



Geographical discrimination of garlic (*Allium Sativum L.*) based on Stable isotope ratio analysis coupled with statistical methods: The Italian case study

Silvia Pianezze^a, Matteo Perini^{a,*}, Luana Bontempo^a, Luca Ziller^a, Angelo Antonio D'Archivio^b

^a Fondazione Edmund Mach, via E. Mach, 1, 38010, San Michele all'Adige, Trento, Italy

^b Dipartimento di Scienze Fisiche e Chimiche, Università degli Studi dell'Aquila, Via Vetoio, 67100, Coppito, L'Aquila, Italy

ARTICLE INFO

Keywords:

Garlic
Stable isotope analysis
Geographical origin
SIMCA model

ABSTRACT

Stable isotope ratio analysis of the major bioelements ($\delta^2\text{H}$, $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$, $\delta^{34}\text{S}$), considered for the first time as a whole, was used to attempt a geographical characterization of red garlic (*Allium sativum L.*) cultivated throughout different Italian territories (Abruzzo, Lazio and Sicily). Up to now, no official methods are available to determine the geographical origin of this type of product. In this context, the Isotope Ratio Mass Spectrometry method (IRMS) represents a powerful analytical technique. The characteristic ranges of variability of the five isotope ratios in 56 red Italian garlic samples are here presented as well as their relationships. The geographical origin has some influence over the different ratios, although their data distribution shows some overlap when Principal Component Analysis (PCA) is applied. In spite of the relative closeness of the sampling sites, a model with very good predictive performance of the geographical classification was achieved by Linear Discriminant Analysis (LDA) and k-nearest neighbours (k-NN) method. Moreover, preliminary class modelling based on Soft Independent Modelling of Class Analogy (SIMCA) supports the ability of stable isotope ratios analysis for the geographical traceability of garlic.

1. Introduction

Garlic (*Allium sativum L.*) is a food condiment used worldwide both as a raw vegetable (fresh leaves or cloves) or as a source of oil, extract or powder after processing it. Water is the largest component (65%) in garlic cloves, followed by carbohydrates (28%, mainly fructans), sulphur compounds (1–4%), proteins (2%), fibres (1.5%) and free amino acids (1–1.5%) (Rahman, 2003). Volatile organo-sulphur compounds are responsible for garlic's typical pungent aroma and for a wide variety of medical properties attributed to it by traditional and modern medicine, including those of an anti-inflammatory, an antimicrobial, a cardioprotective, an anticancer and an antidiabetic (Carson, 1987; Martins et al., 2016; Yun et al., 2014). Alliin (S-allyl-L-cysteine sulfoxide) and other S-alk(en)yl-L-cysteine-S-oxides are the precursors of the bio- and odor-active sulphur components which are generated by the enzymatic action of alliinase when fresh garlic is crushed or cut and successive chain oxidation reactions (Lanzotti, 2006).

Several local Italian garlics are appreciated for their unique flavour and taste compared to the imported product (mainly from Spain, followed by Netherlands, China, France and Argentina) (ISMEA, April,

2014). Great attention nowadays is paid to the origin of foodstuffs, in particular of those with a specific and peculiar geographical identity, which contributes to an increase in their reputation and commercial value. With the aim of preserving and adding value to the traditional ecotypes of the respective territories, and to protect consumers and producers from commercial fraud over the last few decades, production consortia have been constituted in many parts of Italy. Moreover, the link between garlic quality and the geographical origin has been officially recognised with regards to *Aglia di Voghiera* and *Aglia Bianca Polesano* by the European Commission through the attribution of the Protected Designation of Origin (PDO) certification mark ('Quality products registers', 2018). High resolution magic angle spinning-nuclear magnetic resonance (HRMAS-NMR) spectroscopy (Ritota et al., 2012), Fourier-transform infrared spectroscopy (Lu et al., 2011; Ritota et al., 2012), high resolution mass spectrometry (Hrbek et al., 2018), high performance liquid chromatography (Montaño et al., 2011) and electronic nose (Trirongjitmoah et al., 2015) were previously applied to classify garlic according to the cultivars and/or the geographical origin. In addition, the elemental profile determined by atomic spectroscopy was successfully used for the geographical discrimination of garlic

* Corresponding author. Fondazione Edmund Mach, via E. Mach, 1, 38010, San Michele all'Adige, Italy.

E-mail address: matteo.perini@fmach.it (M. Perini).

<https://doi.org/10.1016/j.fct.2019.110862>

Received 31 May 2019; Received in revised form 20 August 2019; Accepted 27 September 2019

Available online 27 September 2019

0278-6915/ © 2019 Elsevier Ltd. All rights reserved.

cultivated in different countries (Smith, 2005) or even in close areas (D'Archivio et al., 2019). In particular, this last work was carried out on Italian red garlic varieties. Linear Discriminant Analysis (LDA) based on the concentration of selected elements (Ba, Ca, Fe, Mg, Mn, Na and Sr) determined by atomic spectroscopy (D'Archivio et al., 2019) was effective in discriminating the different provenances.

Few works on the geographical classification of garlic using stable isotope ratio analysis have been carried out worldwide, one of which was used to discriminate between Asian garlic samples coming from China, South Korea, Vietnam and Taiwan (Liu et al., 2018) and another to characterize garlic samples coming from different parts of Slovenia (Opatić et al., 2017). In both these studies, acceptable discrimination of garlic samples cultivated in different regions or countries was obtained by combining isotopic composition data with the profiles of a number of trace elements. In particular, the stable isotope analysis of $\delta^{13}\text{C}$, $\delta^{15}\text{N}$ and $\delta^{34}\text{S}$ in association with the elemental composition of nine elements was able to differentiate the geographical origin of garlic varieties from Slovenia (Opatić et al., 2017). Moreover, $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values were clearly effective in differentiating Argentinian garlic from Asian samples (Liu et al., 2018).

In the present investigation, for the first time stable isotope ratios of five bioelements ($\delta^2\text{H}$, $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$ and $\delta^{34}\text{S}$) were used all together to characterize, on the basis of the geographical origin, 56 red garlic samples cultivated in three Italian regions, Lazio, Abruzzo (Central Italy) and Sicily (Southern Italy) (Fig. 1). As Italy represents an area which has never been taken under evaluation in this sector, we provided an initial range of data that can be used for further comparisons to other studies. Furthermore, the isotopic composition was used to discriminate the samples through univariate and multivariate statistical approaches.

2. Material and methods

2.1. Sampling

We analysed 56 samples of garlic coming from three different regions of Italy: Abruzzo (A), Lazio (L) and Sicily (S) (Fig. 1). In particular, red garlic samples cultivated in 2017 were acquired from the consortia working in the territories of Sulmona (A-SU) (N = 13), Nubia (S-NU) (N = 9), Procono (L-PR) (N = 15) and Castelliri (L-CL) (N = 10). The producers that are part of these consortia, apart from sharing agronomic practices and marketing policies, breed local bulbs, which should ensure specificity and homogeneity of garlic varieties produced in each of the above areas. In addition, red garlic samples cultivated in three distinct areas near L'Aquila (A-AQ) (Civita di Bagno (N = 3), San Gregorio (N = 3) and Pianola (N = 3), were collected.

2.2. Samples preparation

Two or three cloves were taken from every bulb of garlic, separated from its peel and cut in pieces. Samples thus prepared were lyophilised and grinded in order to obtain a powder.

