Plant, Cell & Environment

WILEY

Plant, Cell & Environment

ORIGINAL ARTICLE OPEN ACCESS

Multi-Elemental Analysis for Geographical Tracing of Chickpeas Produced in Nearby Locations Around a Protected Geographical Indication

Ana Sayago[1] | Raúl González-Domínguez[1,2] | Ángeles Fernández-Recamales[1]

[1]AgriFood Laboratory, Faculty of Experimental Sciences, University of Huelva, Huelva, Spain | [2]Instituto de Investigación e Innovación Biomédica de Cádiz (INiBICA), Hospital Universitario Puerta del Mar, Universidad de Cádiz, Cádiz, Spain

Correspondence: Raúl González-Domínguez (raul.gonzalez@dqcm.uhu.es; raul.gonzalez@inibica.es) | Ángeles Fernández-Recamales (recamale@uhu.es)

Received: 24 October 2024 | **Revised:** 16 June 2025 | **Accepted:** 18 June 2025

Funding: R.G.-D. is recipient of a "Miguel Servet" fellowship (CP21/00120) funded by Instituto de Salud Carlos III.

Keywords: chickpea | geographical origin | machine learning | multi-elemental analysis | protected geographical indication | rare earth elements

ABSTRACT

The multi-elemental profile has repeatedly been proposed as a reliable indicator of the geographical origin of plant-derived foods, as mineral composition accurately reflects the local soil geochemistry and environmental factors. However, this approach may fail in distinguishing specimens from nearby locations, which are expected to be exposed to similar geoclimatic conditions. Herein, we studied 70 chickpea samples collected in four southwestern Spanish provinces, two located within the Protected Geographical Indication 'Garbanzo de Escacena' (i.e., Huelva and Sevilla), as well as other two boundary areas (i.e., Cádiz and Córdoba). Then, inductively-coupled plasma mass spectrometry was employed to simultaneously determine 31 trace elements and 16 rare-earth elements. Interestingly, we found great similarities in the mineral content of chickpeas cultivated in the regions ascribed to the Protected Geographical Indication, but these could be clearly discriminated from the rest of the samples. Afterward, the application of state-of-the-art machine learning tools provided predictive models with good performance in terms of classification accuracy, sensitivity, and specificity. In conclusion, we have demonstrated that the combination of multi-elemental analysis and advanced chemometrics could be a powerful strategy for food authentication and traceability according to the geographical origin.

1 | Introduction

Frauds related to geographical origin are one of the most important food authenticity problems in Europe, with cheap and non-genuine products being branded as high-priced regional crops (González-Domínguez et al. 2022a). Therefore, food traceability and authenticity are essential to guarantee the quality demanded by consumers. For this purpose, it should be stressed that the multi-elemental profile of plant-derived products may depend on multiple factors, such as the plant biological demand, mineral bioavailability and mobility from the soil, agronomical practices (e.g., use of fertilisers and pesticides), climatic factors, and manufacturing processes (Martínez-Ballesta et al. 2010). Thus, metal elements can be used as good markers for tracing food geographical origin. In particular, rare-earth elements (REEs) have been proven to accurately reflect the local geology and environment conditions, as they are usually not added to fertilisers (except in China) (Jiang et al. 2012) and their concentrations minimally change between crop years and according to agricultural practices (Danezis et al. 2016). Accordingly, multi-elemental fingerprinting combined with multivariate data analysis has emerged as a powerful tool to determine the geographical origin of a variety of plant foods, such as fava (Drivelos et al. 2014), lentils (Lymperopoulou et al. 2024), chickpeas (Di Donato et al. 2022), peanut (Zhao and Yang 2022), tea (Ma et al. 2016), caper (Pepi et al. 2018), wine (Aceto et al. 2018), olive

oil (Sayago et al. 2018), saffron (D'Archivio et al. 2019), pear (Coelho et al. 2019), and pistachios (Kalogiouri et al. 2021), among many other matrices. However, this approach is prone to fail when trying to distinguish specimens cultivated in nearby locations, which are expected to be exposed to similar geoclimatic conditions. Thus, further investigation is mandatory to explore the real utility of minerals in food regulation and control.

Pulses are traditional food components of the Mediterranean diet. Their consumption is recommended as part of a healthy diet, eliciting important benefits in cardiovascular and gastrointestinal systems, controlling body weight because of their low-fat content and satiety induction, and providing high supply of nutrients (Angeles et al. 2021; Milán-Noris et al. 2018; Torres-Fuentes et al. 2015; Yust et al. 2012). Moreover, legumes have a strategic value from an agricultural point of view, as their cultivation requires few water resources and increases soil fertility by fixing atmospheric nitrogen through bacterial symbiosis, which altogether reduces water and air pollution by minimising greenhouse gas emissions and the use of fertilisers (Yanni et al. 2023), thus contributing to climate change mitigation and adaptation within the framework of the 2030 Agenda for Sustainable Development. Among them, chickpeas (*Cicer arietinum L.*) are the third most cultivated grain legume species worldwide, after beans (*Phaseolus vulgaris* L.) and peas (*Pisum sativum* L.), being Spain the first producer in Europe with approximately 45.1000 tons per year (FAOSTAT Database n.d.). The quality of Spanish chickpeas is recognised with distinctive brands linked to legal recognition figures, such as the Protected Geographical Indication (PGI) "Garbanzo de Escacena". However, its production is widespread throughout Spain, including other regions neighbouring those protected under the PGI, which highlights the crucial need of accurate methods for ensuring their authenticity.

In the present study, we have investigated the potential of minerals as candidate markers to determine the geographical origin of Spanish chickpeas cultivated in nearby locations, for the first time performing a comprehensive characterisation in terms of major, minor, and REE species (see Figure 1 for a schematic representation of the study design). For this purpose, 70 chickpea samples were collected from neighbouring areas (i.e., two provinces located within the PGI "Garbanzo de Escacena", Huelva and Sevilla; and other two boundary provinces, Cádiz and Córdoba) and were subjected to multi-elemental analysis. Then, univariate and multivariate statistical techniques were employed to assess the discriminant and predictive capacity of minerals as chemical descriptors of origin.

