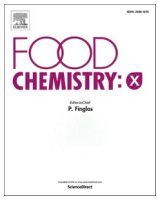
Food Chemistry: X 22 (2024) 101455

Contents lists available at ScienceDirect

Food Chemistry: X

journal homepage: www.sciencedirect.com/journal/food-chemistry-x



Geographical authentication of saffron by chemometrics applied to the ion mobility spectrometry data

Nayereh Shamshiri [a], Reza Fattahi [a], Ahmad Mani-Varnosfaderani [b], Mohsen Barzegar [a,*], Mohammad Ali Sahari [a]

[a] Department of Food Science and Technology, Faculty of Agriculture, Tarbiat Modares University, P. O. Box 14115-336, Tehran, Iran
[b] Department of Chemistry, Faculty of Basic Sciences, Tarbiat Modares University, P. O. Box 14115-175, Tehran, Iran

ARTICLE INFO

Keywords:
Ion mobility spectrometry
Saffron (*Crocus sativus* L.)
Authentication
Multivariate statistical analysis, Geographical origin

ABSTRACT

There is a lack of a reliable tool for quickly determining the geographical origins of saffron (SFR). Ion mobility spectrometry (IMS) has emerged as a promising method for rapid authentication. In this study, 232 Iranian SFR samples harvested in five distinct areas (Khorasan, Azerbaijan, Golestan, Fars, and Isfahan) were analyzed by IMS coupled with chemometric methods. The principal component analysis (PCA) was applied for analyzing the collected IMS data, utilizing three principle components (PCs) that accounted for 81 % of the explained variance. Moreover, the partial least squares-discriminant analysis (PLS-DA) demonstrated the average sensitivity and specificity rates, of 72.3 % to 92.5 % for the exernal test set and 75.5 % to 94.3 % for training set. The accuracy values were $\geq$ 85.0 % for the prediction set for all classes of samples. The results of this study revealed a successful application of IMS and chemometric methods for rapid geographical authentication of saffron samples in Iran.

1. Introduction

Saffron (SFR) in Iran, known as Zaferan, is obtained from the dried stigmas of *Crocus sativus* L. flowers and is used as a spice in food preparations as a coloring or sensing agent (Amirvaresi et al., 2020). Recent studies have also shown its therapeutic properties (Jafari et al., 2022). Due to its low yield rate and high functional value, SFR has become one of the most valuable and expensive spices. As a result, SFR is frequently adulterated, with inferior materials being added to enhance its color properties and the declaration of a different geographical area than the real one, as SFR quality is influenced by climatic and cultivation conditions (Kumari et al., 2021; Biancolillo et al., 2020). SFR has been traditionally cultivated in the the Khorasan (KHN) province of Iran for centuries and more recently in the Azerbaijan (AZN) province as well as other areas in the north, south, and central regions of Iran (provinces of Golestan (GON), Fars (FAS), and Isfahan (ISN)). However, SFR from KHN province is highly esteemed at the national and international levels (Wakefield et al., 2019). In this context, picrocrocin, safranal, and crocetin esters (commonly known as crocin) are considered to be the biological determinants (BDs) responsible for the bitterness, aroma, and color of SFR respectively, and they are used as potential markers for its geographical origin (D'Archivio et al., 2016). BDs fingerprinting is a useful approach for standardizing SFR or detecting mislabeled or fraudulent spices.

Upon obtaining data comprising metabolite signals via analytical techniques, the subsequent task of extracting critical insights from the extensive dataset becomes imperative. Researchers proceed to process this data, employing techniques such as noise filtering, peak detection, peak alignment, and normalization, utilizing either freely available or commercial data processing tools replete with diverse peak processing and annotation capabilities (Zhong et al., 2022). Subsequently, the processed metabolomics data undergoes statistical analysis, encompassing both univariate and multivariate analysis methods (chemometric methods), which play a crucial role in distinguishing and prognosticating characteristics of plant-based foods in metabolomics by identifying trends and biomarkers. Biomarkers, serving as measurable indicators, hold critical importance in discerning metabolic disparities in plant-based foods pertaining to geographical origins, cultivars, or processing methods. Biomarker identification typically entails multivariate statistical analysis, with Principal Component Analysis (PCA), Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA), and Partial Least Squares Discriminant Analysis (PLS-DA) emerging as the

* Corresponding author.
E-mail address: mbb@modares.ac.ir (M. Barzegar).

https://doi.org/10.1016/j.fochx.2024.101455
Received 8 April 2024; Received in revised form 4 May 2024; Accepted 6 May 2024
Available online 10 May 2024

predominant pattern recognition methods in metabolomics. Despite the significance of univariate analysis in identifying metabolites influenced by specific factors, its inability to assess the interrelationships among metabolites from different biological groups necessitates the application of multivariate analysis for comprehensive dataset modeling. Nonetheless, the application of univariate methods to compare two or more groups proves statistically significant, effectively highlighting metabolites significantly affected by the studied factors (Liang et al., 2024).

