper device (Moulinex. Paris, France), IKA® Ultra-Turrax® T18 basic disperser (IKA®-Werke GmbH & Co. KG, Staufen, Germany), syringe filters (17 mm, Nylon 0.45 μm; Nalgene, Rochester, NY, USA), a drying oven and a vibromatic system (J.P. Selecta S.A., Barcelona, Spain) were used for the sample treatment.

Analytical standards of the investigated amino acids (see Table 1) and the internal standard (IS), γ-aminobutyric acid (GABA), were purchased from Sigma Aldrich (St. Louis, MO, USA). Standard stock solution was prepared by dissolving approximately 10 mg of the accurately weighed compounds in 10 mL of HCl (0.1 mol/L), and a final concentration of approximately 1000 mg/L was obtained. Standard in solvent calibration curves were employed for performing the quantification within the range of LOQ (see Table 1) to 30 mg/L (LOQ, 1.0, 5.0, 10.0, 15.0, 20.0, 25.0 and 30 mg/L). Stock, working, and calibration solutions were stored in glass containers in darkness at 4 °C. All solutions remained stable for over two weeks.

2.2. Samples

2.2.1. Sampling

Bee pollen samples were obtained from four experimental apiaries with homogeneous colonies of *Apis mellifera iberiensis* (n = 72). Three of the apiaries, Pistacho (PI), Tío Natalio (TN) and Monte (MO), were located on the Centro de Investigación Apícola y Agroambiental (CIAPA) farm in Marchamalo, and the fourth, Fuentelahiguera (FH), was in the municipality of Fuentelahiguera de Albatages; all of them were in the province of Guadalajara (Spain; see Supplementary Material, Fig. 1S).

The PI apiary consisted of twenty colonies arranged on a plot in which the plant species were for spring and autumn cultivation. The larger number of colonies was determined by several influential factors regarding diversity, such as the different variety of crops present at each period of the year. The hives were distributed in two rows in a central lane facing the crops to facilitate the bees' access to these. The crops were selected for several reasons; they are highly attractive to bees, possess a potentially high nutritional value for these insects, and have significant economic importance in our country. Flowering, which began in April, was staggered depending on the plant species. In each of the other three apiaries (TN, FH and MO), five colonies were selected for the study. These apiaries are in an area primarily occupied by wild vegetation, although their proximity to crops, especially in TN, does not exceed 3 km. The number of colonies in these apiaries was only five, since this has traditionally been considered sufficient, and because the only factor determining diversity was the flowering in succession of the plants surrounding the apiaries.

Bee pollen samples were collected using pollen traps placed at the entrance of the hive. Every two weeks, the pollen trap grid was closed for a period of 24 h in the different hives. The pollen stored in the collection

Table 1
- Sensitivity data in LC-FLD analysis of free amino acids in bee pollen.

Amino acid	LOQ (mg/L)	LOQ (mg/kg)
Aspartic acid (ASP)	0.47	374
Glutamic acid (GLU)	0.18	143
Asparagine (ASN)	0.21	171
Serine (SER)	0.10	81
Glutamine (GLN)	0.44	355
Histidine (HIS)	0.48	380
Glycine (GLY)	0.12	93
Threonine (THR)	0.30	239
Arginine (ARG)	0.36	289
Alanine (ALA)	0.14	117
Tyrosine (TYR)	0.12	96
Valine (VAL)	0.95	68
Methionine (MET)	0.09	131
Tryptophan (TRP)	0.16	144
Phenylalanine (PHE)	0.18	126
Isoleucine (ILE)	0.16	179
Leucine (LEU)	0.23	109
Lysine (LYS)	0.13	312
Proline (PRO)	0.02	15

drawer during this period was collected, and immediately taken to the laboratory, where it was frozen until palynological analysis. In this study, bee pollen samples were collected during three consecutive foraging periods in 2018; these are defined as initial (between April and May), intermediate (June) and final (between July and August).

2.2.2. Palynological analysis

This was performed as previously described (Ares, Tapia, et al., 2022; Cepero et al., 2014), with the difference that the pollen balls were not separated by color. In the present study a representative fraction of each bee pollen sample (≈20 g) was subjected to palynological analysis. Briefly, pollen was extracted by diluting 0.5 g in 10 mL of water and then centrifuged at 2500 rpm for 15 min. Then, the sediment was poured onto the glycerin jelly slide after removal of the supernatant and examined microscopically to identify the pollen (× 250 magnification, Leitz Diaplan microscope, Leitz Messtechnik GmbH, Wetzlar, Germany). The species of plant was identified by means of a photographic atlas (Faegri & Iversen, 1989), together with the reference collection of pollen slides from CIAPA (Cepero et al., 2014).

The results of the contents of corbicular pollens mostly collected in the samples corresponding to each period and colony are summarized in Table 2. Whenever the majority taxon within the composition of the collected sample is well defined (greater than 80%; Campos et al., 2008), the name of that taxon has been applied in the corresponding column, while in those where this requirement is not met, the denomination MF (multifloral) has been applied; this is based on specialized literature (Escuredo et al., 2012). For the different foraging periods under analysis, namely, initial (April–May), intermediate (June), and final (July–August), several samples defined as MF were observed in the "major taxon identification" column; these can be differentiated in accordance with the flowering plants in that period.