2 | Materials and Methods

2.1 | Sample Collection

A total of 70 chickpea samples (milky white chickpea variety) were collected from four nearby southwestern Spanish provinces: Huelva ($N = 26$), Sevilla ($N = 18$), Cádiz ($N = 16$), and

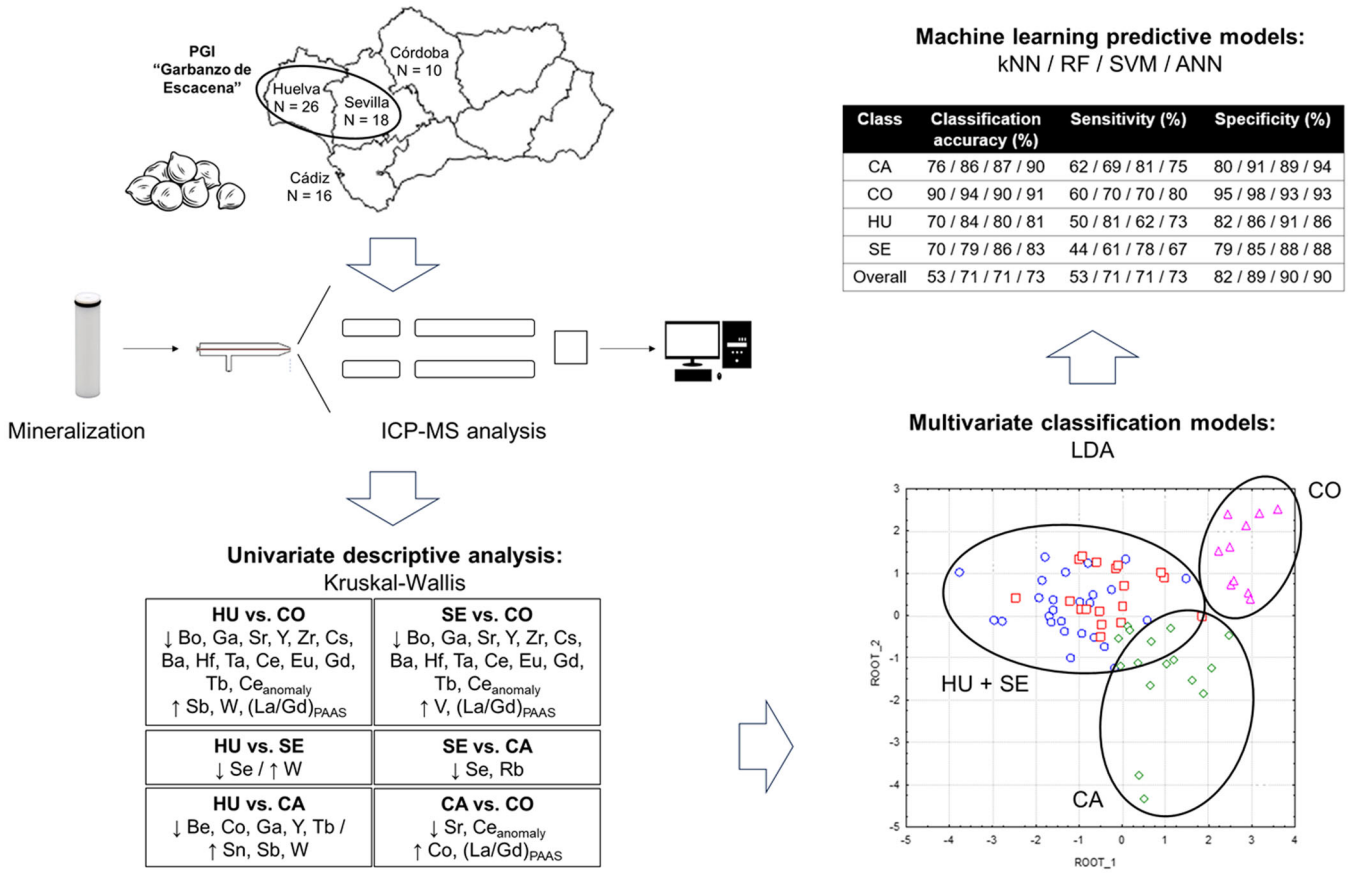

FIGURE 1 | Schematic representation of the study design, comprising sample collection, mineralisation and multi-elemental analysis, univariate descriptive analysis, application of multivariate classification tools, and machine learning predictive modelling.

Córdoba ($N = 10$), as summarised in Supporting Information S1: Table S1. Once collected, the samples were grounded to a powder size under 20 mesh and stored in polyethylene bottles until further analysis. All the samples were kindly provided by local producers.

2.2 | Mineralisation and Multi-Elemental Analysis

Chickpea samples were mineralised following the methodology proposed by Özcan et al. (Özcan et al. 2013). Briefly, 0.50 g of powdered samples were placed in 30 mL Savillex digestion vessels together with 5.0 mL nitric acid (65% v/v) and 2.0 mL hydrogen peroxide (30% v/v) (Sigma-Aldrich, Madrid, Spain). Then, the mixture was heated on a hot plate at 90°C for 24 h. Finally, digested samples were made-up to 25 mL with 2% HNO_3 and filtered through a 0.45 μm PTFE syringe filter. The reliability of the digestion method was corroborated by using a certified reference material (SRM 1573a), being our determinations in good agreement with certified values (i.e., non-significant differences between our experimental results and certified values, $p > 0.05$).

Afterward, multi-elemental analyses were performed by means of inductively coupled plasma mass spectrometry (ICP-MS), using the Agilent 7700X system (Agilent Technologies, CA, USA) equipped with nickel sampler and skimmer cones. Instrumental conditions were daily optimised by using a tuning aqueous solution containing Li, Co, Y, and Tl at 1 $\mu g\,L^{-1}$. The working parameters of the ICP-MS were as follows: radio frequency power, 1.5 kW; plasma gas flow rate, 15 $L\,min^{-1}$; auxiliary gas flow rate, 0.9 $L\,min^{-1}$; carrier gas flow rate, 1.1 $L\,min^{-1}$. The samples were infused at 400 $mL\,min^{-1}$ using a sampling depth of 9.0 mm, and the nebulization chamber temperature was set at 2°C. The isotopes monitored were ^{7}Li, ^{9}Be, ^{11}B, ^{45}Sc, ^{51}V, ^{52}Cr, ^{59}Co, ^{60}Ni, ^{63}Cu, ^{66}Zn, ^{71}Ga, ^{72}Ge, ^{75}As, ^{78}Se, ^{85}Rb, ^{88}Sr, ^{89}Y, ^{90}Zr, ^{93}Nb, ^{95}Mo, ^{111}Cd, ^{118}Sn, ^{121}Sb, ^{133}Cs, ^{137}Ba, ^{139}La, ^{140}Ce, ^{141}Pr, ^{146}Nd, ^{147}Sm, ^{151}Eu, ^{157}Gd, ^{159}Tb, ^{163}Dy, ^{165}Ho, ^{166}Er, ^{169}Tm, ^{172}Yb, ^{175}Lu, ^{178}Hf, ^{181}Ta, ^{182}W, ^{205}Tl, ^{208}Pb, ^{209}Bi, ^{232}Th, and ^{238}U. Analytical calibration curves were built by diluting a multi-elemental standard solution from SCP Science (Paris, France) with 10% HNO_3 in Milli-Q water (Millipore, Watford, UK). The concentration range was 0.2–60 $ng\,mL^{-1}$ for all the elements, with the exception of Sr, V and Zr, which were calibrated in a wider range (10–200 $ng\,mL^{-1}$). All determination coefficients of calibration curves (R^2) were higher than 0.99. Limits of detection (LOD) were calculated as the concentration equivalent to three times the standard deviation of background signal.