Chemometric analysis, classified into supervised and unsupervised approaches, addresses datasets arising from complex, multifactorial studies. Unsupervised analysis serves as an initial data exploration phase, succeeded by supervised methods aimed at developing discriminative and classification models, such as data-driven Soft Independent Modeling by Class Analogy (DD-SIMCA), Random Forest (RF), Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA) (Zhong et al., 2022). Researchers commonly discern markers by integrating Variable Importance in Projection (VIP) values from supervised analyses with *p*-values derived from statistical methods (Liang et al., 2024).

In recent years, the geographical origin of SFR has been successfully determined by fingerprinting of BDs present in the saffron matrix using various analytical methods such as HPLC (D'Archivio et al., 2016), GC/MS (Farag et al., 2020), and Fourier transform/near/mid infrared (FT/N/MIR) spectroscopy (Anastasaki et al., 2010), coupled with MSA. Additionally, Wakefield et al. (2019) were able to discriminate between spices of SFR coming from the KHN province of Iran and the La Mancha province of Spain by using MSA of compositional data obtained by ICP/MS. In a similar study, Masi et al. (2016) used BD concentrations to differentiate SFR grown in different production sites of Italy and Iran. UV–Vis spectroscopy of the aqueous extract of SFR also provides useful information regarding its geographical origin (ISO/TS 3632-2, 2010). Stable isotopes and multielemental metabolomics, a fundamental traceability technique, has undergone rapid evolution in recent years. Under the influence of ecological environments of origin (Li et al., 2024), it has been employed to assess plant origin by examining soil and water conditions through isotopic ratios such as carbon ($^{13}C/^{12}C$), nitrogen ($^{15}N/^{14}N$), and oxygen ($^{18}O/^{16}O$). Also, this technique was applied for the authentication of animal meat (Zhao et al., 2023).

The interfering compounds from food matrix probably cause false quantitative results. Today, various reliable sample preparation techniques have been established for cleaning and purification of food samples. For example, dispersive solid-phase extraction (d-SPE) has been used as a simple operating procedure and less solvent consumption method for elimination of impurities or extraction and preconcentration of analytes in real samples (Mao, et al., 2021). However, most of purification methods have some limitations including the use of large amounts of solvents, being time-consuming, and inclusion of tedious operation conditions (Liu et al., 2016). In addition, most of the analytical methods for assessing SFR authenticity are often too time-consuming, needing sample purification, and labor-intensive to be applied routinely (Amirvaresi et al., 2021). To address this issue, ion mobility spectrometry (IMS) could be a convenient alternative for fingerprinting due to its rapidity, sensitivity, simplicity, and low cost. In our recent study, the potential of IMS in SFR characterization was highlighted, with a focus on the "cationic" fingerprints of SFR metabolites (Fattahi et al., 2022).

The combination of IMS and MSA (IMS-MSA) provides analytical technique that is suitable for in-field, real-time detection and analysis. In practice, gaseous metabolites undergo ionization in the reaction chamber, and ions are separated by an electric field at ambient pressure in the IMS drift tube. This separation process not only depends on the charge of cations/anions but also on the shape and size of the metabolites in the carrier gas (García-Nicolás et al., 2020). The applications of IMS-MSA have expanded in recent years, including botanical fingerprinting and safety control analyses in foodomics.

Previous studies (Li et al., 2020) have explored IMS aromatic fingerprinting approaches to assess the authenticity of green tea using multivariate classification methods. The MSA models utilized PCA, PLS-DA, and HCA. In addition, Fattahi et al., 2022 have demonstrated that IMS-based fingerprinting is a reliable method for rapid authenticity testing of agro-foods with high discriminative power.

In this study, we aimed to investigate (i) the influence of some CD-IMS experimental conditions on the separation and relative intensities of the SFR metabolites peaks and developing a rapid, simple, and sensitive analysis method; (ii) the possibility of using volatile BDs to discriminate SFR from different provinces of Iran using CD-IMS and applying variable importance in projection (VIP) to the IMS peak intensities to screen the potential BDs related to SFR origin; (iii) the practical feasibility of the suggested tool by multivariate clustering and classification methods, such as PCA and VIP-PLS-DA, for classifying and predicting the geographical origin of SFR.

2. Experimental

2.1. The SFR samples

A total of 232 SFR spices, all cultivated and harvested in 2021, were obtained from different regions of Iran. Specifically, we collected 90 samples from KHN (north-eastern of Iran), 43 samples from AZN (north-western of Iran), 20 samples from GON (north of Iran), 60 samples from FAS (southern of Iran), and 19 samples from ISN (central of Iran). These spices were directly obtained from local farmers who guaranteed their geographical origin and authenticity. To insure that any potential degradations of biological determinants (BDs) in various spices was avoided, we carefully selected a blend of both old and newly harvested spices, covering a wide range of Iranian SFR.