Thus, in the first period, the MF samples corresponded to combinations of Brassica t., which includes different species of the same genus and some species of other genera, like wild and cultivated rapeseed (*Brassica napus* L.), wild radish (Raphanus t.) and wild rocket (Diplotaxis t.), with typical spring flowering plants such as almond trees, rockroses, and shepherd's sacks (Prunus t., Cistaceae, Cistus, Cistus ladanifer, Capsella t.). Meanwhile, the MF of the intermediate period differed between apiaries (MO, PI and TN). In MO, the combinations mainly comprised poppies, brambles, or broom (Papaver, Rubus, Retama t.), typical late spring or early summer plants in this geographical area. In PI, the combinations were much more diverse, and the number of taxa making up each combination was also greater. In this case, the taxa clearly corresponded to those with the highest flowering rate for the period and the location, such as Rosaceae, poppies or gorses (Rosaceae: Rosa t., Rubus, Prunus t.; Papaver and Genista t.). In TN, the total number of samples analyzed for each colony were MF combinations, comprising mainly poppies (Papaver); no predominant taxon was observed. For the final period (July–August), except for the hives of the FH apiary, where Rosa t. Was the principal taxon, MF comprised the rest; common taxa were sunflower, rose and bramble, (Helianthus, Rosa t., Rubus).

2.2.3. Sample treatment

Bee pollen samples were individually mixed, ground and pooled for optimum sample homogeneity. Next, bee pollen was dried until the mass stabilized (humidity was between 9% and 12%), and subsequently it was stored in the dark at − 20 °C until analysis. The proposed procedure, which is described in Fig. 1, was optimized in a previous study (Ares, Martín, et al., 2022). The resulting extract obtained after performing this procedure was filtered through a 0.45 μm nylon filter and transferred to a 2 mL vial, which was placed in the automatic injector for performing the derivatization procedure. OPA and FMOC-Cl were the derivatization reagents, and they were prepared as stated in Table 1S (see Supplementary Material). The online derivatization procedure is described in Table 2S (see Supplementary Material).

Table 2
- Major taxon and harvesting period data (April–May, AM; June, JN; July–August, JA) of the bee pollen samples from four different apiaries (Monte, MO; Pistacho, PI; Tío Natalio, TN; Fuentelahiguera, FH).

Sample	Major taxon	Harvesting period
MO-1	Brassica t. + MF	AM
MO-2	Brassica t.	AM
MO-3	Quercus	AM
MO-4	Brassica t.+ MF	AM
MO-5	Papaver + Rubus	JN
MO-6	MF	JN
MO-7	Papaver + MF	JN
MO-8	MF	JN
MO-9	Papaver + Retama t.	JN
MO-10	MF	JA
MO-11	Rosa t.	JA
MO-12	MF	JA
PI-1	Brassica t.	AM
PI-2	Brassica t.	AM
PI-3	Brassica t.	AM
PI-4	Brassica t.	AM
PI-5	Brassica t.	AM
PI-6	Brassica t.	AM
PI-7	Brassica t.	AM
PI-8	Brassica t.	AM
PI-9	Brassica t.	AM
PI-10	Quercus ilex t.	AM
PI-11	Brassica t.	AM
PI-12	Brassica t.	AM
PI-13	Brassica t.	AM
PI-14	MF	JN
PI-15	MF	JN
PI-16	Papaver + Rosa t.	JN
PI-17	Teucrium + Rosa t.	JN
PI-18	MF	JN
PI-19	MF	JN
PI-20	Rubus	JN
PI-21	Papaver + Rosaceae	JN
PI-22	Reseda + Retama t.	JN
PI-23	Reseda	JN
PI-24	Papaver + Retama t.	JN
PI-25	Cytisus t.	JN
PI-26	MF	JN
PI-27	MF	JN
PI-28	MF	JN
PI-29	Papaver + Rosa t.	JN
PI-30	MF	JN
PI-31	MF	JN
PI-32	MF	JA
PI-33	MF	JA
PI-34	MF	JA
PI-35	Rosa t.	JA
PI-36	MF	JA
PI-37	MF	JA
PI-38	Rosa t.	JA
PI-39	MF	JA
PI-40	Retama t.	JA
PI-41	MF	JA
PI-42	MF	JA
PI-43	MF	JA
PI-44	Rubus	JA
PI-45	MF	JA
TN-1	MF	JN
TN-2	MF	JN
TN-3	MF	JN
TN-4	MF	JN
TN-5	MF	JN
TN-6	MF	JA
TN-7	MF	JA
TN-8	Rosa t	JA
FH-1	Brassica t.	AM
FH-2	MF	AM
FH-3	Brassica t.	AM
FH-4	Vicia t.	JN
FH-5	Rosaceae	JN
FH-6	Rosa t.	JA
FH-7	Rosa t.	JA

MF, multifloral.

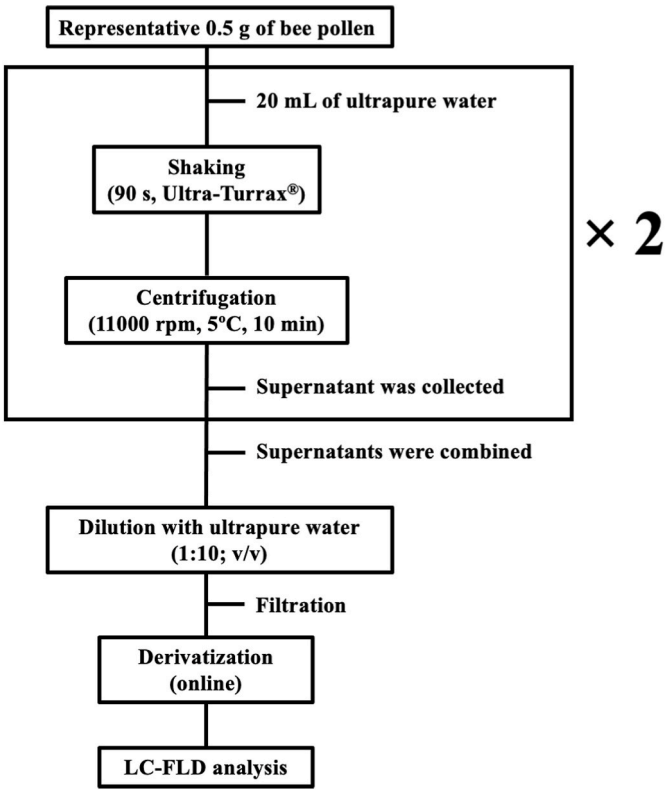

Fig. 1. Analytical procedure work-up flow chart.