2.3 | Rare Earth Elements Data Preprocessing

The raw data were employed to determine various REE ratios and anomalies, according to Gallello et al. (Gallello et al. 2019). First, REE concentrations were normalised based on Post-Archean Australian Shale (PAAS) values. Then, we computed the following ratios: $(La/Yb)_{PAAS}$, as an indicator of light REE enrichment over heavy REE; $(La/Gd)_{PAAS}$ and $(La/Sm)_{PAAS}$, as indicators of light REE enrichment over medium REE; and $(Sm/Yb)_{PAAS}$, as an indicator of medium REE enrichment over heavy REE. Moreover, REE anomalies were used to assess differences between measured and expected concentrations based on neighbouring REEs, according to the following formula (1 and 2):

$$\text{Ce anomaly} = 3 \times Ce_{PAAS}/(2 \times La_{PAAS} + Nd_{PAAS}) \quad (1)$$

$$\text{Eu anomaly} = Eu_{PAAS}/(Sm_{PAAS} \times Gd_{PAAS})^{1/2} \quad (2)$$

2.4 | Data Analysis

The normal distribution of the data was tested by applying the Shapiro Wilk's test using a W statistical parameter and Skewness index, which measures the deviation of the distribution from symmetry. As most variables showed a skewed distribution and variances were not homogeneous (checked by Levene's test), the Kruskal-Wallis non-parametric multiple comparison test was employed to look for differences between the four study groups, followed by a post hoc comparison to highlight the pairs of groups responsible for those differences. Significance was set at a H-value exceeding the critical chi-squared cutoff, that is, $\chi^2\ (3,\ 0.05) = 7.815$. Then, linear discriminant analysis (LDA) was used to build multivariate classification models. Finally, the performance of various complementary supervised machine learning methods, including k nearest neighbour (kNN), support vector machines (SVM), random forest (RF), and artificial neural network (ANN), was compared to assess the predictive value of minerals as markers of food authenticity (González-Domínguez et al. 2022b). The optimisation and validation of these models was performed by 10-fold cross-validation. Their statistical performance was evaluated in terms of classification accuracy (i.e., proportion of correct predictions relative to all predictions made by the model), sensitivity (i.e., ratio of correctly classified cases belonging to a specific class), and specificity (i.e., percentage of correctly classified nonmembers of that class). All statistical analyses were performed using the Statistica 8.0 software (StatSoft, Tulsa, OK).

3 | Results

The multi-elemental profile of chickpeas was determined by ICP-MS, as summarised in Table 1. The minimum concentrations detected for some elements (Be, Ge, As, Se, Cd, Sn, Sb, Hf, Tl, Bi, Sm, Eu, Dy, Ho, Er, Yb, Lu, and U) were below their LODs. In this respect, thallium, bismuth, and arsenic were only found in less than 50% of the samples, so they were not considered in subsequent analyses. However, the rest of monitored elements were detected in quantifiable amounts in all the samples under study. Zinc was found to be the most abundant element in chickpeas (19.37 $mg\,kg^{-1}$), followed by barium (11.10 $mg\,kg^{-1}$), boron (10.62 $mg\,kg^{-1}$), and copper (7.48 $mg\,kg^{-1}$). Moderate amounts were also found for rubidium (4.40 $mg\,kg^{-1}$), strontium (4.10 $mg\,kg^{-1}$), molybdenum (1.90 $mg\,kg^{-1}$), and nickel (1.57 $mg\,kg^{-1}$), accounting for 20% of the total mineral fraction. The remaining trace elements determined in the present study showed average values below 1.0 $mg\,kg^{-1}$, thus representing less than 4% of the total mineral content. Among REEs, cerium was the most abundant element (0.65 $mg\,kg^{-1}$, 82% of the total rare earth fraction), followed by

TABLE 1 | Descriptive analysis of trace elements and rare earth elements concentrations (expressed as mg kg^{-1}) in chickpea samples.

Elements	Limit of detection	Mean	Minimum	Maximum	Standard deviation	Skewness
Trace elements						
Li	0.054	0.37	0.098	0.78	0.13	1.05
Be	0.0024	0.0013	< LOD	0.0038	0.00090	0.82
B	0.49	10.62	4.78	19.08	3.54	0.55
Sc	0.0032	0.0069	0.0012	0.047	0.0069	3.67
V	0.0027	0.021	0.0048	0.082	0.012	2.14
Cr	0.0034	0.32	0.053	1.17	0.26	1.34
Co	0.0019	0.043	0.010	0.091	0.017	0.68
Ni	0.036	1.57	0.18	8.42	1.90	2.28
Cu	0.17	7.48	1.16	27.30	4.14	2.24
Zn	0.62	19.37	4.19	40.57	6.44	1.39
Ga	0.004	0.041	0.015	0.14	0.021	2.22
Ge	0.0068	0.0019	< LOD	0.0092	0.0017	3.06
As	0.012	0.0099	< LOD	0.10	0.017	3.88
Se	0.047	0.013	< LOD	0.063	0.012	1.97
Rb	0.0069	4.40	0.57	7.48	1.74	-0.48
Sr	0.011	4.10	0.29	9.98	2.46	0.52
Y	0.0045	0.063	0.0046	0.20	0.039	1.56
Zr	0.0081	0.099	0.025	0.34	0.062	1.25
Nb	0.00045	0.0027	0.00060	0.016	0.0024	3.13
Mo	0.095	1.90	0.21	4.41	0.87	0.44
Cd	0.0027	0.0094	< LOD	0.092	0.014	4.34
Sn	0.10	0.067	< LOD	0.38	0.086	1.59
Sb	0.0096	0.0025	< LOD	0.013	0.0033	1.79
Cs	0.0021	0.024	0.0070	0.052	0.0092	0.67
Ba	0.016	11.10	0.69	20.20	3.95	-0.27
Hf	0.0013	0.0029	< LOD	0.014	0.0022	2.40
Ta	0.000095	0.00080	0.00010	0.0025	0.00060	0.78
W	0.00068	0.049	0.0021	1.34	0.19	5.89
Tl	0.0013	0.00010	< LOD	0.0011	0.00010	7.34
Pb	0.0063	0.49	0.037	8.85	1.26	5.59
Bi	0.00094	0.00080	< LOD	0.0095	0.0013	4.20
Rare earth elements						
La	0.00065	0.024	0.0066	0.15	0.025	3.78
Ce	0.0037	0.65	0.23	1.64	0.30	1.31
Pr	0.00035	0.0044	0.0011	0.030	0.0054	3.77
Nd	0.0011	0.016	0.0037	0.11	0.020	3.79
Sm	0.0014	0.0027	< LOD	0.020	0.0039	3.78
Eu	0.00036	0.0015	< LOD	0.0056	0.0010	2.81
Gd	0.00077	0.014	0.0056	0.048	0.0079	1.79
Tb	0.00017	0.00070	0.00020	0.0037	0.0007	3.39
Dy	0.00041	0.0020	< LOD	0.020	0.0037	3.83
Ho	0.000038	0.0004	< LOD	0.0040	0.00070	3.81