2.2. Extraction of SFR metabolites for IMS analysis

To extract BDs from the saffron samples, a procedure similar to the one described by Amirvaresi, et al., (2020) was followed, with a few modifications. Firsty, the SFR spices were finely grounded into a powder using ball mills. This ensured that the particle size had a maximum impact on the yield of BDs extraction. To eliminate any heterogeneity among the SFR spices, different spices with the same geographical origins were blended into a single representative spice, capturing the typical characteristics of the origin. After that, 0.02 g of ground SFR was weighed from each origin and placed in a 15 mL centrifuge tube. Then, 10 mL of methanol (MeOH) was added. The samples were subjected to extraction using an ultrasonic bath sonicator (FAPAN, 1200UPS, Iran) for 20 min at a fixed frequency of 20 kHz. During the extraction, the temperature of the water bath was kept below 25 °C. Next, the entire alcoholic extract was transferred to the centrifuge tube and centrifuged (Sigma 3-30KS, Germany) for 15 min at 4500 g. Subsequently, the supernatants were filtered using 0.45 μm acetate cellulose syringe filters. The extraction of the SFR metabolites was performed in ten replicates for each SFR sample to assess the the extraction uncertainty. The filtrated samples were stored in the dark at 4 °C. Before conducting CD-IMS analysis, each sample was homogenized by vortexing for 10 s. Finally, 2 μL of the filtrate was injected into the CD-IMS using a 10 μL Hamilton syringe.

2.3. Spectrometric examination

The saffron metabolites were analyzed using a 300 series CD-IMS apparatus that was manufactured by TOF Tech. Pars Co., Ltd, Isfahan Science and Technology Town (ISTT), Iran. The IMS operated in the positive mode (CD^+-IMS) at an electric field strength of 500 V cm^{-1}, scannig the range of 0–20 ms at a rate of 100 scans per second. The developed CD^+-IMS method was based on a previously reported IMS method for the determination of ochratoxin A with some modifications (Khalesi et al., 2011). In the CD^+ ionization mode, the hydronium ions undergo proton transfer with all biogenic BDs having a proton affinity

higher than H_2O (7.16 eV). The first CD ionization source for IMS was designed by Tabrizchi et al. (2000). Following CD^+ ionization, the Bradbury-Nielsen shutter grid was opened periodically for 40 μs. Cationic ions were then transported to a 10.70 cm drift tube at ambient pressure using a 150 mL min^{-1} air flow at 230 °C. Additionally, a drift gas (900 mL min^{-1} counter flow) was used to collide with the secondary ions of ionized BDs, resulting in the separation of ions. The separated ions were collected at the end of the IMS drift tube (cell) by an aperture grid collector. The ion signal was amplified with a gain of 10^9 V A^{-1} and sent to the data acquisition system through an A/D converter provided by Pico Technology (digital oscilloscope, 2000 series). Under the optimized conditions, various saffron samples exhibited cationic spectra. A program developed using Picoscope 2204A software (Pico Technology Co., UK) was run on a laptop computer to control the device operation and collect spectra. The IMS spectrum consists of signals corresponding to each component (BDs) for the saffron samples against the drift time (DTs). On the other hand, CD-IMS yeilded 2-dimensional (2D) data matrices representing intensity versus DT. For each origin, BD fingerprints were determined in ten replicates and the average spectra were displayed. Multiple injections of saffron samples were performed to assess the instrumental precision. The dataset obtained from Picoscope 2204A software was exported as a TXT file and subsequently imported into MATLAB (release R2017a; The MathWorks Inc., Natick, USA) for further chemometrics analysis.

2.4. The IMS spectral data analysis

2.4.1. Pre-processing of data

The pre-processing of the IM spectral data including autoscaling, alignment, and noise correction was performed based on what reported in a previous study (Garrido-Delgado et al., 2018), with certain modifications. The CD^+-IMS analysis resulted a spectrum where the x-axis represented the DT in milliseconds (ms), and the y-axis represented the intensity (in volts (V)) of each BD. The entire IMS spectra consisted of 1955 variables. After collecting and analyzing all samples, the resulting matrix had dimension of 232 (samples) × 1955 (variables). To construct a reliable predictive model using chemometric analysis, the dimensionality reduction step was employed to enhance the accuracy and interpretability of MSA (Chen et al., 2019a). In this regard, a sub-dataset of the drift times from 4.54 to 7.84 ms was extracted to eliminate interferences from reactant ions (*i.e.,* hydronium (H_3O^+), ammonium (NH_4^+), and nitrosonium (NO^+)) and MeOH. Meanwhile, the acquisition time from 7.84 to 15.28 ms (excluding the BDs signal) was retained for subsequent analysis. Consequently, a matrix with the dimension of 232 (samples) × 1188 (variables) was constructed, containing the relevant information from each spectrum. Additionally, smoothing was performed as an important step prior to MSA to reduce noise amplification in the IMS spectra. To meet this requirement, a Savitzky-Golay filtering tool (window size: 17, order: 2) was applied to maximize the signal-to-noise ratios of all spectra of samples. Following this step, Hotelling's T^2 digits and Q residuals were coded in Matlab 2017a to detect possible outliers (Uarrota ey al., 2014).