2.3. LC-FLD conditions

An Agilent Technologies (Palo Alto, CA, USA) 1100 LC system equipped with a vacuum degasser, a quaternary solvent pump, an autosampler, a column oven, and a FLD was employed in this study. All were controlled by an Agilent ChemStation software. It must be remarked that the LC-FLD conditions were taken from a previous publication (Ares, Martín, et al., 2022). A Gemini® C_{18} (150 × 4.6 mm; 5-μm) analytical column and a C_{18} security guard cartridge (4 × 3.0 mm i. d.), both form Phenomenex (Torrance, CA, USA) were used for the analyses. The mobile phase was a mixture of sodium acetate (25 mmol/L at pH 8.0; solvent A), acetonitrile (solvent B), and methanol (solvent C), which was applied at a flow rate of 1.0 mL/min in the following gradient elution mode: i) 0 min (A-B-C, 100:0:0, v/v/v); ii) 2 min (A-B-C, 85:0:15, v/v/v); iii) 25 min (A-B-C, 67:18:15, v/v/v); iv) 32 min (A-B-C, 0:85:15, v/v/v); v) 35 min (A-B-C, 0:85:15, v/v/v); vi) A-B-C, 100:0:0, v/v/v). The injection volume and temperature were set at 1 μL and 40 °C, respectively. It was employed a detection program in which the OPA derivatives were detected at 240 nm (excitation) and 450 nm (emission) from 0 to 28.5 min, while FMOC derivatives were monitored from 28.5 min at 266 nm (excitation) and 313 nm (emission). A representative LC-FLD chromatogram of a bee pollen sample analyzed under the proposed conditions is shown in Fig. 2.

2.4. Statistical analysis

The normal distribution of data, analysis of variance (ANOVA) and the CDA required in this paper were studied by means of SAS PROC CANDISC (version 9.4; SAS Institute Inc., Cary, NC, USA). The CANDISC procedure carried out CDA, computed squared Mahalanobis distances between class means, and both univariate and multivariate one-way

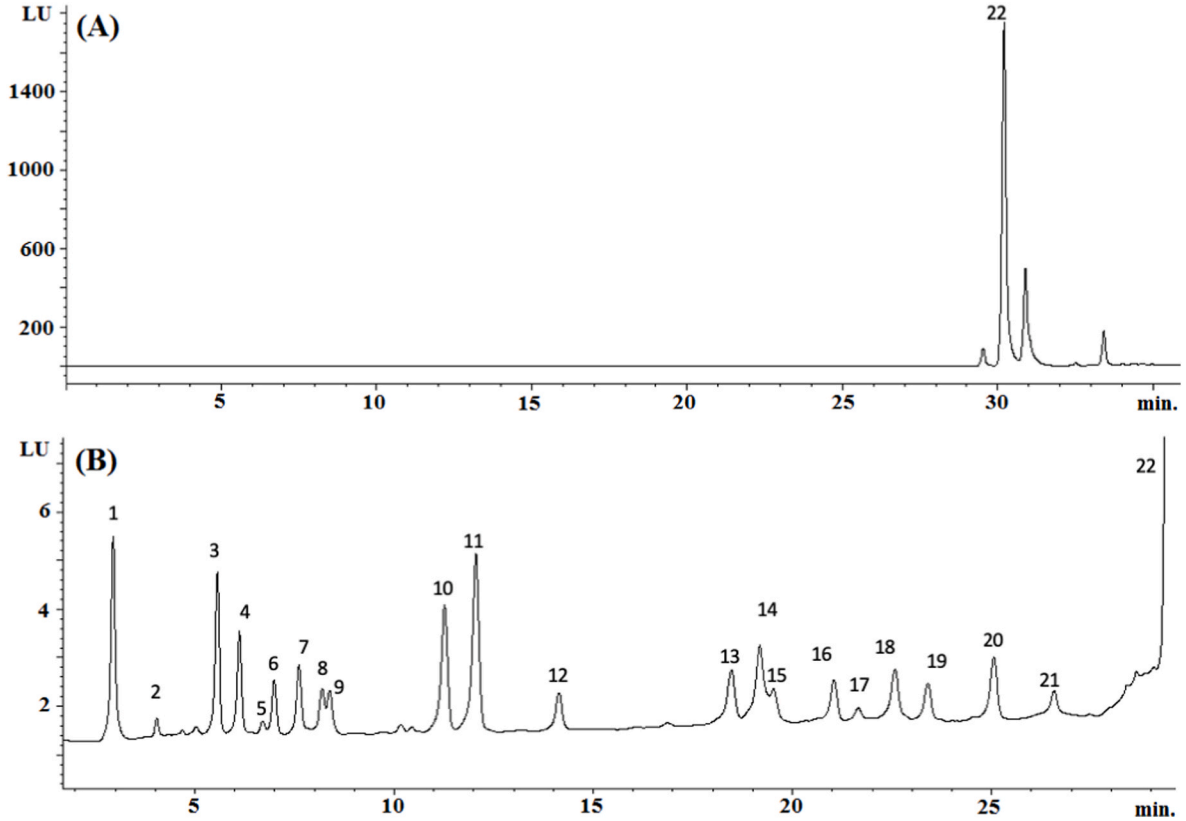

Fig. 2. Representative LC-FLD chromatograms obtained from multifloral bee pollen sample: (**A**) normal; (**B**) enlarged chromatogram without PRO (1. ASP; 2. GLU; 3. ASN; 4. SER; 5. GLN; 6. HIS; 7. GLY I; 8. THR; 9. ARG; 10. ALA; 11. GABA I; 12. TYR; 13. GLY II; 14. VAL; 15. MET; 16. GABA II; 17. TRIP; 18. PHE; 19. ILE; 20. LEU; 21. LYS; 22. PRO). The LC-FLD conditions are summarized in Subsection 2.4.