2.3. Stable isotope ratio analysis

The $^{13}\text{C}/^{12}\text{C}$, $^{15}\text{N}/^{14}\text{N}$ and $^{34}\text{S}/^{32}\text{S}$ ratios were measured using an isotope ratio mass spectrometer (IsoPrime, Isoprime Limited, Germany) after total combustion in an elemental analyser (VARIO CUBE, Isoprime Limited, Germany). The $^2\text{H}/^1\text{H}$ and $^{18}\text{O}/^{16}\text{O}$ ratios were measured using an isotope ratio mass spectrometer (IRMS) (Finnigan DELTA XP, Thermo Scientific) coupled with a pyrolyser (Finnigan DELTA TC/EA, high temperature conversion elemental analyser, Thermo Scientific). To analyse the samples, the amount introduced in the mentioned

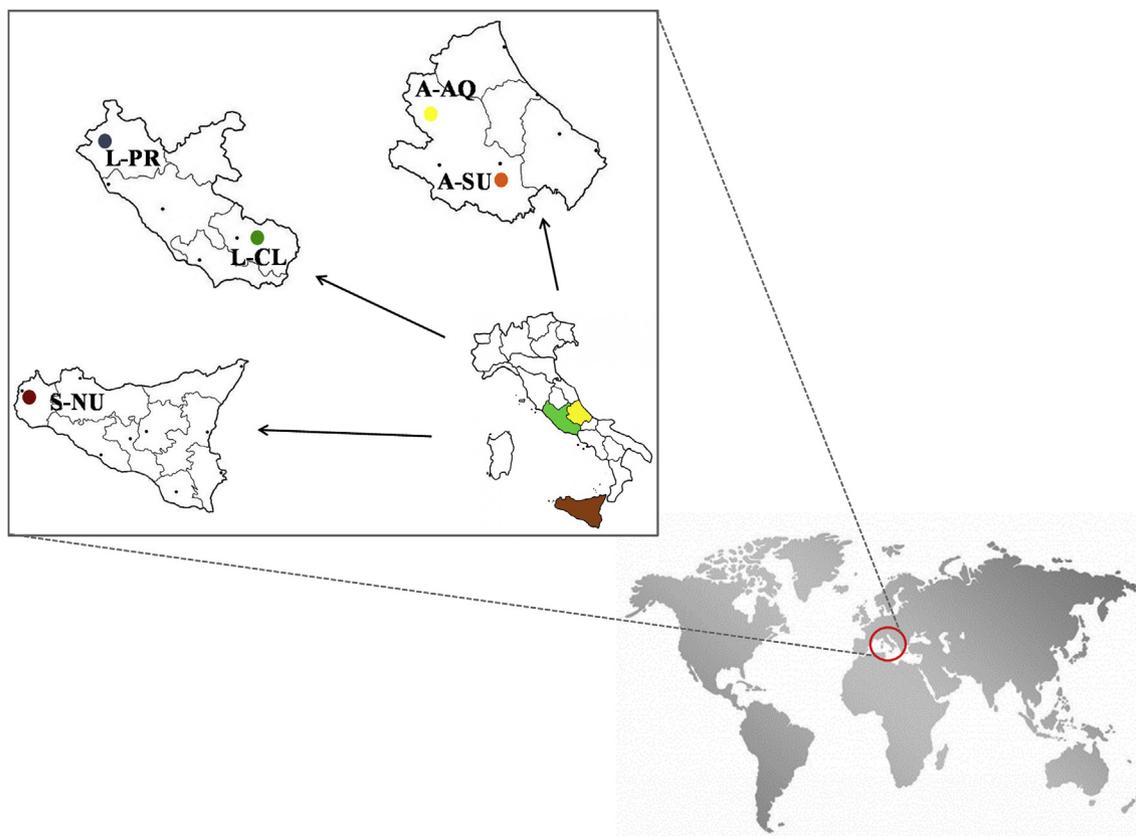


Fig. 1. Geographical position of sampling sites: A-AQ (yellow spot) and A-SU (orange spot) in Abruzzo, L-PR (blue spot) and L-CL (dark green spot) in Lazio and finally S-NU (red spot) in Sicily. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

instruments was, respectively, 0.5 and 0.2 mg.

According to the IUPAC protocol, the values are denoted in delta in relation to the international V-PDB (Vienna-Pee Dee Belemnite) for $\delta^{13}\text{C}$, V-SMOW (Vienna-Standard Mean Ocean Water) for $\delta^2\text{H}$ and $\delta^{18}\text{O}$, Air for $\delta^{15}\text{N}$ and V-CDT (Vienna-Canyon Diablo Troilite) for $\delta^{34}\text{S}$ according to the following general equation:

$$\delta_i E = (iR_{\text{sample}} - iR_{\text{reference}}) / iR_{\text{reference}}$$

where i is the mass number of the heavier isotope of element E (for example, ^{13}C); R_{SA} is the respective isotope ratio of a sample (such as for C : number of ^{13}C atoms/number of ^{12}C atoms or as approximation $^{13}\text{C}/^{12}\text{C}$); R_{REF} is the respective isotope ratio of internationally recognized reference material.

The delta values are multiplied by 1000 and expressed in units “per mil” (‰).

The isotopic values $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ were calculated against two working in-house standards (caseins), the first with $\delta^{13}\text{C} = -21.98\text{‰}$ and $\delta^{15}\text{N} = 7.38\text{‰}$, the second with $\delta^{13}\text{C} = -28.60\text{‰}$ and $\delta^{15}\text{N} = -3.40\text{‰}$. They have themselves been calibrated against international reference materials: fuel oil NBS-22 with $\delta^{13}\text{C} = -30.03\text{‰}$, sucrose IAEA-CH-6 with $\delta^{13}\text{C} = -10.45\text{‰}$ (IAEA-International Atomic Energy Agency, Vienna, Austria), and L-glutamic acid USGS 40 with $\delta^{13}\text{C} = -26.39\text{‰}$ and $\delta^{15}\text{N} = -4.52\text{‰}$ (U.S. Geological Survey, Reston, VA, USA) for $^{13}\text{C}/^{12}\text{C}$ and $^{15}\text{N}/^{14}\text{N}$ and potassium nitrate IAEA-NO₃ ($\delta^{15}\text{N} = +4.7\text{‰}$) from IAEA for $^{15}\text{N}/^{14}\text{N}$.

Keratins CBS (Caribou Hoof Standard $\delta^2\text{H} = -157 \pm 2\text{‰}$ and $\delta^{18}\text{O} = +3.8 \pm 0.1\text{‰}$) and KHS (Kudu Horn Standard, $\delta^2\text{H} = -35 \pm 1\text{‰}$ and $\delta^{18}\text{O}/^{16}\text{O} = +20.3 \pm 0.2\text{‰}$) from U.S. Geological Survey, were used to obtain $^{18}\text{O}/^{16}\text{O}$ and $^2\text{H}/^1\text{H}$ values.

Barium sulphates IAEA-SO-5 ($\delta^{34}\text{S} = +0.5\text{‰}$) and NBS 127 ($\delta^{34}\text{S} = +20.3\text{‰}$) from IAEA were used to obtain $^{34}\text{S}/^{32}\text{S}$ values.