2.4.2. Discriminative analysis by PCA

After implementation of pre-preocessing methods, unsupervised pattern recognition through PCA was employed for the preliminary analysis of the spectral data. According to Chen et al., (2019b), PCA is defined as the linear transformation y = Ab, where it transforms the *J*-dimensional original factor b into *N*-dimensional output factor y ($N \ll J$). The factor y captures the most important ions from the original data. The transformation matrix A in PCA ensures that the factors are mutually orthogonal and form new descriptors (PCs), allowing for easier visualization of the data without significant information loss. In this approach, the first PC describes the largest variance, followed by the second PC describing the next largest variance, and so on. To gain a comprehecive understanding of this relationship, only the first 3 PCs were retained. By visualizing the 3D plot of PC1 vs. PC2 vs. PC3, the relative similarity of SFR samples (in the scores plot) can be studied. Hence, this PCA can play an important role in visualizing differeces bwetween saffron samples produced in vbarious growing regions of Iran using the IMS instrument.

2.4.3. The BD selection

In the contex of MSA tools, it is important to acknowledge the potential risk of inappropriate BDs. According to Zhang et al. (2020), the selection of the most significant BDs within the IMS spectrum, specially the region of interest, plays a critical role in developing a reliable PLS-DA model. To identify the most discriminatory regions within the IMS spectrum, BDs with a VIP value greater than "1″ were deemed important. In this case, VIP was used to determine the most relevant ions for SFR classification across different geographical origins. The VIP values for each BD in the dataset (r) are proportional to their accumulated PLS weights from all selected latent variables (LVs). This relationship is expressed by equation (Eq. (1)):

$$VIP_r = \sqrt{\frac{\sum_{j=1}^{T} W_{rj}^2 \times SS_j \times N}{SS_l \times T}} \rangle 1 \quad (1)$$

where SS_j represents the sums of squares of variance for the j^{th} constituent, SS_l represents the total sum of squares given by the dependent BD, W_{rj} denotes the weight value for BD, and the component j is drived from the PLS model. N represents the number of BDs, while T is the total number of LVs.

2.4.4. The PLS-DA models

After explorating the collected data using PCA, supervised PLS-DA was utilized to identify distinct geographical classes of Iranian SFR based on IMS spectra and BD intensities. The PLS-DA carried out with one vs all strategy. As mentioned earlier, the important variables were selected using VIP > 1 before subjecting all BDs to PLS-DA. Subsequently, VIP-PLS-DA was performed to validate the findings of PCA, using the combination of the lowest number of BDs but with highest influence. It is important to note that PCA is an unsupervised analysis and does not define classes in advance. In the PLS-DA model, a rotation occurs on the PC axis to maximize the covariance of X (IMS signal) and Y (SFR origin), and these rotated PCs are referred to as LVs. The purpose of this MSA was to identify the LVs that effectively differentiate between different origin classes based on their spectra (X matrix) and their maximum covariance with a target origin defined in the Y matrix (Bassbasi et al., 2014). The optimal numbers of LVs was determined based on the minimum value obtained through cross-validation, aiming to minimize the classification error.

2.4.5. Evaluation of PLS-DA performance

The evaluation of PLS-DA performance involved determining the percentage of correctly assigned external data samples to their predefined classes. To achieve this, the data was divided randomely into a training set (70 % of the spectral data within each class) and a testing set (30 % of the spectral data within each class) using the Duplex algorithm (Reale et al., 2023). This resulted in 162 samples assigned to the calibration set and 70 samples assigned to the validation set. The Duplex algorithm ensures a uniform distribution of testing objects within the range of training objects. Following the data splitting, the overall performance of classification methods in external validation was quantified using sensitivity (SE, shows true positive rate), specificity (SP, shows true negative rate), and accuracy (AC) percentages. SE, also referred to as the true positive (TP) fraction, represents the proportion of samples correctly classified as belonging to a specific class Y. SP, on the other hand, known as the true negative (TN) fraction, represents the proportion of samples not belonging to class Y that were classified as foreign by the PLS-DA. Additionally, AC represents the proportion of samples

correctly classified as both TP and TN. In all cases, a value closer to one indicates higher accuracy of the MSA. The relevant parameters were calculated using the following equations (2), (3), and (4):

$$SE = \frac{NTP}{NTP + NFN} \tag{2}$$

$$SP = \frac{NTN}{NTN + NFP} \tag{3}$$

$$AC = \frac{NTP + NTN}{NTP + NTN + NFN + NFP} \tag{4}$$

where NTP and NFN represent the number of TPs and the number of false negatives (FNs), respectively. Similarly, NTN and NFP indicate the number of TNs and the number of false positives (FPs), respectively.