ANOVA. CDA obtains quadratic combinations of the factor variables that emphasize the differences among the groups (Ares, Tapia, et al., 2022; Jobson, 1991). The data base used in the present study comprised the response of each sample to the experimental groups (apiary of origin and harvesting period) and to the fifteen free amino acids measured. Thus, in two CDAs the discriminant factors were the apiary of origin and the harvesting period. In both of these, firstly the number of canonical functions needed to further explain the original variability of the data was determined; the scatterplot of the samples in the canonical functions and in each discriminant factor aptly showed the appropriacy of using the canonical functions selected for discrimination. Secondly, a quadratic discriminant function assuming normality tested equality of covariance matrixes (use within covariance matrix), computing squared Mahalanobis distances between class means, via cross-validation. Finally, a confusion matrix was obtained.

3. Results and discussion

3.1. Determination of the free amino acid content

The free amino acid content was determined in 72 samples of bee pollen from four apiaries located in Marchamalo (PI, n = 45; MO, n = 12; TN, n = 8; FH, n = 7). All the samples were analyzed in triplicate, and in certain cases it was necessary to dilute the extracts with ultrapure water (1:100; v/v), due to the high concentrations being outside the linear range. Interestingly, free amino acids were not detected in only one sample (TN-3). The results are listed in Tables 3 and 4 and **3S–8S** (see Supplementary Material), where the frequency (the number of samples in which a free amino acid was detected/the total number of samples) and concentration intervals (mg/kg) are shown.

In each discrimination factor (apiary of origin and harvesting periods) we have studied the normality of the data base, involving the response of the samples to the number of canonical functions determined. The null hypothesis for the Shapiro-Wilk test is that a variable is normally distributed among some population. As a rule of thumb, we reject the null hypothesis if $p < 0.05$. Results of the univariate and multivariate normality tests showed that for the apiary of origin normality did not apply regarding the canonical functions data; meanwhile, for the harvesting period, normality was observed only in the case of the second canonical function (see Supplementary Material, Tables 9S and 10S). Multi-normality of the data is required to guarantee a solution that minimizes the expected error. In our case, non-normal data results in uncertainty in terms of finding the optimal solution. However, the results of the classification matrix show that the solution obtained is acceptable. In addition, one-way ANOVA was performed for the canonical functions data for evaluating the apiary and harvesting effect. It was seen that, in both cases and for all the canonical functions, average responses varied significantly (see Supplementary Material, Figs. 2S and 3S), which implies that there were differences between the apiaries of origin and the harvesting periods.

Given the data obtained from the concentrations for each of the free amino acids grouped by apiary and month of collection, it can be concluded that: i) the largest number and concentration of free amino acids was found in the PI apiary (3245728 mg/kg), followed by MO (846674 mg/kg), TN (398203 mg/kg) and, finally, FH (146932 mg/kg); while, samples with the highest content per sample were those from PI (71127.3 mg/kg) and MO apiaries (70556.2 mg/kg); ii) the largest total concentration of free amino acids was observed for the samples harvested in June (2506372 mg/kg), followed by the period April–May (1321242 mg/kg) and, lastly, July–August (8099923 mg/kg), which matched the order of the mean values per sample; iii) the vast majority of free amino acids in all the apiaries and for all the collection months was proline, in a concentration range of 298–569989 mg/kg; this is consistent with the related literature (Ghosh & Jung, 2022; Themelis et al., 2019). This could be explainable, as it has been demonstrated that honeybees prefer proline when choosing nectar, and because proline accumulates at a higher concentration in the nectars of many angiosperms, like some of the Brassica t. plants from which the pollen analyzed in this study derives. (Ghosh & Jung, 2022); iv) threonine was the least detected free amino acid in the samples (29 out of the 72 samples; 40% frequency), while methionine and leucine were found in 36 (50% frequency) and 40 (56% frequency) samples, respectively.

Table 3
- Overall frequency and concentration range of each amino acid in the bee pollen samples from the four apiaries.

Amino acid	Frequency[a] (%)	Concentration range[b] (mg/kg; dry weight)
ASP	98.6	594–12319
GLU	90.2	< LOQ-6608
ASN	69.4	178–15433
SER	81.9	114–17819
GLN	75.0	367–391358
HIS	98.6	675–51149
GLY	98.6	237–32545
THR	43.0	287–6044
ARG	61.1	319–8763
ALA	88.8	128–30376
TYR	98.6	131–4616
VAL	69.4	70–5888
MET	50.0	170–5411
TRP	59.7	148–2042
PHE	98.6	180–4394
ILE	98.6	204–2973
LEU	55.5	126–9896
LYS	98.6	461–4585
PRO	98.6	298–569990

[a] Number of samples in which an amino acid was detected/total number of samples (n = 72).
[b] Content over the limit of detection.