(Continues)

TABLE 1 | (Continued)

Elements	Limit of detection	Mean	Minimum	Maximum	Standard deviation	Skewness
Er	0.00072	0.0012	< LOD	0.012	0.0022	3.78
Tm	0.000017	0.00010	< LOD	0.0016	0.00030	3.72
Yb	0.00041	0.0010	< LOD	0.011	0.0020	3.79
Lu	0.00046	0.00010	< LOD	0.0016	0.00030	3.78
Th	0.00042	0.0057	0.00080	0.035	0.0059	3.58
U	0.00044	0.00080	< LOD	0.0025	0.00040	2.58

lanthanum (24.3 $\mu g\,kg^{-1}$), neodymium (15.6 $\mu g\,kg^{-1}$) and gadolinium (14.1 $\mu g\,kg^{-1}$), while others were found in concentrations below 10 $\mu g\,kg^{-1}$.

After multi-elemental analysis, univariate, multivariate, and machine learning statistical tools were applied to evaluate the discriminant and predictive capacity of minerals as chemical descriptors of origin. The Kruskal-Wallis non-parametric multiple comparison test evidenced that many of the trace elements under investigation showed statistically significant H-values exceeding the critical chi-squared cutoff, that is, $\chi^2(3, 0.05) = 7.815$, with the exception of Li, Sc, Cr, Ni, Cu, Zn, Ge, Nb, Mo, Cd, and Pb (Table 2). In contrast, among REEs, only Ce, Eu, Gd, and Tb, as well as the $(La/Gd)_{PAAS}$ ratio and the Ce anomaly, were found to differ between, at least, two study groups. Then, forward stepwise LDA was applied to multi-elemental data (preselecting only significant elements according to the Kruskal-Wallis test) to build classification models. Using this method, three discriminant functions were obtained, which explained 68.7%, 23.8% and 7.4% of the total variability of the data, respectively.

Finally, a variety of complementary supervised machine learning methods, namely kNN, SVM, RF, and ANN, were used to build predictive models. For this purpose, a 10-fold cross-validation approach was employed to optimise the crucial parameters driving models' performance. Briefly, the kNN algorithm was carried out by selecting the Euclidean distance as the metric to calculate between-sample similarity, and the optimal number k of neighbours was set at 5. Four different kernel functions were compared to build SVM models (i.e., sigmoid, polynomial, linear, and radial basis functions), best results being yielded by the linear kernel function and a C-value of 5. In RF, the lowest misclassification rate was obtained when using 100 decision trees and seven predictors for each node (i.e., as the square root of the total number of features), thereby providing good sample classification. Finally, optimum ANN modelling required the following parameters: two-layer neural networks with hyperbolic tangent transfer function, four output neurons and as many input neurons as variables, 100 neurons per hidden layer, and 200 iterations during back-propagation training. After models' optimisation and validation, their statistical performance was compared in terms of classification accuracy (i.e., proportion of correct predictions relative to all predictions made by the model), sensitivity (i.e., ratio of correctly classified cases belonging to a specific class), and specificity (i.e., percentage of correctly classified non-members of that class), as summarised in Table 3.

4 | Discussion

Herein, we demonstrate the potential of multi-elemental analysis in combination with cutting-edge chemometrics techniques with the aim of tracing the geographical origin of chickpeas cultivated in nearby regions. To this end, we collected 44 samples from the two provinces belonging to the PGI 'Garbanzo de Escacena' (i.e., Huelva and Sevilla) and 26 samples from boundary areas (i.e., Cádiz and Córdoba). Then, ICP-MS was employed to determine 31 trace elements and 16 REEs, as summarised in Table 1. Interestingly, major minerals showed similar levels to those reported by other authors in chickpeas from different locations around the world (Di Donato et al. 2022; Kafaoğlu et al. 2014; Gunes et al. 2009; Cabrera et al. 2003; Kaya et al. 2018). However, this is the first study performing such a comprehensive multi-elemental characterisation of chickpeas, so no reference values are available to our knowledge for other minor trace elements and REEs, although these were within concentration ranges reported in other legume species (Drivelos et al. 2014; Lymperopoulou et al. 2024).

Afterward, the Kruskal-Wallis non-parametric multiple comparison test was applied to look for differences in the multi-elemental profile between the four study groups (Table 2). Overall, the chickpea samples collected in the provinces of Huelva and Sevilla (both partially located within the PGI 'Garbanzo de Escacena' geographical area) were very similar in terms of mineral content, with slight differences being exclusively observed in selenium and wolframium levels. Interestingly, these two sampling areas showed a characteristic multi-elemental profile when compared to specimens from the easternmost province (i.e., Córdoba), as mirrored in lower contents of boron, gallium, strontium, yttrium, zirconium, caesium, barium, hafnium, tantalum, cerium, europium, gadolinium, and terbium, as well as lower Ce anomaly and higher $(La/Gd)_{PAAS}$ ratio. Furthermore, chickpeas from Córdoba had lower levels of vanadium than samples from Sevilla, as well as reduced concentrations of antimony and wolframium in comparison with samples from Huelva. To a lesser extent, univariate statistical analysis also enabled to discriminate chickpeas collected from Cádiz with respect to the samples from Huelva (i.e., Be, Co, Ga, Y, Sn, Sb, W, Tb), Sevilla (i.e., Se, Rb), and Córdoba (i.e., Co, Sr, $(La/Gd)_{PAAS}$ ratio, Ce anomaly).

After this preliminary descriptive analysis, multivariate statistical tools were employed to investigate the discriminant and predictive capacity of trace elements and REEs as chemical descriptors of chickpea geographical origin. As shown in the

TABLE 2 | Kruskal-Wallis multiple comparison of element concentrations between the four study groups.