The uncertainty of the method (calculated as one standard deviation when analysing the same sample at least ten times under reproducible conditions) was 0.1‰ for $\delta^{13}\text{C}$ values, 0.2‰ for $\delta^{15}\text{N}$ values, 0.3‰ for $\delta^{18}\text{O}$ and $\delta^{34}\text{S}$ values and 1‰ for $\delta^2\text{H}$ values.

2.4. Statistical analysis

The data were statistically evaluated using Statistica For Windows v 13.1 (StatSoft Inc., Tulsa, OK, USA) and V-PARVUS 2010. Statistically significant differences were found using a Tukey HSD test. In all the cases, the cutoff value was set at $p < 0,05$, which is associated with a significant difference between groups of values.

Preliminary exploration of stable isotope ratio data matrix was conducted by Principal Component Analysis (PCA). PCA identifies orthogonal directions, or principal components (PCs), the first describing the largest variance, the second describing the second-largest variance, and so on. Patterns in a data matrix can be emphasized by projecting objects and variables into the space of few significant PCs with minimal loss of information. LDA was then applied in order to attempt a geographical classification of the garlic samples. LDA (Härdle and Simar, 2015) assumes that the various classes collecting similar objects (here garlic samples produced in a specific territory) are described by multivariate normal distributions having the same covariance but a different location of centroids within the variable domain. Separations

between classes are hyperplanes and the allocation of a given object within one of the classes is based on a *maximum likelihood discriminant rule*. Before LDA, the samples were split into a calibration set and a prediction set by a duplex Kennard-Stone algorithm (Kennard and Stone, 1969; Snee, 1977) separately applied to each garlic class. The generalization capability of the LDA model built with the calibration data was finally tested on the prediction set. Classification performance was quantified by the percentage of calibration or external data samples correctly assigned to the predefined classes. The results of LDA classification were further visualized by projecting the classes into the space of canonical variables or discriminant functions (DFs), namely the direction that maximizes the ratio of between-class to within-class variance. Geographical discrimination of garlic samples was also attempted by using k-nearest neighbours (k-NN) method, a non-parametric approach based on a distance measure of the object to classify from all the other objects of the training set. The object is assigned to the most common class among its k nearest neighbours. Soft independent modelling of class analogy (SIMCA) (Wold and Sjöström, 1977), the first class modelling method introduced in chemistry, was also applied to test the ability of stable isotope ratio analysis for authentication purposes. SIMCA assumes that a principal component model of appropriate dimensionality can describe the typical class variability. The number of significant PCs is generally different for the various classes and it is identified by a cross-validation scheme. The range of the scores delimits the model. It follows that model class shapes are rectangles (two PCs), parallelepipeds (three PCs) or hyperparallelepipeds (more than three PCs). The decision on the acceptance or refusal of a new object is based on an F-test in which the distance of the object is compared with the mean distance of the objects used to develop the model (Forina et al., 2008). The program package V-PARVUS 2010 (Forina, M., Lanteri, S., Armanino, C., Casolino, C., Casale, M., & Oliveri, P., 2010) was used to perform multivariate statistical analyses.

3. Results and discussion

3.1. Isotopic composition

The $\delta^{13}\text{C}$ values range from -27.1 to -24.6‰ (Table 1). Similar variability ranges were reported by Liu et al. (from -29.9 to -24.4‰) for samples from Argentina and Asia and by Opatić et al. (from -28.5 to -25.6‰) for samples from Slovenia (Liu et al., 2018; Opatić et al., 2017).

This behavior is in line with the botanical origin of the matrix, indeed while C4 plants have $\delta^{13}\text{C}$ values between -14 and -12‰ , C3 plants (such as garlic) range from -30 to -23‰ (Camin et al., 2008, 2004; Crittenden et al., 2007; Knobbe et al., 2006; Molkentin, 2009; Molkentin and Giesemann, 2007).

The $\delta^{13}\text{C}$ analysis is quite effective in discriminating between two macro areas: Central Italy (with Abruzzo and Lazio) and the South (with Sicily). However, as it is to be expected, a complete differentiation of the five locations is not feasible using just this parameter. In particular, Sicily showed the higher values while Abruzzo the lower ones, probably due to the climatic effect (higher mean temperature in the South of Italy) on the stable isotope composition (Lerman et al., 1967). Indeed,

Table 1

Means and standard deviations of the isotopic ratios in Italian garlic samples, divided into the five groups under analysis (A-AQ and A-SU in Abruzzo, L-CL and L-PR in Lazio and S-NU in Sicily).

Region	Sampling Zone	N. samples	$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^{15}\text{N}$	$\delta^{13}\text{C}$	$\delta^{34}\text{S}$
Abruzzo	AQ	9	-56 ± 3	30.6 ± 0.9	3.4 ± 1.9	-25.8 ± 0.6	2.3 ± 3.8
	SU	13	-61 ± 6	30.2 ± 1.2	-0.3 ± 2.4	-26.3 ± 0.6	-0.5 ± 1.2
Lazio	CL	10	$-44 \pm$	31.6 ± 0.7	1.4 ± 1.2	-26.0 ± 0.5	7.4 ± 2.3
	PR	14	-41 ± 3	33.1 ± 1.2	0.5 ± 2.3	-25.6 ± 0.3	-1.8 ± 1.6
Sicily	NU	9	-38 ± 4	30.9 ± 0.4	0.9 ± 0.8	-25.0 ± 0.3	-0.1 ± 0.7