Table 4
- Total (sum of all free amino acid content of all samples) and mean (content per sample) free amino acid and content (mg/kg; dry weight) in the bee pollen samples grouped by apiary (Monte, MO; Pistacho, PI; Tío Natalio, TN; Fuentelahiguera, FH) and harvesting period (April–May, AM; June, JN; July–August, JA).

Apiary of origin			Harvesting period		
Apiary	Total content	Mean content	Period	Total content	Mean content
FH	146932	20990.3	AM	1321242	66062.1
MO	846674	70556.2	JN	2506372	83545.7
TN	398203	49775.4	JA	809923	36814.7
PI	3245728	71127.3			

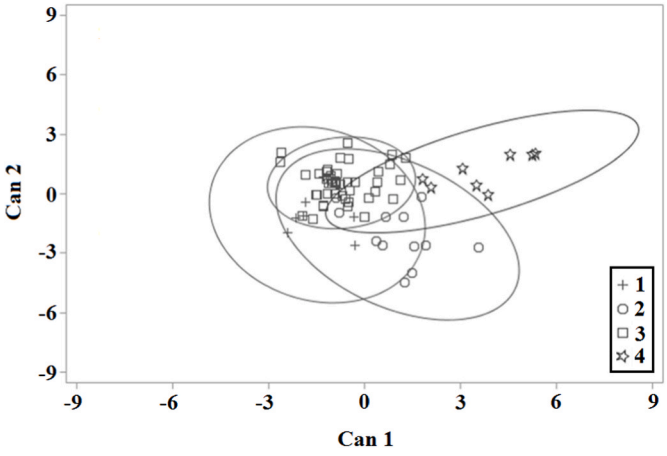

Fig. 3. Representation of the 95% prediction ellipse for each apiary origin (FH, 1; MO, 2; PI, 3; TN, 4) for the two first canonical functions.

As occurs with other bee pollen constituents, amino acid composition is strongly dependent on botanical and geographical origins (Al-Kahtani et al., 2020; Bayram, 2021; Themelis et al., 2019), although other factors such as storage and processing conditions could also affect the amino acid content of bee pollen (Ares et al., 2018). However, the samples analyzed in the present study were not processed prior to analysis, being stored in the same conditions, and obtained from the same geographical area. Differences between apiaries and harvesting periods have already been commented on in subsection 2.2, and they mainly concern the flowering plants of the surrounding areas and those which are predominant at each period of the year. Consequently, the differences observed in amino acid composition depending on the apiary of origin and harvesting periods could be tentatively attributed to the flowering plants from which the bee pollen originated (see Table 2). This is a common finding that has been extensively discussed, and our results concur with the previously reported data (Ares et al., 2022; Bayram et al., 2021; Ghosh & Jung, 2022; Stabler et al., 2018; Sommano et al., 2020; Taha et al., 2019; Thakur & Nanda, 2018; Themelis et al., 2019).

For example, it has been reported that in many cases the predominant source was Brassica t, the preeminent amino acids of which are glutamine and glutamic acid (Korus, 2014). This is of particular interest, as glutamic acid is easily converted to proline, which is the predominant amino acid. Both compounds are considered indicative of the freshness-aging characteristics of the pollen, with glutamic acid content being more abundant in fresh samples and proline content greater in dry or aged samples (Themelis et al., 2019; Yan et al., 2019). In addition, the main amino acid found in the *Cistus* genus is proline, as was the case in many of the samples analyzed (Moerman et al., 2016). Nevertheless, there is a lack of information relating to free amino acid content in the other botanical sources, and therefore it is not possible to embark on a more extensive discussion of the results regarding the botanical origin.

Finally, the differences observed in amino acid composition at different harvesting periods area concur with previous publications in which seasonal variations in the amount of bee pollen collected and its nutrient content have been reported Al-Kahtani et al., 2020; Negrão & Orsi, 2018; Al-Kahtani & Taha, 2020). In one of these studies, it was observed that the concentrations of proteins and amino acids in the bee pollen samples analyzed were significantly influenced by the harvesting season, which the authors attributed to the same reason we have already pointed out, namely, the diversity of the dominant botanical origins during each season (Al-Kahtani et al., 2020).

3.2. Canonical discriminant analysis

The main goal of the present study is to demonstrate whether free amino acids could be biomarkers of the geographical origin, apiary, and/or harvesting period. Therefore, two different statistical analyses (canonical discriminant analysis) were performed in accordance with the apiary of origin (PI, MO, TN, and FH) or the harvesting period (April–May, June, July–August). In both cases, quadratic combinations of the quantitative variables (free amino acid concentrations) were found, as these provide greater discrimination between the different groups to which they belong. All the bee pollen samples were analyzed in triplicate.

3.2.1. Apiary of origin

In this study, the sample size was 72, which were classified into four factors/groups (apiaries: FH, MO, PI, and TN) with 57 variables (19 amino acids per sample analyzed in triplicate). As can be seen, there were many variables, yet with our statistical approach we were able to reduce the dimensions without losing information, thereby obtaining an acceptable graphic representation in two or three dimensions. Based on a classification variable and several quantitative ones, the weights of the canonical functions were obtained, and their values are summarized in Table 11S (see Supplementary Material). Amino acid_1, Amino acid_2 and Amino acid_3 indicate the values obtained in each of three analyses carried out on the bee pollen samples for. It should be mentioned that the first three canonical functions explained the 100% of the variability of the original data.