2.4.6. Chemometrics software

In the process of chemometric analysis, the following steps were take. Firstly, the data was exported using the Picoscope software version 2000 from Pico Technology Co. Ltd. (UK). Subsequently, MSA techniques including data pre-processing (Savitzky-Golay, Hotelling's T^2, Q residuals, Duplex algorithm, and VIP), clustering (PCA), classification (PLS-DA), and validation (SE, SP, and AC) were carried out using MATLAB R2017a from the Mathworks Inc., Natick, Massachusetts, U.S. (Uarrota et al., 2014).

3. Results and discussion

3.1. Ionization mechanism of CD^+-IMS

Based on the design proposed by Fattahi et al., 2023, the BDs of saffron were identified using CD ionization in positive ion mode. This approach allowed for better identification of saffron metabolites based on the IMS spectra. In the CD^+-IMS signal, there are five peaks centered at 4.88, 5.23, 5.64, 6.14, and 6.76 ms (refer to Fig. 1A). It is important to note that the identification of reactant ions and their assignments were based on the literature (Darzi and Tabrizchi, 2013). Upon careful examination of the spectra in Fig. 1A, the first reactant ion is likely due to ammonia protonation (NH_4^+) around 4.88 ms. The second reactant ion, known as nitrosonium (NO^+), is observed at 5.23 ms and is formed in air through reaction between nitrogen and secondary oxygen emitted from the CD source. Additionally, in Fig. 1A, a third prominent peak at 5.64 ms corresponds to one of the dominant and important protonation ions (hydronium, H_3O^+) of water molecules in the carrier gas. As reported by Darzi and Tabrizchi (2013), ionization in CD^+ begins with high-energy protons released from the H_3O^+ ion, facilitating protonation of an BD with a proton affinity higher than 7.16 eV. Consequently, protons quickly interact with proton-affinitive metabolites in saffron extract, resulting the formation of positive product ions $[BD + H]^+$. Fig. 1A also shows two additional significant peaks at 6.14 and 6.76 ms, which are attributed to the protonation of the MeOH. According to the report by Khalesi et al. (2011), MeOH generates very short-drifted ions close to the reactant ions without interfering with those of BDs. Moreover, MeOH was employed as solvent, since it shows a much smaller proton affinity in comparison with SFR components. This allows the protanation of the BD in the presence of much amounts of the MeOH. Base on previous study (Gazerani et al., 2013), the purification of BDs revealed that the use of MeOH as an extraction solvent for SFR also presents several benefits such as higher recoveries and lower costs while it is more eco-friendliness.

3.2. Detection of BDs in saffron extract by CD^+-IMS

The enhancement of food safety is not inherently achieved through traceability of the origin alone. When the quality of plant-based food is influenced by geographical environment, growing conditions, and cultivation methodologies, metabolomics emerges as a crucial tool for studying the complex molecular composition within a biological matrix. In all reports, both untargeted and targeted metabolomics strategies are employed for qualitative analysis, with certain studies also quantifying and validating results. However, the dynamic nature of plant growth and development, affected by diverse factors, has led to the diversification of metabolic marker analysis and enriched metabolomics investigations (Liang et al., 2024).

The BDs of saffron extract were easily identified by comparing their spectral profiles and DTs with those of reference standards (refer to