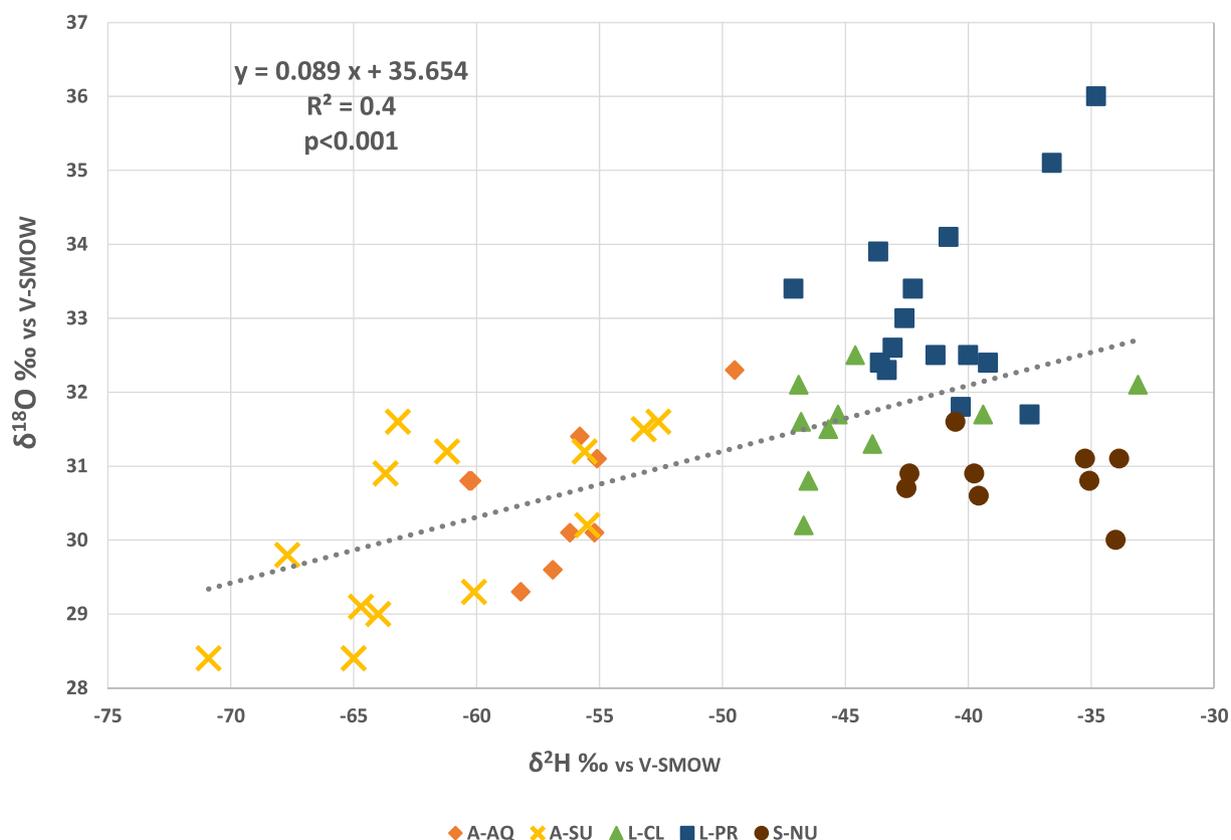


Fig. 2. Scatterplot of $\delta^2\text{H}$ vs $\delta^{18}\text{O}$.

the carbon isotopic composition shows a correlation with variations in the regional temperature conditions during the life of the plant (Lerman et al., 1967). The dependence on temperature can be explained considering both the changes in the coefficient of CO_2 utilization at different temperature conditions, and the result of the thermodynamically ordered distribution of isotopes in biological systems (Galimov, 1985).

As reported in literature, oxygen and hydrogen isotopic ratios depend on many variables, such as geographical origin, altitude, latitude, proximity to the sea (Bowen et al., 2007).

It is well known that hydrogen and oxygen isotopic values are strictly correlated in precipitation/meteoric water (Araguas et al., 1996). With regards to plants, this correlation is still valid, even if it is not that strict. Indeed, in a plant, the only source of hydrogen is water, however, oxygen also derives from O_2 and CO_2 (Barbour, 2007). Our results confirm this behavior (Fig. 2) with a significant linear relationship (R^2 value of 0.4 and a $p < 0.001$) expressed by the equation $\delta^2\text{H} = 0.0908 \delta^{18}\text{O} + 35.777$.

Despite an incomplete differentiation between Lazio and Abruzzo $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values (Table 1), it is possible to notice that Abruzzo samples tend to have lower values than Lazio ones. Abruzzo and Lazio are at approximately the same latitude (42°N), which is one of the parameters that most influences oxygen and hydrogen isotopic ratios. Therefore, the altitude of the sampling points of L'Aquila (714 m ASL) and Sulmona (405 m) (in Abruzzo) and of Castelliri (261 m) and Procono (418 m) (in Lazio) was considered, to check if this parameter could be the cause of the difference between Abruzzo and Lazio's $\delta^2\text{H}$ and $\delta^{18}\text{O}$. With the exception of L'Aquila, the similar altitudes between the sampling points seem to not justify significant differences in the hydrogen and oxygen isotopic compositions.

To evaluate correlation with the isotopic signature of rainfall, in the absence of direct measurement of the $\delta^2\text{H}$ and $\delta^{18}\text{O}$ of rainwater, we used water isotope data from the WaterIsotope database administered by Gabriel Bowen. The data available in the database [http://wateriso.](http://wateriso.utah.edu)

[utah.edu](http://wateriso.utah.edu) are the monthly weighted average precipitation values for sites all over the world.

Abruzzo and Lazio, despite their analogous latitudes and altitudes, have different $\delta^2\text{H}$ (on average -32‰ vs -29‰) and $\delta^{18}\text{O}$ values (on average -5.4‰ vs -5.1‰) of rainfall. This entails that the garlic samples grown in the two regions have different isotopic composition too. The different isotopic composition of rainfall between Tyrrhenian and Adriatic coasts is a feature that has already been detected in another work on olive oil (Bontempo et al., 2009). This is principally due to different climatic conditions of the two coasts (differences in average temperature between 1 and 2°C and humidity between 4.8 and 8.5%).

As reported in Fig. 3 and Table 1, Nubia (in Sicily) samples have the highest $\delta^2\text{H}$ values of all the dataset. This is due to the lower altitude (about 5 m ASL) of Sicilian sampling points compared to those in Lazio and Abruzzo, in combination with lower values of latitude and a shorter distance to the sea. On the contrary, Nubia $\delta^{18}\text{O}$ values do not follow the growing trend of the rest of the dataset. It is difficult to explain this behavior as the only source of hydrogen is water while oxygen also derives from O_2 and CO_2 (Barbour, 2007).

Liu et al. carried out a study about the discrimination between garlic samples coming from Argentina and some countries of Asia (Liu et al., 2018). Hydrogen and oxygen isotopic ratios were fundamental to discriminate between Argentinian and Asian samples. As the first ones had $\delta^2\text{H}$ values ranging between -103 and -86‰ and $\delta^{18}\text{O}$ values ranging between 21.1 and 25.5‰, while the second ones had $\delta^2\text{H}$ values ranging between -70 and -36‰ and $\delta^{18}\text{O}$ values ranging between 25.1 and 32.9‰, the two groups were definitely distinguishable. Both $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values of Italian groups of samples (ranging between -71 and -33‰ for hydrogen and between 27.8 and 36.0‰ for oxygen) are significantly different from the samples from Argentina, however, it is not possible to discriminate them against Asian garlic samples. With regards to the work of Opatić et al., a comparison in $\delta^{18}\text{O}$ cannot be made, as they measured this parameter in garlic water and not in the