	Concentration (mg kg^{-1}) Mean ± standard deviation				Kruskal-Wallis						
						p values					
	Huelva ($N = 26$)	Sevilla ($N = 18$)	Cádiz ($N = 16$)	Córdoba ($N = 10$)	*H*	Huelva-Sevilla	Huelva-Cádiz	Huelva-Córdoba	Sevilla-Cádiz	Sevilla-Córdoba	Cádiz-Córdoba
Trace elements											
Li	0.38 ± 0.15	0.36 ± 0.15	0.34 ± 0.083	0.39 ± 0.13	1.20	NS	NS	NS	NS	NS	NS
Be[b]	0.00089 ± 0.00058	0.0013 ± 0.00084	0.0018 ± 0.0011	0.0017 ± 0.0012	8.16	NS	0.049	NS	NS	NS	NS
B[c,e]	9.51 ± 3.94	9.63 ± 2.78	11.38 ± 2.82	13.86 ± 2.78	14.71	NS	NS	0.0037	NS	0.017	NS
Sc	0.0064 ± 0.0066	0.0047 ± 0.0025	0.0099 ± 0.011	0.0074 ± 0.0040	6.71	NS	NS	NS	NS	NS	NS
V[e]	0.021 ± 0.016	0.024 ± 0.011	0.021 ± 0.021	0.013 ± 0.0095	8.62	NS	NS	NS	NS	0.034	NS
Cr	0.41 ± 0.29	0.32 ± 0.27	0.29 ± 0.24	0.17 ± 0.034	5.41	NS	NS	NS	NS	NS	NS
Co[b,f]	0.038 ± 0.018	0.045 ± 0.014	0.053 ± 0.018	0.034 ± 0.0091	11.65	NS	0.025	NS	NS	NS	0.038
Ni	1.84 ± 1.95	2.02 ± 2.60	0.95 ± 0.74	1.16 ± 1.57	4.88	NS	NS	NS	NS	NS	NS
Cu	8.11 ± 3.95	8.28 ± 5.51	5.28 ± 1.74	8.08 ± 3.97	7.72	NS	NS	NS	NS	NS	NS
Zn	19.26 ± 7.32	20.49 ± 7.11	16.88 ± 2.26	21.72 ± 21.72	4.16	NS	NS	NS	NS	NS	NS
Ga[b,c,e]	0.036 ± 0.026	0.035 ± 0.012	0.048 ± 0.021	0.053 ± 0.012	18.73	NS	0.031	0.00090	NS	0.017	NS
Ge	0.0024 ± 0.0022	0.0015 ± 0.00071	0.0023 ± 0.0020	0.0011 ± 0.00054	6.95	NS	NS	NS	NS	NS	NS
Se[a,d]	0.016 ± 0.016	0.0053 ± 0.0034	0.020 ± 0.010	0.010 ± 0.0058	21.66	0.016	NS	NS	0.00030	NS	NS
Rb[d]	4.51 ± 1.78	3.50 ± 1.44	5.29 ± 1.80	4.44 ± 1.48	8.94	NS	NS	NS	0.020	NS	NS
Sr[c,e,f]	4.11 ± 2.47	3.63 ± 2.22	3.39 ± 1.39	6.88 ± 3.33	9.74	NS	NS	0.016	NS	0.0074	0.0070
Y[b,c,e]	0.048 ± 0.042	0.052 ± 0.018	0.076 ± 0.045	0.092 ± 0.029	20.05	NS	0.024	0.00020	NS	0.015	NS
Zr[c,e]	0.091 ± 0.071	0.076 ± 0.047	0.11 ± 0.058	0.15 ± 0.050	11.53	NS	NS	0.047	NS	0.0081	NS
Nb	0.0025 ± 0.0022	0.0025 ± 0.0033	0.0032 ± 0.0018	0.0024 ± 0.0016	4.57	NS	NS	NS	NS	NS	NS
Mo	1.75 ± 0.88	2.05 ± 0.74	1.88 ± 1.09	2.05 ± 0.72	2.05	NS	NS	NS	NS	NS	NS
Cd	0.0078 ± 0.0065	0.014 ± 0.024	0.0079 ± 0.0082	0.0075 ± 0.0073	0.79	NS	NS	NS	NS	NS	NS
Sn[b]	0.12 ± 0.10	0.056 ± 0.075	0.014 ± 0.012	0.043 ± 0.056	14.17	NS	0.0022	NS	NS	NS	NS
Sb[b,c]	0.0042 ± 0.0035	0.0027 ± 0.0038	0.00097 ± 0.0013	0.00010 ± 0.00029	21.45	NS	0.011	0.00030	NS	NS	NS
Cs[c,e]	0.021 ± 0.0090	0.021 ± 0.0065	0.027 ± 0.010	0.032 ± 0.0068	15.38	NS	NS	0.0023	NS	0.0054	NS
Ba[c,e]	9.09 ± 4.46	10.85 ± 2.51	12.11 ± 3.26	14.96 ± 2.34	18.50	NS	NS	0.00020	NS	0.0096	NS
Hf[c,e]	0.0024 ± 0.0016	0.0020 ± 0.0011	0.0034 ± 0.0022	0.0051 ± 0.0034	14.32	NS	NS	0.0084	NS	0.0041	NS
Ta[c,e]	0.00069 ± 0.00064	0.00050 ± 0.00031	0.00098 ± 0.00051	0.0012 ± 0.00048	14.36	NS	NS	0.033	NS	0.0076	NS

(Continues)

TABLE 2 | (Continued)