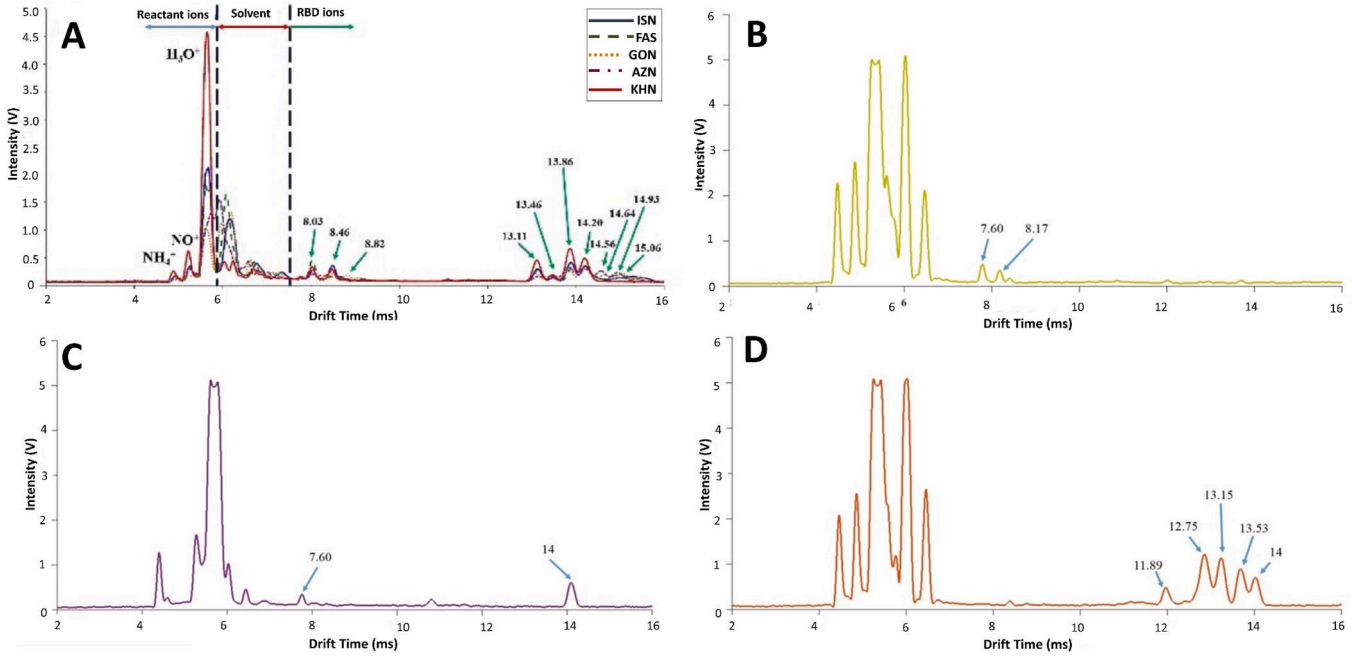


Fig. 1. CD^+-IMS fingerprint spectra of (A) saffron samples originated from Azerbaijan (AZN), Golestan (GON), Khorasan (KHN), Fars (FAS), and Isfahan (ISN) provinces and the pure standards of (B) safranal, (C) picrocrocin, and (D) crocetin esters. CD^+-IMS: corona discharge-ion mobility spectrometry in positive mode.

Fig. 1A-D). From a geographical perspective, it is widely recognized that safranal (aroma), picrocrocin (taste), and crocetin derivatives (color), including crocin, crocetin-di-(β-D-glucosyl)-ester, and crocetin-mono-(β-D-glucosyl)-ester, are closely associaed BDs in saffron. These BDs are influenced by variables such as soil type, climatic conditions, and more (Nescatelli et al., 2017). As depicted in Fig. 1, IMS successfully separated the primary chemical components of saffron. The mean peaks of all BDs are highlighted in Fig. 1B-D. Notably, safranal, and picrocrocin exhibited two distinct peaks at DTs of 7.60 and 8.17 ms, or 7.60 and 14.00 ms, respectively. Within the range of 11.72–14.34 ms, the IMS spectra of crocetin derivatives (cis- and *trans*-polyene dicarboxylic acid sugar esters) displayed five peaks (DTs of 11.89, 12.75, 13.15, 13.53, and 14.00 ms) (Fig. 1D). Previous studies have identified crocetin esters as a result of the thermal decomposition of crocin (Karasu et al., 2019). The thermal sensitivity of crocin can be attribiuted to the light-adsorbing ability and the presence of a long, unsaturated chain in its structure. According to a previously published report (Liao and Liang, 2020), IMS is capable of separating and identifying individual esters, enabling the determination of the geographical origin of samples. Under the optimized instrumental conditions (see section 2.3), the peaks of the main consitituents of the saffron are well-separated. The CD^+-IMS spectra indicate that the drift times ranging from 7.44 to 14.34 ms can be considered as a fingerprint region for saffron. The varying intensities of the main peaks in this region are dependent on the geographical variations of the BDs and have been associated with the primary metabolites of saffron, such as safranal, picrocrocin, and crocetin derivatives. Consequently, it is feasible to determine the geographical origin of saffron based on the spectra of the selected BDs.