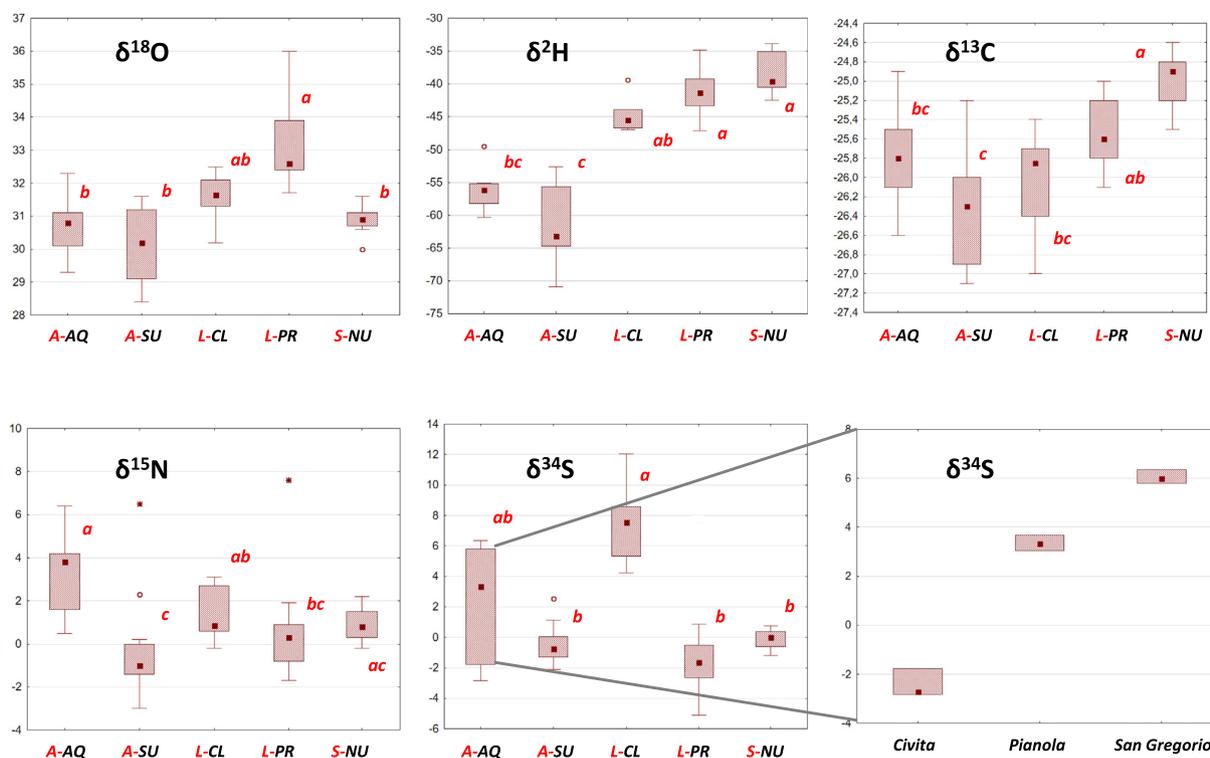


Fig. 3. Box plots of the five parameters under study ($\delta^2\text{H}$, $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$ and $\delta^{34}\text{S}$) for the five groups of the dataset (A-AQ, A-SU, L-CL, L-PR and S-NU) together with a variation of $\delta^{34}\text{S}$ boxplot, in which A-AQ group was split into three components that correspond to three different sampling zones (Civita, Pianola and San Gregorio); significantly different values (Tukey's HSD, $p < 0.05$) between groups are indicated with different letters.

bulk sample.

The growing tendency of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ according to the geoclimatic conditions seems to be the same as $\delta^{13}\text{C}$ (Fig. 3). Indeed, a significant correlation has been noticed between $\delta^2\text{H}$ and both $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ ($p < 0.001$) whereas $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ do not show a significant correlation. Both $\delta^2\text{H}$ and $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ in plants, in fact, depend on factors of influence in the same way (Galimov, 1985).

The main factor affecting $\delta^{15}\text{N}$ in cultivated plants is the fertilization process used. Synthetic fertilizers, produced from atmospheric nitrogen via the Haber process, have $\delta^{15}\text{N}$ values between -4 and $+4\text{‰}$, while organic fertilizers are characterized by values between $+0.6$ and $+36.7\text{‰}$ (Bateman et al., 2007; Otero et al., 2004). $\delta^{15}\text{N}$ values of our Italian samples range from -3.0‰ to $+7.6\text{‰}$. As we can see in Fig. 3 and Table 1, the L'Aquila (A-AQ) group samples are characterized by higher $\delta^{15}\text{N}$ values that can be attributed to the use of organic fertilizers in garlic cultivations. To explain lower or even negative values, we can consider both the possible use of synthetic fertilizers or that of organic fertilizers deriving from nitrogen-fixing plants (Yoneyama and others, 2005). Moreover, it is interesting to notice how L'Aquila (A-AQ) and Sulmona (A-SU) values vary, despite both coming from the Abruzzo region.

The sulphur isotopic ratio is influenced by many different factors, such as abundance of sulphides in soil, plants' aerobic and anaerobic growth, local bedrocks (Rubenstein and Hobson, 2004), active microbial process in the soil, fertilization procedures, active deposition (Krouse et al., 2000). $\delta^{34}\text{S}$ values of our Italian samples range from -5.1‰ to $+12.0\text{‰}$ (Table 1). As we can see in $\delta^{34}\text{S}$ boxplot (Fig. 3), there is an interesting difference between Castelliri (L-CL) samples and the other dataset. This may be due to geological features which differentiate this sampling point from the others, or even to particular cultivating practices which involve the use of sulphur fertilizers (Clark, 2013; Krouse et al., 1991). Moreover, since the L'Aquila group (A-AQ) shows a wider variability of $\delta^{34}\text{S}$ values, we decided to look into this aspect further. Thus, displaying the A-AQ data according to its three

main sampling zones (Civita, Pianola and San Gregorio), we observed that $\delta^{34}\text{S}$ noticeably discriminates the garlic samples (Fig. 3). The complexity of the components that contribute to determining the sulphur isotopic ratio is huge, and the zone under consideration is quite particular, as L'Aquila is a seismic territory, characterized by several seismic points, thermal sulphuric sources and some important faults, that may contribute to defining a particular $\delta^{34}\text{S}$ landscape.

Liu et al. reported a higher isotopic range for $\delta^{15}\text{N}$ (between 4.4 and 9.8‰) for samples from Argentina probably due to the use of non-chemical fertilizers (Liu et al., 2018). On the contrary, the range reported for Asia (between -2.4 and $+8.4\text{‰}$) overlaps that found in the Italian garlic data (between -3.0 and $+7.6\text{‰}$), probably indicating a mixed use of organic and chemical fertilizers.

The results found by Opatić et al. (2017) are perfectly analogous to ours for $\delta^{34}\text{S}$ (Slovenian values between -4.0 and 7.1‰), while nitrogen isotopic ratios are significantly higher (Slovenian garlic values range between 0.1 and 12.3‰ with a mean value of $6.5 \pm 2.6\text{‰}$). This is due to the use of wholly organic fertilizers in Slovenian cultivars, as it is stated in Opatić's article.

3.2. Multivariate statistical analysis and validation

The multivariate statistical analysis carried out on the stable isotope data matrix consisted in a preliminary PCA exploration of the whole dataset and in a successive LDA and k-NN classification limited to red garlic varieties coming from A-SU, L-PR, L-CL and S-NU. Each of these four categories can be assumed as homogeneous on the basis of the geographical origin of the garlic samples that in each investigated site are obtained by propagation of local bulb varieties. Class modelling by SIMCA was finally attempted.