Once the values have been found, the averages of the first three canonical functions for the four original apiaries are obtained (see Supplementary Material, Table 12S), and the 95% confidence ellipses of each experimental group were graphically represented (see Fig. 3). It should be specified that the location of the points in the graphic representation is the result of the weight of the canonical functions, either positive or negative, of the different amino acids. Canonical function 1 is represented on the x-axis; therefore, if the values of the function are on the right, it implies that the positive values will have greater weight than the negative ones; conversely, if the values of the canonical functions are represented on the left, the negative values will be more significant than the positive ones. Canonical function 2 is represented on the y-axis, which implies that, if the values are at the top, the positive ones will be more important than the negative ones; when these are lower, the negative values will have greater weight than the positive ones. As can be appreciated in Fig. 3, the observation of the 95% prediction ellipse for each apiary origin for the two first canonical functions does not show as clear a differentiation of the groups as the central ones do. This indicates that two canonical functions are insufficient. However, if three canonical functions were considered, the Wilk's λ statistic (0.075) and the p-value <0.0001 demonstrate that the discriminant analysis must be performed with three canonical functions.

Next, a CDA was performed by the quadratic discriminant function and the first three canonical functions, with the results shown in Table 13S (see Supplementary Material). Cross-validation concluded that the quadratic discriminant analysis permitted a correct classification of the bee pollen samples regarding their apiary of origin, as more than 75% were correctly assigned, while for two apiaries (FH and TN), the success rate was 100%. The reason why some samples were not correctly classified may fundamentally be due to most of them being multifloral; in this regard, there is a great difference in terms of plant origin, and the related amino acid content. Nevertheless, these results have confirmed the potential of amino acids as markers of the apiary of origin, because many of the bee pollen samples were accurately attributed to apiaries in the same area. This pioneer finding has scientific relevance, since, unlike previous studies where the relationship between amino acid content and botanical and/or geographical origin has been investigated, it has for the first time been shown that amino acid content can be used to differentiate the apiary of origin of the pollen samples; this increases the possibility of enhancing specification of the origin of the bee pollen.

3.2.2. Harvesting period

The sample size and number of variables were the same as in the above-mentioned study, namely, 72 and 57, respectively. In this case, however, samples were categorized in three classes that corresponded to the harvesting periods (April–May, June, July–August). The weights of the canonical functions were obtained as previously described, and these are listed in Table 14S (see Supplementary Material). Unlike the previous analysis, 100% variability of the original data could be explained with only the first two canonical functions. Subsequently, the average values of canonical functions 1 and 2 were obtained (see Supplementary Material, Table 15S), and the 95% confidence ellipses of each experimental group were graphically represented (see Fig. 4). The observation of the 95% prediction ellipse for each harvesting period for the two first canonical functions shows a clear differentiation of the three factors/groups. Although the three ellipses do not have an empty intersection, a differentiated grouping of the samples according to this factor is clearly seen in Fig. 4, this being corroborated with the Wilk's λ statistic (0.08796) and the p-value (0.034). This apparently acceptable classification of the samples was confirmed by CDA (see Supplementary Material, Table 16S), and cross validation, as more than 90% of the samples were correctly grouped to their corresponding harvesting period solely

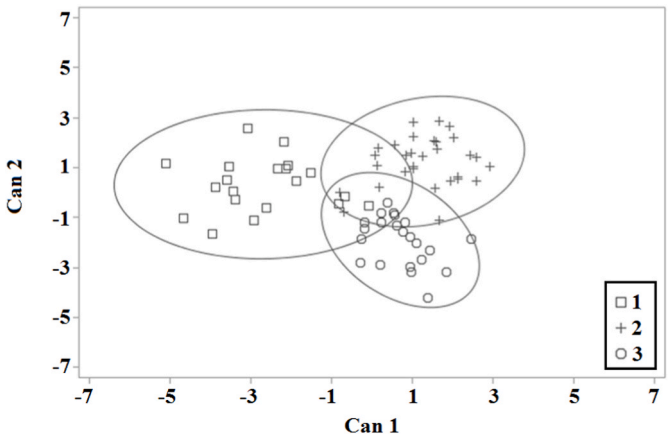

Fig. 4. Representation of the 95% prediction ellipse for each harvesting period (April–May, 1; June, 2; July–August, 3) for the two first canonical functions.

because of their free amino acid content; in one period (June), a 100% classification rate was obtained.

This is quite a relevant result, as, to the best of our knowledge, it is the first time that bee pollen samples have been differentiated with respect to the harvesting period by determining individual free amino acid content. Therefore, the number of existing options for differentiating and characterizing bee pollen samples has now increased, which may facilitate the authentication and traceability of this bee product.

4. Conclusions

An analytical study of free amino acid content by LC-FLD was carried out on 72 samples of bee pollen from four different apiaries, located in Marchamalo (Guadalajara, Spain). Variable amounts and numbers of free amino acids were found in the samples analyzed, with proline predominating in all the samples. The largest number of amino acids analyzed was found in the PI apiary, whilst the greatest concentrations of free amino acids were observed in the samples collected in June. These results were expected, as the largest number of samples corresponded to this apiary and harvesting period, although it should be mentioned that the highest mean concentration values per sample were also observed for these parameters. The differences observed in amino acid composition can be related to the flowering plants from which the bee pollen samples originated. CDAs were conducted based on the free amino acid content in the 72 bee pollen samples; it was possible to differentiate between the four apiaries or the three harvesting periods by means of the first three or two canonical functions, respectively. In addition, it was possible to assign more than 75% of the samples to the corresponding apiary, with 100% correct classification for two of them (FH and TN). It has also been shown that an even greater number of samples were correctly classified according to their harvesting period (>90%), with a 100% classification rate being obtained in one period (June). These are quite relevant findings, since, to the best of our knowledge, these studies have never previously been carried out by focusing attention on the free amino acid content of bee pollen. Finally, a new perspective for classifying pollen samples has been provided, by demonstrating for the first time that the amino acid content can be used to discriminate bee pollen samples in relation to the harvesting period. This may serve as a starting point for ascertaining whether the relationship can be extended to longer time periods, perhaps by analyzing samples for two or more consecutive years to account for variation in climatic conditions that could affect the flowering of the plants, and to samples from different regions and countries.