3.3. Detection of different saffron origins based on the CD^+-IMS fingerprints

The primary objective of this study was to address concerns related to the rapid authentication of the origin of SFR stigmas cultivated in different provinces. Prior to conducting MSA of IMS spectra, the spectral characteristics of SFR from Iranian provinces were analyzed to identify the functional ions responsible for the samples. As explained in section 3.2, it is expected that different geographical origins would influence the height and position of SFR ions peaks. To visually observe the changes in BDs among stigmas with different cultivated origins, a characteristic fingerprint of each SFR was obtained using only six ions (refer to Fig. 1). This approach allows for easy and rapid characterization of origin. By visually comparing the main ions among the five types of SFR (KHN, AZN, GON, FAS, and ISN), it was possible to individually distinguish the different SFR samples based on their distinct ionization fingerprints in IMS. These changes in spectral patterns and ionization intensities are believed to be a result of geographical variations in the stigmas, as suggested by previous studies (Anastasaki et al., 2010; Tabtabaei et al., 2019). The application of the fingerprint pattern revealed that the ionization fingerprint of SFR extracts slightly shifts the BDs ionization to higher DTs compared to reference standards, potentially due to the presence of SFR tissue interferences. Notably, crocetin derivatives were found to be the most abundant ions (DT from 12.86 to 14.54 ms), exhibiting good discrimination between AZN and GON samples, and partial discrimination between FAS and ISN samples. Additionally, the intensity of crocetin derivatives differed in SFR samples originating from KHN compared to other regions. These crocetin derivatives, also reported by Masi et al., (2016), appear to serve as excellent geographical marker for the authentication Iranian SFR, as they are present in higher quantities in Iranian samples. In AZN and GON samples, several small ions were observed beside the main BDs ions. The ions characterizing the AZN origin ranged from 14.39 to 15.67 ms, while for GON origin, they ranged from 14.47 to 15.85 ms. A distinct ion at 8.82 ms was also observed in GON samples. The ion feature safranal exhibited higher ionization in FAS and ISN samples (DTs at 8.03 and 8.46 ms, respectively), but significantly diminished in AZN and FAS samples (DTs at 8.03 and 8.46 ms, respectively). Similarly, SFR originating from KHN displayed a distinguishable spectrum. Regarding picrocrocin, its IMS peak intensity was relatively similar among SFR samples from Iranian provinces, except for KHN origin (DT at 14.20 ms). Based on these findings, SFRs from different geographical origins exhibited variations in both IMS spectra and biological compositions, potentially influenced by environmental conditions, temperature, and soil type in various regions. Despite these slight spectral differences, the intensity changes of BDs were not consistent, making spectral characterization challenging. Therefore, futher application of MSA is necessary to uncover hidden spectral differences and predict and visualize the authentication origins in SFR samples.

3.4. The MSA of CD^+-IMS data

3.4.1. Discrimination of SFR origins based on PCA

Despite the similarity of SFR origins spectra, there are discernible in the ionization peaks, which can be better observed by utilizing appropriate classification tools such as PCA and VIP-PLS-DA. In this section, PCA was employed in the fingerprint region (refer to section 3.3) to examine the clustering of data points within the multi-block space of intensities. To achieve this, a PCA algorithm was applied to the data sets (232 samples × 1188 variables), and the explained variance by each PC was calculated. The 3 dimensional representation of the first three PCs is presented in Fig. 2A. PCA provides scores that describe the relationships between origins, as well as loadings that describe the correlations between intensities. As can be seen in Fig. 2A, the PCA achieved describing 81 % of the variance of the data for identitying the five saffron origins. PC1 accounted for 51 % of the total variance in the data set, PC2 explained 18 %, while PC3 explained 12 %. The contribution of the remaining PCs was not significant. A study by Song et al. (2021) reported that the PCA algorithm is considered a preferred discrimination model when the cumulative variance contribution rate exceeds 60 %. Observing the multi-block PCA model, the score plot demonestrates a satisfactory discrimination of the AZN (marked by a purple circle) and GON (marked by a pink circle) stigmas, although incomplete, between Iranian SFR stigmas produced in FAS (marked by a yellow circle) and those of ISN (marked by a blue circle). Additionally, a clear discrimination was observed between KHN (marked by a green circle) and the other SFR origins. The distinctness of the KHN stigma cluster in this plot indicates the potential use of this approach in detecting the authentication of KHN with any of the SFRs used in this investigation (refer to Fig. 2A). These results are satisfactory since importers and exporters are mostly concerned with the discriminating pure KHN from other types of SFR origins. In good agreement with the previous section 3.3, the key ions detected with IMS analysis have been confirmed to be useful for the classification of KHN origin. The exploratory PCA suggests that variability within the collected IMS spectra is dominated by the authenticity class of SFR, which masks the possible differences related to SFR. In summary, the PCA results that BDs fingerprinting by IMS approach in combination with MSA represents an effective method for the rapid authentication of different SFR origins, particularly those with subtle differences such as AZN, GON, and KHN origins. To further investigate the differences, the raw IMS ionization spots were analyzed by VIP-PLS-DA to obtain a supervised model.