The results of the PCA carried out on the autoscaled variables, in order to emphasize the role of quantities with low variance, are summarised in Fig. 4, which displays the objects and the variables simultaneously projected in the space of the first four PCs explaining as a

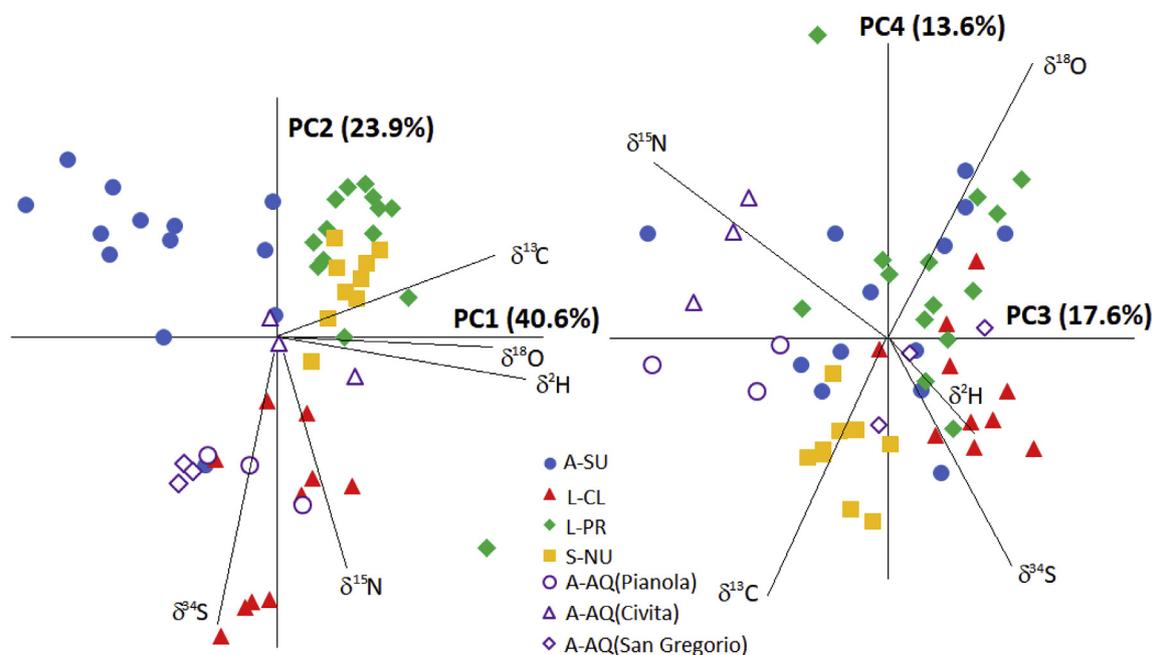


Fig. 4. PCA biplot showing the garlic samples and the stable isotope ratios projected in the space of the significant PCs extracted from the autoscaled data.

whole 95.7% of the total variance. Although no information on the garlic origin is imported into the PCA, the biplots reveal a clear clustering of the samples coming from the same territory. Garlic produced in S-NU and L-PR areas, in particular, give rise to compact clusters with the exception of a single isolated L-PR sample. These two groups partially overlap in the PC1-PC2 plane, but are perfectly separated in the PC3-PC4 plane. A-SU and L-CL garlic groups are a bit more spread-out than S-NU and L-PR yet are still separated in the PC1-PC2 plane. The observed greater variability of A-SU and L-CL samples as compared with the S-NU and L-PR ones likely reflects the extension of the territories where garlic is cultivated. Another interesting thing to notice is that PCA can distinguish the samples coming from the same region. In fact, further inspection of the sample distribution within the PC1-PC2 plane reveals the almost complete separation of A-SU samples from the A-AQ ones (coming from the same region, Abruzzo). Moreover, the A-AQ garlic samples, although cultivated in very close territories (San Gregorio, Civita and Pianola), are also distinctly separated in the PC1-PC2 plane and such separation is confirmed in the PC3-PC4 plane.

A projection of the variable loadings on the significant PCs allows the observed separation of garlic categories in terms of variations in the stable isotope ratios to be explained. Separation of A-SU garlic samples from S-NU and L-PR categories, in particular, occurs in a direction almost parallel to PC1 and $\delta^{13}\text{C}$, $\delta^{18}\text{O}$ and $\delta^2\text{H}$ seem to be the discriminant variables responsible for such separation. On the other hand, $\delta^{34}\text{S}$ and $\delta^{15}\text{N}$, being collinear with PC2, account for the isolation of L-CL garlic with respect to all the other categories. $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$, being collinear with the direction along which S-NU and L-PR garlic groups are separated in the PC3-PC4 plane, seem potentially useful for the differentiation of these two categories. As for the intra-region differentiation of L-SU garlic from A-AQ samples which occurs along the bisecants of PC1 and PC2 axes, all the measured stable isotope ratios seem to be involved.

LDA was performed on garlic samples cultivated in A-SU, L-CL, L-PR and S-NU sites to attempt a classification based on geographical origin. On the other hand, garlic samples coming from L'Aquila are too heterogeneous to be considered as a group in LDA. One L-PR sample, that in PCA appeared as an isolated spot, was removed from the data set. LDA classification was finally performed on 44 samples, namely 13 A-SU, 10 L-CL, 12 L-PR and 9 S-NU. A strict validation of LDA classification performance was carried out on an external data set designed by

duplex Kennard-Stone data splitting. Approximately 30% of the available objects, namely 13 data samples, were included in the prediction set; four of these were from A-SU, three from L-CL, three from L-PR and three from A-NU. A forward stepwise procedure based on the Wilk's lambda (W), an inverse index of the separation goodness, was applied to avoid including in LDA the variables with poor discriminant power. The variable with the highest W was selected first, while those producing the maximum W decrease were selected in the successive steps. Their significance p was determined by F statistics. Following this procedure, $\delta^{34}\text{S}$ ($W = 0.1210$, $p < 10^{-4}$), $\delta^2\text{H}$ ($W = 0.0243$, $p < 10^{-4}$) and $\delta^{18}\text{O}$ ($W = 0.0119$, $p = 0.0004$) were sequentially incorporated in the LDA model. The next candidate, $\delta^{13}\text{C}$, produced a non-significant W decrease ($W = 0.0108$, $p = 0.5035$), but it was anyway included, as its removal would cause a deterioration of the classification performance evaluated by leave-one-out cross-validation. Finally, only $\delta^{15}\text{N}$ ($W = 0.0107$, $p = 0.9711$) was excluded. The results of LDA are summarised in Fig. 5 and Table 2, displaying the canonical plot and the rate of correct classification of calibration/prediction samples, respectively. Fig. 5 displays the garlic samples projected in the plane of the first two discriminant functions (DFs), or canonical variables, representing uncorrelated directions of best class separation, ordered according to their discriminant power. Analogously to PCA biplots, the variable loadings can be reported together with the data samples, which gives a measure of their relative importance. The L-CL group appears in Fig. 5 as a cluster, distinctly separated from the others along DF1. As for S-NU, a partial overlapping with L-PR samples can be observed in the plane DF1 – DF2, however, these two classes are distinctly separated along DF3 (not shown in Fig. 5). A-SU and L-PR samples, on the other hand, are almost superimposed in Fig. 5. According to the reciprocal location of garlic classes in the canonical plot, most of the observed misclassifications in calibration and cross-validation arise from wrong attribution of some A-SU samples to the L-PR class and of L-PR samples to the S-NU class. Nevertheless, a single sample of the prediction set coming from A-SU was erroneously attributed to the L-PR class and the misclassified sample was very close to the delimiter between these two categories. As a whole, the LDA results suggest a clear isolation of L-CL and S-NU garlic categories, being both characterised by 100% sensitivity (ability to accept its own objects) in calibration and prediction. However, while the L-CL class also exhibits 100% specificity (ability to reject extraneous objects), some L-PR

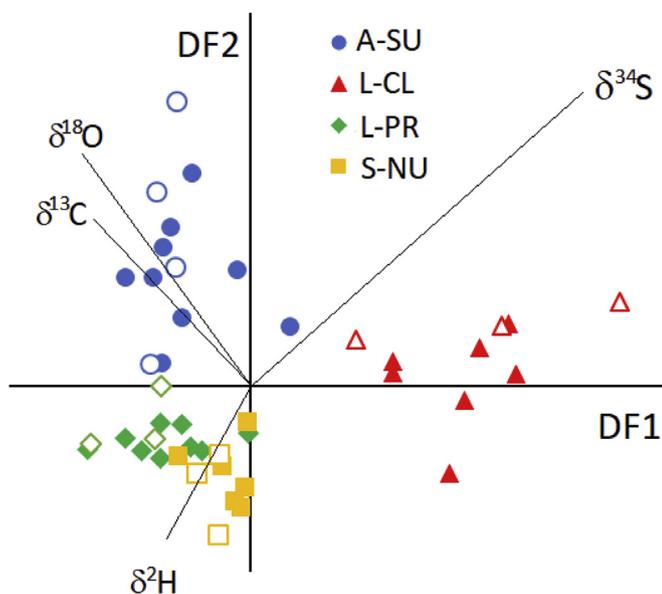


Fig. 5. Projection of the LDA classes on the plane of the first two discriminant functions (DFs). Open symbols represent the external garlic samples.