Author statement

Ana M. Ares: Conceptualization, Methodology, Investigation, Supervision, Validation, Visualization, Writing - original draft; **Laura Toribio:** Methodology, Investigation, Validation, Visualization, Writing - original draft; **Jesús A. Tapia:** Conceptualization, Formal analysis, Software, Visualization; Writing - review & editing; **Amelia V. González-Porto:** Conceptualization, Methodology, Investigation, Resources, Visualization, Writing - original draft; **Mariano Higes:** Conceptualization, Resources, Visualization; **Raquel Martín-Hernández:** Conceptualization, Funding acquisition, Project administration, Resources, Visualization Writing - original draft; **José Bernal:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Visualization, Writing - original draft; Writing - review & editing.

Funding

This work was supported by the Spanish "*Ministerio de Economía y Competitividad*" and the "*Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria*" (Project numbers RTA 2015-00013-C03-01 and 03).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors wish also to thank David Rixham (White Rose English School, Valladolid, Spain) for performing the English revision, and M. Palmero for her technical support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fbio.2022.102092.

References

Al-Kahtani, S. N., & Taha, E. A. (2020). Seasonal variations in nutritional composition of honeybee pollen loads. *Journal of the Kansas Entomological Society, 93*, 105–112.

Al-Kahtani, S. N., Taha, E. K., Khan, K. A., Ansari, M. J., Farag, S. A., Shawer, D. M. B., & Elnabawy, E. S. M. (2020). Effect of harvest season on the nutritional value of bee pollen protein. *PLoS One, 15*, Article e0241393.

Araújo, J. S., Chambó, E. D., Costa, M., Cavalcante da Silva, S., Lopes de Carvalho, C. A., & Estevinho, L. M. (2017). Chemical composition and biological activities of mono- and heterofloral bee pollen of different geographical origins. *International Journal of Molecular Sciences, 18*, 921.

Ares, A. M., Martín, M. T., Toribio, L., & Bernal, J. (2022). Determination of free amino acids in bee pollen by liquid chromatography with fluorescence detection. *Food Analytical Methods, 15*, 2172–2180.

Ares, A. M., Tapia, J., González-Porto, A. V., Higes, M., Martín-Hernández, R., & Bernal, J. (2022). Glucosinolates as markers of the origin and harvesting period for discrimination of bee pollen by UPLC-MS/MS. *Foods, 11*, 1446.

Ares, A. M., Valverde, S., Bernal, J. L., Nozal, M. J., & Bernal, J. (2018). Extraction and determination of bioactive compounds from bee pollen. *Journal of Pharmaceutical and Biomedical Analysis, 147*, 110–124.

Bayram, N. E. (2021). Vitamin, mineral, polyphenol, amino acid profile of bee pollen from *Rhododendron ponticum* (source of "mad honey"): Nutritional and palynological approach. *Journal of Food Measurement and Characterization, 15*, 2659–2666.

Bayram, N. E., Gercek, Y. C., Çelik, S., Mayda, N., Kostić, A.Ž., Dramićanin, A. M., & Özkök, A. (2021). Phenolic and free amino acid profiles of bee bread and bee pollen with the same botanical origin – similarities and differences. *Arabian Journal of Chemistry, 14*, Article 103004.

Bleha, R., Shevtsova, T. V., Živčáková, M., Korbářová, A., Ježková, M., Saloň, I., Brindza, J., & Synytsya, A. (2021). Spectroscopic discrimination of bee pollen by composition, color, and botanical origin. *Foods, 10*, 1682.

Campos, M. G. R., Bogdanov, S., de Almeida-Muradian, L. B., Szczesna, T., Mancebo, Y., Frigerio, C., & Ferreira, F. (2008). Pollen composition and standardisation of analytical methods. *Journal of Apicultural Research, 47*, 154–161.

Cepero, A., Ravoet, J., Gómez-Moracho, T., Bernal, J. L., Del Nozal, M. J., Bartolomé, C., Maside, X., Meana, A., González-Porto, A. V., de Graaf, D., Martín-Hernández, & R., & Higes, M. (2014). Holistic screening of collapsing honey bee colonies in Spain: A case study. *BMC Research Notes, 7*, 649.

Chica, M., & Campoy, P. (2012). Discernment of bee pollen loads using computer vision and one-class classification techniques. *Journal of Food Engineering, 112*, 50–59.

Conte, G., Benelli, G., Serra, A., Signorini, F., Bientinesi, M., Nicolella, C., Mele, M., & Canale, A. (2017). Lipid characterization of chestnut and willow honeybee-collected pollen: Impact of freeze-drying and microwave-assisted drying. *Journal of Food Composition and Analysis, 55*, 12–19.

De-Melo, A. A. M., Estevinho, L. M., Moreira, M. M., Delerue-Matos, C., de Freitas, A. S., Barth, O. M., & Almeida-Muradian, L. B. (2018). A multivariate approach based on physicochemical parameters and biological potential for the botanical and geographical discrimination of Brazilian bee pollen. *Food Bioscience, 25*, 91–110.

Escuredo, O., Silva, L. R., Valentão, P., Seijo, M. C., & Andrade, P. B. (2012). Assessing Rubus honey value: Pollen and phenolic com-pounds content and antibacterial capacity. *Food Chemistry, 130*, 671–678.

Faegri, K., & Iversen, J. (1989). *Textbook of pollen analysis*. New York, NY, USA: Wiley.