	Concentration (mg kg^{-1}) Mean ± standard deviation				Kruskal-Wallis *p* values						
	Huelva ($N = 26$)	**Sevilla ($N = 18$)**	**Cádiz ($N = 16$)**	**Córdoba ($N = 10$)**	***H***	**Huelva-Sevilla**	**Huelva-Cádiz**	**Huelva-Córdoba**	**Sevilla-Cádiz**	**Sevilla-Córdoba**	**Cádiz-Córdoba**
W[a,b,c]	0.11 ± 0.31	0.021 ± 0.048	0.0093 ± 0.0032	0.0059 ± 0.0012	29.65	0.0014	0.0078	0.00040	NS	NS	NS
Pb	0.94 ± 1.98	0.22 ± 0.15	0.18 ± 0.14	0.31 ± 0.44	3.96	NS	NS	NS	NS	NS	NS
Rare earth elements											
La	0.030 ± 0.036	0.018 ± 0.0052	0.028 ± 0.025	0.016 ± 0.0075	3.62	NS	NS	NS	NS	NS	NS
Ce[c,e]	0.51 ± 0.22	0.59 ± 0.20	0.76 ± 0.37	0.93 ± 0.24	19.56	NS	NS	0.00010	NS	0.010	NS
Pr	0.0056 ± 0.0077	0.0029 ± 0.0010	0.0054 ± 0.0057	0.0028 ± 0.0019	3.65	NS	NS	NS	NS	NS	NS
Nd	0.020 ± 0.029	0.010 ± 0.0036	0.019 ± 0.022	0.010 ± 0.0073	3.16	NS	NS	NS	NS	NS	NS
Sm	0.0033 ± 0.0055	0.0016 ± 0.00061	0.0035 ± 0.0043	0.0020 ± 0.0017	2.77	NS	NS	NS	NS	NS	NS
Eu[c,e]	0.0014 ± 0.0013	0.0012 ± 0.00031	0.0018 ± 0.0011	0.0018 ± 0.00051	13.92	NS	NS	0.012	NS	0.026	NS
Gd[c,e]	0.012 ± 0.0070	0.012 ± 0.0048	0.017 ± 0.011	0.020 ± 0.0050	16.67	NS	NS	0.0012	NS	0.0091	NS
Tb[b,c,e]	0.00066 ± 0.00086	0.00045 ± 0.00016	0.00085 ± 0.00082	0.00079 ± 0.00044	15.12	NS	0.033	0.013	NS	0.044	NS
Dy	0.0024 ± 0.0050	0.00086 ± 0.00041	0.0027 ± 0.0042	0.0018 ± 0.0027	6.11	NS	NS	NS	NS	NS	NS
Ho	0.00047 ± 0.00099	0.00015 ± 0.00075	0.00051 ± 0.00085	0.00035 ± 0.00058	5.16	NS	NS	NS	NS	NS	NS
Er	0.0014 ± 0.0029	0.00048 ± 0.00022	0.0016 ± 0.0026	0.0012 ± 0.0018	7.29	NS	NS	NS	NS	NS	NS
Tm	0.00018 ± 0.00040	0.000004 ± 0.000003	0.00020 ± 0.00038	0.00014 ± 0.00026	5.59	NS	NS	NS	NS	NS	NS
Yb	0.0012 ± 0.0025	0.00040 ± 0.00018	0.0014 ± 0.0026	0.0010 ± 0.0016	5.95	NS	NS	NS	NS	NS	NS
Lu	0.00018 ± 0.00037	0.00005 ± 0.00003	0.00019 ± 0.00038	0.00013 ± 0.00024	4.29	NS	NS	NS	NS	NS	NS
Th	0.0065 ± 0.0071	0.0042 ± 0.0015	0.0069 ± 0.0078	0.0047 ± 0.0044	2.59	NS	NS	NS	NS	NS	NS
U	0.00068 ± 0.00026	0.00068 ± 0.00020	0.00084 ± 0.00050	0.00098 ± 0.00068	1.23	NS	NS	NS	NS	NS	NS
Rare earth elements ratios and anomalies											
$(La/Yb)_{PAAS}$	4.33 ± 3.18	3.71 ± 1.63	2.62 ± 1.23	2.58 ± 1.31	5.11	NS	NS	NS	NS	NS	NS
$(La/Gd)_{PAAS}$[c,e,f]	0.27 ± 0.12	0.22 ± 0.11	0.22 ± 0.11	0.10 ± 0.02	14.69	NS	NS	0.00090	NS	0.015	0.032
$(La/Sm)_{PAAS}$	1.85 ± 0.68	1.77 ± 0.36	1.50 ± 0.39	1.52 ± 0.44	4.70	NS	NS	NS	NS	NS	NS
$(Sm/Yb)_{PAAS}$	2.07 ± 0.87	2.06 ± 0.70	1.70 ± 0.55	1.59 ± 0.59	4.20	NS	NS	NS	NS	NS	NS

(Continues)

TABLE 2 | (Continued)

	Concentration (mg kg^{-1}) Mean ± standard deviation				Kruskal-Wallis					
					p values					
	Huelva ($N = 26$)	Sevilla ($N = 18$)	Cádiz ($N = 16$)	Córdoba ($N = 10$)	*H*	Huelva-Sevilla	Huelva-Cádiz	Huelva-Córdoba	Sevilla-Cádiz	Sevilla-Córdoba	Cádiz-Córdoba
Ce anomaly[c,f]	13.82 ± 7.87	21.09 ± 12.46	19.92 ± 13.29	34.90 ± 10.90	15.21	NS	NS	0.00080	NS	0.015	0.014
Ey anomaly	1.23 ± 0.39	1.40 ± 0.22	1.28 ± 0.25	1.50 ± 0.32	3.83	NS	NS	NS	NS	NS	NS

Note: Superscript letters highlight significant differences between study groups in pairwise comparisons.
[a]Huelva versus Sevilla.
[b]Huelva versus Cádiz.
[c]Huelva versus Córdoba.
[d]Sevilla versus Cádiz.
[e]Sevilla versus Córdoba.
[f]Cádiz versus Córdoba.

LDA scores plot that was built using the two first discriminant functions (Figure 2), an overlapping was observed between the samples from Huelva and Sevilla, probably because of their geographical proximity and similar mineral levels. However, these samples were in great extent separated from other study areas along the first discriminant function. On the other hand, samples collected in non-PGI areas (i.e., Cádiz and Córdoba) were orthogonally separated along the second discriminant function. Interestingly, the first discriminant function was positively associated with Ba and Gd levels and negatively associated with Sb levels, whereas the most significant associations with the second discriminant function were found for Sr (positive) and Se (negative), in close agreement with univariate findings aforementioned.

Once demonstrated the discriminant power of minerals on the basis of univariate (i.e., Kruskal-Wallis test) and multivariate (i.e., LDA) findings, we implemented supervised machine learning methods to evaluate their predictive value as markers of authenticity (Table 3) (González-Domínguez et al. 2022b). In general, the worst results were provided by the kNN method, probably because this algorithm does not train a model (i.e., it does not optimise weights), but simply compares how similar the known cases are to the new cases to make a prediction. The rest of the models showed similar performance when considering the overall data set, with classification accuracies and sensitivities above 70%, and specificities around 90%. However, the most notable findings were obtained when modelling each study area separately. The highest level of accuracy (91.4%), sensitivity (80.0%), and specificity (93.3%) was achieved when applying the ANN method to Córdoba samples, in line with the aforementioned findings reporting that chickpea samples from this province show the most distinctive multi-elemental profile (Table 2). In contrast, statistical performance was lower for models considering PGI areas, regardless of the machine learning method used (classification accuracy < 85%, sensitivity < 80%, specificity < 90%), in line with the overlapping observed in the LDA model. Overall, although these parameters were slightly lower to those reported in other studies aimed at exploring the utility of multi-elemental analysis for geographical characterisation of different legume species (Drivelos et al. 2014; Di Donato et al. 2022; Foschi et al. 2020), the predictive performance was still adequate, especially considering the close proximity of the study areas.