Table 2

Correct assignments (%) of garlic samples in calibration (calculated classes) and external prediction (predicted classes) provided by LDA and k-NN classification methods. Sensitivity (%) and specificity (%) observed in SIMCA class modelling.

	calculated classes				predicted classes			
	L-CL	A-SU	S-NU	L-PR	L-CL	A-SU	S-NU	L-PR
LDA	100	92	100	89	100	75	100	100
k-NN	86	100	83	89	100	100	100	100
SIMCA	sensitivity				specificity			
		calc	pred		to L-CL	to A-SU	to S-NU	to L-PR
	L-CL	100	33	–	100	100	100	100
	A-SU	100	100	100	–	100	75	–
	S-NU	100	67	100	100	–	100	–
L-PR	100	67	100	100	100	78	–	

samples were erroneously assigned to the S-NU category in LDA calibration.

Apart from LDA, a geographical classification of the garlic samples was performed by k-NN, a non-parametric method in which similarities between objects are evaluated by distance functions of a preselected number k of closest neighbours. In this work, usual Euclidean metrics was used to compute the distances from the autoscaled data and the optimal k value ($k = 3$) was determined by cross-validation. Descriptive and predictive performances of k-NN classification of the garlic samples according to their geographical origin are collected in Table 2. These data reveal a relatively good classification of the training data, since all the A-SU garlic samples were correctly assigned and only a single sample of each of the other three classes was misclassified. Furthermore, all the external garlic samples were correctly classified by k-NN.

In SIMCA class-modelling, unlike LDA and k-NN classification methods, each class is modelled separately and assignment of an external sample to a given class is evaluated independently from the other classes. This permits to establish if a sample which is supposed to belong to class A it is really part of that class or not, which is the typical question in the control of food genuineness and authentication (Forina et al., 2008; Rodionova et al., 2016). Application of class-modelling techniques implies that an object could be either accepted by a specific class or accepted by more classes or rejected by any class in the problem. In spite of the relatively small number of garlic samples representing some geographical areas, SIMCA class modelling was applied here to explore the real potentiality of stable isotope ratio analysis

for the authentication of garlic ecotypes. The optimal number of PCs representing the four garlic classes and the kind of data scaling were determined by leave-one-out cross validation. Finally, a different number of PCs extracted from the centered variables was used to model the L-CL (two PCs), A-SU (three PCs), S-NU (two PCs) and L-PR (one PC) garlic classes. The results of SIMCA class-modelling expressed in terms of sensitivity (the percentage of objects correctly assigned to a given class) and specificity (the percentage of objects extraneous to a given class which are correctly refused by the class model) are displayed in Table 2. It can be observed that all the objects of the calibration set were correctly assigned to the expected classes. The sensitivity in external prediction is still quantitative for the A-SU class or acceptable for both L-PR and S-NU classes (67%), but it is considerably lower for the L-CL category (33%). However it must be reminded that because of the small number of garlic samples in the external set (three or four for each class), misclassification of few samples has a great impact on the classification percentage rate. This is the case of L-CL, for which two of the three external samples fall outside of the class model. Focusing on specificity, the class models developed for L-CL and S-NU garlic varieties refuse 100% of garlic samples coming from all the other territories. On the other hand, only a moderate decrease of specificity is observed for the class model of A-SU, that accepts three L-PR samples, and for the class model of L-PR that accepts two S-NU samples.

4. Conclusions

The usefulness of stable isotope ratios of bioelements in detecting the geographical origin of red garlic from Italy, an area that has never been taken under evaluation in this sector, was confirmed. This paper presents an unprecedented study in which stable isotope ratios analysis of the five bioelements ($\delta^2\text{H}$, $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$ and $\delta^{34}\text{S}$) has been applied for the first time all together to discriminate between bulk red garlic samples coming from Abruzzo, Lazio and Sicily. In spite of the close distance of some of the production areas investigated in this work, both linear discriminant analysis and k-NN method were able to classify almost all the garlic samples on the basis of their origin. Although the two classification strategies provide comparable performances, k-NN may be preferred to linear discriminant analysis because it does not require any preliminary assumptions about the shape of the decision boundary. Nevertheless, the latter approach can provide useful information on the role of the individual bioelements in the discrimination of garlic varieties cultivated in different territories, which is not easily evident in k-NN classification.

Preliminary class modelling performed by SIMCA confirms that stable isotope ratios of the bioelements, apart from having a good discriminant ability, are also useful variables to build class models for typical garlic varieties cultivated in specific territories and assess the genuineness of the garlic samples relating to their geographical origin. The SIMCA models developed for each of the red garlic varieties cultivated in four different Italian sites exhibit an acceptable sensitivity and good specificity. It follows that SIMCA class modelling is the recommended statistical treatment for the authenticity verification of garlic geographical origin based on the stable isotope ratios of the bioelements. These first results are encouraging but further investigations, involving a greater number of samples and other parameters to be combined with isotopic ones, should be addressed to improve the method.

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.