Gardana, C., Del Bo, C., Quicazán, M. C., Corrrea, A. R., & Simonetti, P. (2018). Nutrients, phytochemicals and botanical origin of commercial bee pollen from different geographical areas. *Journal of Food Composition and Analysis, 73*, 29–38.

Ghosh, S., & Jung, C. (2022). Temporal changes of nutrient composition from pollen patty to bee bread with special emphasis on amino and fatty acids composition. *Journal of Asia-Pacific Entomology, 25*, Article 101873.

Gonçalves, P. J. S., Estevinho, L. M., Pereira, A. P., Sousa, J. M. C., & Anjos, O. (2018). Computational intelligence applied to discriminate bee pollen quality and botanical origin. *Food Chemistry, 267*, 36–42.

Isopescu, R. D., Spulber, R., Josceanu, A. M., Mihaiescu, D. E., & Popa, O. (2020). Romanian bee pollen classification and property modelling. *Journal of Apicultural Research, 59*, 443–451.

Jobson, J. D. (1991). *Applied multivariate data analysis. Regression and experimental design*. New York, NY, USA: Springer.

Kafantaris, I., Amoutzias, G. D., & Mossialos, D. (2021). Foodomics in bee product research: A systematic literature review. *European Food Research and Technology, 247*, 309–331.

Kaškonienė, V., Adaškevičiūtė, V., Kaškonas, P., Mickienė, R., & Maruška, A. (2020). Antimicrobial and antioxidant activities of natural and fermented bee pollen. *Food Bioscience, 34*, Article 100532.

Korus, A. (2014). Amino acid retention and protein quality in dried kale (*Brassica oleracea* L. var. *Acephala*). *Journal of Food Processing and Preservation, 38*, 676–683.

Milla, L., Schmidt-Lebuhn, A., Bovill, J., & Encinas-Viso, F. (2022). Monitoring of honey bee floral resources with pollen DNA metabarcoding as a complementary tool to vegetation surveys. *Ecological Solutions and Evidence, 3*, Article e12120.

Moerman, R., Vanderplanck, M., Roger, N., Declèves, S., Wathelet, B., Rasmont, P., Fournier, D., & Michez, D. (2016). Growth rate of bumblebee larvae is related to pollen amino acids. *Journal of Economic Entomology, 109*, 25–30.

Negrão, A. F., & Orsi, R. O. (2018). Harvesting season and botanical origin interferes in production and nutritional composition of bee pollen. *Anais da Academia Brasileira de Ciências, 90*, 325–332.

Prđun, S., Svečnjak, L., Valentić, M., Marijanović, Z., & Jerković, I. (2021). Characterization of bee pollen: Physico-chemical properties, headspace composition and FTIR spectral profiles. *Foods, 10*, 2103.

Restrepo, L. C. C., Alabat, I. E. G., Grosso, G. S., & dos Santos, F. D. R. (2021). Markers for the spatial and temporal differentiation of bee pollen harvested by *Apis mellifera* L. in the Eastern Andes of Colombia. *Journal of Apicultural Research* (in press).

Sommano, S., Bhat, F. M., Wongkeaw, M., Sriwichai, T., Sunanta, P., Chuttong, B., & Burgett, M. (2020). Amino acid profiling and chemometric relations of black dwarf honey and bee pollen. *Frontiers in Nutrition, 7*, Article e558579.

Stabler, D., Power, E. F., Borland, A. M., Barnes, J. D. M., & Wright, G. A. (2018). A method for analysing small samples of floral pollen for free and protein-bound amino acids. *Methods in Ecology and Evolution, 9*, 430–438.

Taha, E. K. A., Al-Kahtani, S., & Taha, R. (2019). Protein content and amino acids composition of bee-pollens from major floral sources in Al-Ahsa, eastern Saudi Arabia. *Saudi Journal of Biological Sciences, 26*, 232–237.

Thakur, M., & Nanda, V. (2018). Assessment of physico-chemical properties, fatty acid, amino acid and mineral profile of bee pollen from India with a multivariate perspective. *Journal of Food and Nutrition Research, 57*, 328–340.

Themelis, T., Gotti, R., Orlandini, S., & Gatti, R. (2019). Quantitative amino acids profile of monofloral bee pollens by microwave hydrolysis and fluorimetric high performance liquid chromatography. *Journal of Pharmaceutical and Biomedical Analysis, 173*, 144–153.

Tutun, H., Aluç, Y., Kahraman, H. A., Sevin, S., Yipel, M., & Ekici, H. (2022). The content and health risk assessment of selected elements in bee pollen. *Journal of Food Composition and Analysis, 105*, Article 104234.

Wang, Z., Ren, P., Wu, Y., & He, Q. (2021). Recent advances in analytical techniques for the detection of adulteration and authenticity of bee products – a review. *Food Additives and. Contaminants Part A, 38*, 533–549.

Wang, R. Y., Zhong, M. C., Hao, N., Wang, T., & Wang, H. T. (2022). Botanical origin authenticity control of pine pollen food products using multiplex species-specific PCR method. *Food Analytical Methods, 15*, 421–427.

Yan, S., Li, Q., Xue, X., Wang, K., Zhao, L., & Wu, L. (2019). Analysis of improved nutritional composition of bee pollen (*Brassica campestris* L.) after different fermentation treatments. *International Journal of Food Science and Technology, 54*, 2169–2181.

Zeghoud, S., Rebiai, A., Hemmami, H., Ben Seghir, B., Elboughdiri, N., Ghareba, S., Ghernaout, D., & Abbas, N. (2021). ATR–FTIR spectroscopy, HPLC chromatography, and multivariate analysis for controlling bee pollen quality in some Algerian regions. *ACS Omega, 6*, 4878–4887.