5 | Conclusions

In summary, the combination of multi-elemental analysis and advanced chemometrics tools stands out as an excellent strategy for food authentication according to the geographical origin, even when studying very nearby regions. Using this approach, we found chickpea samples from two different provinces belonging to the PGI "Garbanzo de Escacena" (i.e., Huelva and Sevilla) to share a characteristic multi-elemental profile, which is clearly distinguishable from that of other samples collected in boundary areas (i.e., Cádiz and Córdoba). The application of state-of-the-art machine learning methods enabled us to build predictive models showing adequate performance in terms of classification accuracy, sensitivity, and specificity. However, future studies using larger number of samples and considering

TABLE 3 | Comparison of the statistical performance of predictive models built with multi-elemental data.

Method	Class	Classification accuracy	Sensitivity	Specificity
k nearest neighbour (kNN)	Cádiz ($N = 16$)	0.757	0.625	0.796
	Córdoba ($N = 10$)	0.900	0.600	0.950
	Huelva ($N = 26$)	0.700	0.500	0.818
	Sevilla ($N = 18$)	0.700	0.444	0.788
	Overall ($N = 70$)	0.529	0.529	0.824
Random forest (RF)	Cádiz ($N = 16$)	0.857	0.688	0.907
	Córdoba ($N = 10$)	0.943	0.700	0.983
	Huelva ($N = 26$)	0.843	0.808	0.864
	Sevilla ($N = 18$)	0.786	0.611	0.846
	Overall ($N = 70$)	0.714	0.714	0.886
Support vector machine (SVM)	Cádiz ($N = 16$)	0.871	0.812	0.889
	Córdoba ($N = 10$)	0.900	0.700	0.933
	Huelva ($N = 26$)	0.800	0.615	0.909
	Sevilla ($N = 18$)	0.857	0.778	0.885
	Overall ($N = 70$)	0.714	0.714	0.902
Artificial neural network (ANN)	Cádiz ($N = 16$)	0.900	0.750	0.944
	Córdoba ($N = 10$)	0.914	0.800	0.933
	Huelva ($N = 26$)	0.814	0.731	0.864
	Sevilla ($N = 18$)	0.829	0.667	0.885
	Overall ($N = 70$)	0.729	0.729	0.897

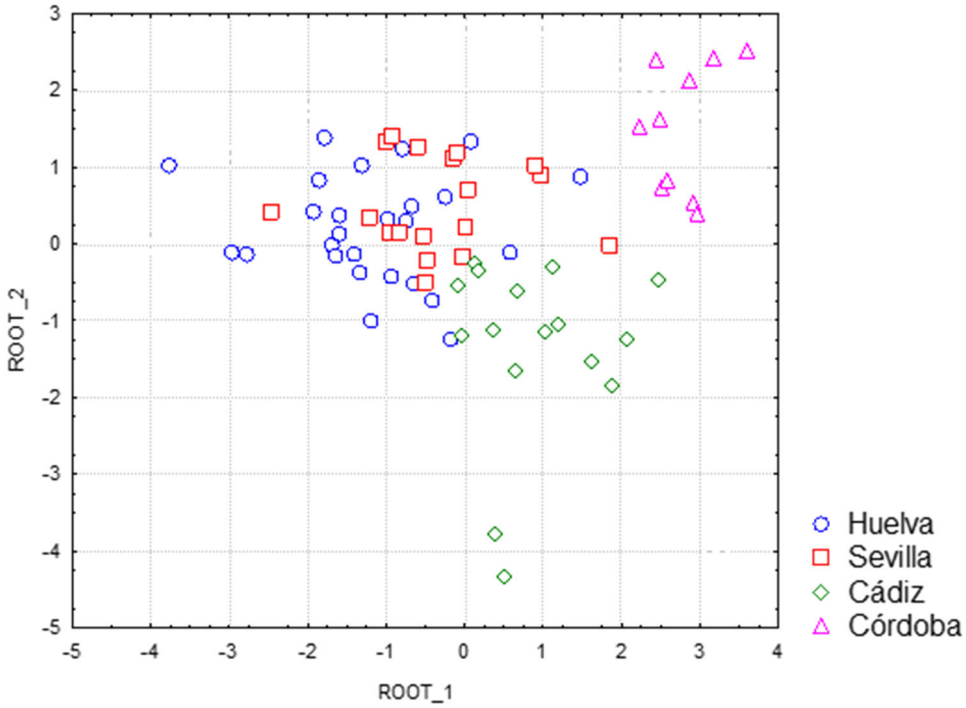

FIGURE 2 | Linear discriminant analysis (LDA) scores plot based on multi-elemental data showing the distribution of chickpea samples according to their geographical origin in the space defined by the two first discriminant functions. Legend: Huelva, blue circles; Sevilla, red squares; Cádiz, green diamonds; Córdoba, pink triangles.

other sampling locations are needed to validate our findings and to prove the robustness and reproducibility of trace elements and REEs as markers of authenticity for chickpeas. Noteworthy, it should be stressed that this methodology is of potential applicability to any plant-based food, opening the door to the widespread use of minerals as reliable tracers of authenticity and traceability in the food industry.

Acknowledgements

The authors acknowledge the Protected Designation of Origin "Garbanzo de Escacena" for kindly providing samples. R.G.-D. is recipient of a "Miguel Servet" fellowship (CP21/00120) funded by Instituto de Salud Carlos III. Funding for open access charge: Universidad de Huelva/CBUA.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

Aceto, M., F. Bonello, D. Musso, C. Tsolakis, C. Cassino, and D. Osella. 2018. "Wine Traceability With Rare Earth Elements." *Beverages* 4: 23.

Angeles, J. G. C., J. C. Villanueva, L. Y. C. Uy, et al. 2021. "Legumes as Functional Food for Cardiovascular Disease." *Applied Sciences* 11: 5475.

Cabrera, C., F. Lloris, R. Giménez, M. Olalla, and M. C. López. 2003. "Mineral Content in Legumes and Nuts: Contribution to the Spanish Dietary Intake." *Science of the Total Environment* 308: 1–14.

Coelho, I., A. S. Matos, R. Teixeira, et al. 2019. "Combining Multi-element Analysis and Chemometrics to Trace the Geographical Origin of Rocha Pear." *Journal of Food Composition and Analysis* 77: 1–8.

Danezis, G. P., A. S. Tsagkaris, F. Camin, V. Brusic, and C. A. Georgiou. 2016. "Food Authentication: Techniques, Trends & Emerging Approaches." *TrAC, Trends in Analytical Chemistry* 85: 123–132.