References

- Araguas, L.A., Araguas Araguas, L., Danesi, P., Froehlich, K., Rozanski, K., 1996. Global monitoring of the isotopic composition of precipitation. *J. Radioanal. Nucl. Chem. Artic.* 205, 189–200.
- Barbour, M.M., 2007. Stable oxygen isotope composition of plant tissue: a review. *Funct. Plant Biol.* 34, 83.
- Bateman, A.S., Kelly, S.D., Woolfe, M., 2007. Nitrogen isotope composition of organically and conventionally grown crops. *J. Agric. Food Chem.* 55, 2664–2670.
- Bontempo, L., Camin, F., Larcher, R., Nicolini, G., Perini, M., Rossmann, A., 2009. Coast and year effect on H, O and C stable isotope ratios of Tyrrhenian and Adriatic Italian olive oils. *Rapid Commun. Mass Spectrom.* 23, 1043–1048.
- Bowen, G.J., Ehleringer, J.R., Chesson, L.A., Stange, E., Cerling, T.E., 2007. Stable isotope ratios of tap water in the contiguous United States. *Water Resour. Res.* 43. <https://doi.org/10.1029/2006wr005186>.
- Camin, F., Perini, M., Colombari, G., Bontempo, L., Versini, G., 2008. Influence of dietary composition on the carbon, nitrogen, oxygen and hydrogen stable isotope ratios of milk. *Rapid Commun. Mass Spectrom.* 22, 1690–1696.
- Camin, F., Wietzerbin, K., Cortes, A.B., Haberhauer, G., Lees, M., Versini, G., 2004. Application of multielement stable isotope ratio analysis to the characterization of French, Italian, and Spanish cheeses. *J. Agric. Food Chem.* 52, 6592–6601.
- Carson, J.F., 1987. Chemistry and biological properties of onions and garlic. *Food Rev. Int.* 3, 71–103.
- Clark, I.D., 2013. Environmental Isotopes in Hydrogeology.
- Crittenden, R.G., Andrew, A.S., LeFourmou, M., Young, M.D., Middleton, H., Stockmann, R., 2007. Determining the geographic origin of milk in Australasia using multi-element stable isotope ratio analysis. *Int. Dairy J.* 17, 421–428.
- D'Archivio, A.A., Foschi, M., Aloia, R., Maggi, M.A., Rossi, L., Ruggieri, F., 2019. Geographical discrimination of red garlic (*Allium sativum* L.) produced in Italy by means of multivariate statistical analysis of ICP-OES data. *Food Chem.* 275, 333–338.
- Forina, M., Lanteri, S., Armanino, C., Casolino, C., Casale, M., Oliveri, P., 2010. Parvus - DIFAR università di Genova. WWW Document. <http://www.difar.unige.it/index.php/it/download-parvus.html>, Accessed date: 18 January 2019.
- Forina, M., Oliveri, P., Lanteri, S., Casale, M., 2008. Class-modeling Techniques, Classic and New, for Old and New Problems. *Chemometrics and Intelligent Laboratory Systems*. <https://doi.org/10.1016/j.chemolab.2008.05.003>.
- Galimov, E.M., 1985. Fractionation of isotopes in biological systems. In: *The Biological Fractionation of Isotopes*, pp. 174–197.
- Härde, W.K., Simar, L., 2015. *Applied Multivariate Statistical Analysis*. Springer.
- Hrbek, V., Rektorisova, M., Chmellarova, H., Ovesna, J., Hajslova, J., 2018. Authenticity assessment of garlic using a metabolomic approach based on high resolution mass spectrometry. *J. Food Compos. Anal.* 67, 19–28.
- Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. *Technometrics* 11, 137.
- Knobbe, N., Vogl, J., Pritzkow, W., Panne, U., Fry, H., Lochotzke, H.M., Preiss-Weigert, A., 2006. C and N stable isotope variation in urine and milk of cattle depending on the diet. *Anal. Bioanal. Chem.* 386, 104–108.
- Krouse, H.R., Grinenko, V.A., 1991. International Council of Scientific Unions. Scientific Committee on Problems of the Environment, Stable Isotopes: Natural and Anthropogenic Sulphur in the Environment. John Wiley & Son Limited.
- Krouse, H.R., Roy Krouse, H., Mayer, B., 2000. Sulphur and oxygen isotopes in sulphate. *Environmental Tracers in Subsurface Hydrology*. pp. 195–231.
- Lanzotti, V., 2006. The analysis of onion and garlic. *J. Chromatogr. A* 1112, 3–22.
- Lerman, J.C., Mook, W.G., Vogel, J.C., 1967. Effect of the tunguska meteor and sunspots on radiocarbon in tree rings. *Nature* 216, 990–991.
- Liu, T.-S., Lin, J.-N., Peng, T.-R., 2018. Discrimination of geographical origin of asian garlic using isotopic and chemical datasets under stepwise principal component analysis. *J. Forensic Sci.* 63, 1366–1373.
- Lu, X., Ross, C.F., Powers, J.R., Aston, D.E., Rasco, B.A., 2011. Determination of total phenolic content and antioxidant activity of garlic (*Allium sativum*) and elephant garlic (*Allium ampeloprasum*) by attenuated total reflectance-Fourier transformed infrared spectroscopy. *J. Agric. Food Chem.* 59, 5215–5221.
- Martins, N., Petropoulos, S., Ferreira, I.C.F.R., 2016. Chemical composition and bioactive compounds of garlic (*Allium sativum* L.) as affected by pre- and post-harvest conditions: a review. *Food Chem.* 211, 41–50.
- Molkentin, J., 2009. Authentication of organic milk using $\delta^{13}\text{C}$ and the α -linolenic acid content of milk fat. *J. Agric. Food Chem.* 57, 785–790.
- Molkentin, J., Giesemann, A., 2007. Differentiation of organically and conventionally produced milk by stable isotope and fatty acid analysis. *Anal. Bioanal. Chem.* 388, 297–305.
- Montaño, A., Beato, V.M., Mansilla, F., Orgaz, F., 2011. Effect of genetic characteristics and environmental factors on organosulfur compounds in garlic (*Allium sativum* L.) grown in Andalusia, Spain. *J. Agric. Food Chem.* 59, 1301–1307.
- Opatić, A.M., Nečemer, M., Kocman, D., Lojen, S., 2017. Geographical origin characterization of slovenian organic garlic using stable isotope and elemental composition analyses. *Acta. Chim. Slov.* 64, 1048–1055.
- Otero, N., Soler, A., Canals, A., 2004. Fertilizer characterization: isotopic data (N, S, O, C, and Sr). *Environ. Sci. Technol.* 38, 3254–3262.
- Quality products registers, 2018. European commission - european commission. WWW Document. https://ec.europa.eu/info/food-farming-fisheries/food-safety-and-quality/certification/quality-labels/quality-products-registers_en, Accessed date: 18 January 2019.
- Rahman, K., 2003. Garlic and aging: new insights into an old remedy. *Ageing Res. Rev.* 2, 39–56.
- Ritota, M., Casciani, L., Han, B.-Z., Cozzolino, S., Leita, L., Sequi, P., Valentini, M., 2012. Traceability of Italian garlic (*Allium sativum* L.) by means of HRMAS-NMR spectroscopy and multivariate data analysis. *Food Chem.* 135, 684–693.
- Rodionova, O.Y., Titova, A.V., Pomerantsev, A.L., 2016. Discriminant analysis is an inappropriate method of authentication. *Trac. Trends Anal. Chem.* <https://doi.org/10.1016/j.trac.2016.01.010>.
- Rubenstein, D.R., Hobson, K.A., 2004. From birds to butterflies: animal movement patterns and stable isotopes. *Trends Ecol. Evol.* 19, 256–263.
- Smith, R.G., 2005. Determination of the country of origin of garlic (*Allium sativum*) using trace metal profiling. *J. Agric. Food Chem.* 53, 4041–4045.
- Snee, R.D., 1977. Validation of regression models: methods and examples. *Technometrics* 19, 415.
- Trirongjitmoah, S., Juengmunkong, Z., Srikulnath, K., Somboon, P., 2015. Classification of garlic cultivars using an electronic nose. *Comput. Electron. Agric.* 113, 148–153.
- Wold, S., Sjöström, M., 1977. SIMCA: a method for analyzing chemical data in terms of similarity and analogy. *ACS (Am. Chem. Soc.) Symp. Ser.* <https://doi.org/10.1021/bk-1977-0052.ch012>.
- Yun, H.-M., Ban, J.O., Park, K.-R., Lee, C.K., Jeong, H.-S., Han, S.B., Hong, J.T., 2014. Potential therapeutic effects of functionally active compounds isolated from garlic. *Pharmacol. Ther.* 142, 183–195.