D'Archivio, A. A., M. L. Di Vacri, M. Ferrante, M. A. Maggi, S. Nisi, and F. Ruggieri. 2019. "Geographical Discrimination of Saffron (*Crocus sativus* L.) Using ICP-MS Elemental Data and Class Modeling of PDO Zafferano Dell'aquila Produced in Abruzzo (Italy)." *Food Analytical Methods* 12: 2572–2581.

Di Donato, F., F. Squeo, A. Biancolillo, L. Rossi, and A. A. D'Archivio. 2022. "Characterization of High Value Italian Chickpeas (*Cicer arietinum L.*) by Means of ICP-OES Multi-Elemental Analysis Coupled With Chemometrics." *Food Control* 131: 108451.

Drivelos, S. A., K. Higgins, J. H. Kalivas, S. A. Haroutounian, and C. A. Georgiou. 2014. "Data Fusion for Food Authentication. Combining Rare Earth Elements and Trace Metals to Discriminate 'Fava Santorinis' From Other Yellow Split Peas Using Chemometric Tools." *Food Chemistry* 165: 316–322.

Foschi, M., A. A. D'Archivio, and L. Rossi. 2020. "Geographical Discrimination and Authentication of Lentils (*Lens culinaris* Medik.) by ICP-OES Elemental Analysis and Chemometrics." *Food Control* 118: 107438.

FAOSTAT Database. n.d. Accessed October 21, 2024. https://www.fao.org/faostat/en/#data.

Gallello, G., C. Ferro-Vázquez, S. Chenery, et al. 2019. "The Capability of Rare Earth Elements Geochemistry to Interpret Complex Archaeological Stratigraphy." *Microchemical Journal* 148: 691–701.

González-Domínguez, R., M. Santos-Martín, A. Sayago, and Á. Fernández-Recamales. 2022a. "Authenticity and Chemometrics of Fruits and Vegetables." In *Chemometrics and Authenticity of Foods of Plant Origin*, edited by S. Agriopoulou, M. Tarapoulouzi, and T. Varzakas, 90–108. CRC Press.

González-Domínguez, R., A. Sayago, and Á. Fernández-Recamales. 2022b. "An Overview on the Application of Chemometrics Tools in Food Authenticity and Traceability." *Foods* 11: 3940.

Gunes, A., A. Inal, and Y. K. Kadioglu. 2009. "Determination of Mineral Element Concentrations in Wheat, Sunflower, Chickpea and Lentil Cultivars in Response to P Fertilization by Polarized Energy Dispersive X-Ray Fluorescence." *X-Ray Spectrometry* 38: 451–462.

Jiang, D. G., J. Yang, S. Zhang, and D. J. Yang. 2012. "A Survey of 16 Rare Earth Elements in the Major Foods in China." *Biomedical and Environmental Sciences: BES* 25: 267–271.

Kafaoğlu, B., A. Fisher, S. Hill, and D. Kara. 2014. "Chemometric Evaluation of Trace Metal Concentrations in Some Nuts and Seeds." *Food Additives & Contaminants: Part A* 31: 1529–1538.

Kalogiouri, N. P., N. Manousi, D. Klaoudatos, T. Spanos, V. Topi, and G. A. Zachariadis. 2021. "Rare Earths as Authenticity Markers for the Discrimination of Greek and Turkish Pistachios Using Elemental Metabolomics and Chemometrics." *Foods* 10: 349.

Kaya, M., A. Kan, A. Yilmaz, R. Karaman, and A. Sener. 2018. "The Fatty Acids and Mineral Composition of Different Chickpea Cultivars Cultivated." *Fresen Environ Bull* 27: 1240–1247.

Lymperopoulou, T., K. Balta-Brouma, L. A. Tsakanika, C. Tzia, A. Tsantili-Kakoulidou, and F. Tsopelas. 2024. "Identification of Lentils (*Lens culinaris* Medik) From Eglouvi (Lefkada, Greece) Based on Rare Earth Elements Profile Combined With Chemometrics." *Food Chemistry* 447: 138965.

Ma, G., Y. Zhang, J. Zhang, et al. 2016. "Determining the Geographical Origin of Chinese Green Tea by Linear Discriminant Analysis of Trace Metals and Rare Earth Elements: Taking Dongting Biluochun as an Example." *Food Control* 59: 714–720.

Martínez-Ballesta, M. C., R. Dominguez-Perles, D. A. Moreno, et al. 2010. "Minerals in Plant Food: Effect of Agricultural Practices and Role in Human Health. A Review." *Agronomy for Sustainable Development* 30: 295–309.

Milán-Noris, A. K., J. A. Gutiérrez-Uribe, A. Santacruz, S. O. Serna-Saldívar, and C. Martínez-Villaluenga. 2018. "Peptides and Isoflavones in Gastrointestinal Digests Contribute to the Anti-Inflammatory Potential of Cooked or Germinated Desi and Kabuli Chickpea (*Cicer arietinum L.*)." *Food Chemistry* 268: 66–76.

Özcan, M. M., N. Dursun, and F. Al Juhaimi. 2013. "Macro- and Microelement Contents of Some Legume Seeds." *Environmental Monitoring and Assessment* 185: 9295–9298.

Pepi, S., A. Sardella, A. Bonazza, and C. Vaccaro. 2018. "Geochemical Caper Fingerprints as a Tool for Geographical Origin Identification." *Environmental Geochemistry and Health* 40: 1385–1403.

Sayago, A., R. González-Domínguez, R. Beltrán, and Á. Fernández-Recamales. 2018. "Combination of Complementary Data Mining Methods for Geographical Characterization of Extra Virgin Olive Oils Based on Mineral Composition." *Food Chemistry* 261: 42–50.

Torres-Fuentes, C., M. M. Contreras, I. Recio, M. Alaiz, and J. Vioque. 2015. "Identification and Characterization of Antioxidant Peptides From Chickpea Protein Hydrolysates." *Food Chemistry* 180: 194–202.

Yanni, A. E., S. Iakovidi, E. Vasilikopoulou, and V. T. Karathanos. 2023. "Legumes: A Vehicle for Transition to Sustainability." *Nutrients* 16: 98.

Yust, M. M., M. C. Millán-Linares, J. M. Alcaide-Hidalgo, F. Millán, and J. Pedroche. 2012. "Hypocholesterolaemic and Antioxidant Activities of Chickpea (*Cicer arietinum* L.) Protein Hydrolysates." *Journal of the Science of Food and Agriculture* 92: 1994–2001.

Zhao, H., and Q. Yang. 2022. "Study on Influence Factors and Sources of Mineral Elements in Peanut Kernels for Authenticity." *Food Chemistry* 382: 132385.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.