3.4.2. Selection of main BDs in SFR by using VIP model

For the development of the PLS-DA model, VIP analysis was combined with instrumental spectrometry to ensure the robustness of the PLS-DA model with the highest discrimination potential. This was achieved by selecting the main related spectral ions,which were represented by the raw variables as peak intensities of these ions. Following the embedded procedure described in section 2.4.3, BDs presenting VIP values greater than 1 considered relevant for the model selection. As mentioned in the previous section (3.3), the ions in the range of 2 and 16 ms were consistently detected. Based on the calculated VIP values, a

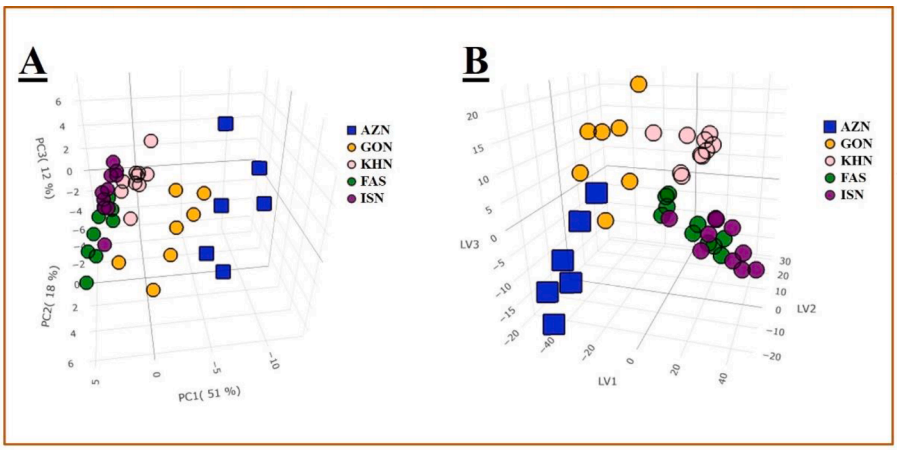


Fig. 2. Discrimination of saffron stigmas from different provinces, Azerbaijan (AZN), Golestan (GON), Khorasan (KHN), Fars (FAS), and Isfahan (ISN) of Iran. (A) PCA scores plot (PC1 vs. PC2 vs. PC3) in the analysis of the IMS spectra of saffron stigmas; (B) PLS-DA scores plot (LV1 vs. LV2 vs. LV2) on the selected (VIP > 1) data set. PCA: principal component analysis; IMS: ion mobility spectrometry; PLS-DA: partial least squares discriminant analysis; LV: latent variables; VIP: variable importance in projection.

graphical representation of the selected spectral ions is presented in Fig. 1A. Upon analyzing the spectra obtained through our MSA approach, the ionization peaks detected for authenticating FAS, ISN, and KHN had DTs of 8.03, 8.46, 13.11, 13.46, 13.86, and 14.20 ms. For the GON origin, the DTs for the best classification model were 8.03, 8.46, 8.82, 13.11, 13.46, 13.86, 14.20, 14.64, and 15.06 ms. The selected ions resulted in DTs of 8.03, 8.46, 13.11, 13.46, 13.86, 14.20, 14.56, and 14.93 ms for SFR originated from AZN province. The interpretation of the IMS spectra presented earlier generally selected BDs between 7.85 and 8.64 ms, which correspond to the ionization of safranal, and between 12.86 and 14.48 ms, primarily dependent on the ionization of picrocrocin and crocetin esters. In a more recent study, Masi et al. (2016) demonesrtated that SFR stigmas from Iran and Italy could be differentiated using metabolites extracted by alcoholic method and analyzed via proton-transfer-reaction time-of-flight (PTR-TOF)-MS/HPLC. They also showed that picrocrocin and safranal were the most useful BDs for describing stigmas authenticity. Overall, the authentication results using a characteristic spectral range of metabolites tend to be better than using the entire region (Senizza et al., 2019). Consequently, only ions at selected DTs were retained, and the selected data matrix (232 × 727) at each point was used as the input of the VIP-PLS-DA model.

3.4.3. *Discrimination of SFR by origin: PLS-DA model using selected ions*

To ensure the reliability of the PCA model, PLS-DA with selected ions was applied for the recognition and classification of the five geographical origins of SFR. A large number of protonized IMS spectra related to SFR origins were acquired and processed, as described in section 3.4.2. A total of 727 spots for each spectrum were extracted and considered for the PLS-DA analysis. The projections of the samples to the first three LVs of the optimized PLS-DA model built using the VIP-selected ions is shown in Fig. 2B. It can be observed that KHN was better discriminated in the PLS-DA plot compared to the PCA plot. The LV3 values for the KHN samples are mostly positive and had a prominent role for discrimination of KHN from other samples. As mentioned previously, the intensities of picrocrocin and total crocetin derivatives in KHN were significantly higher than in comparison to other origins. These compounds can be considered as discriminating BDs for KHN, consistent with a prior study (Moras et al., 2022). However, an overlap between FAS and ISN still existed, and they could be relatively discriminated along with LV1 in the positive area of LV2. This overlap could be attributed to the geographic proximity of the FAS and ISN cultivating regions, as ISN samples were obtained from a nearby region to FAS. In the established PLS-DA plot, a weak overlap was observed between AZN and GON. Nonetheless, the AZN origin was discriminated by the negative part of LV1 and the positive part of LV3, while the GON origin was discriminated by the negative part of LV1 and LV3. Specifically, the results of confusion matrix obtained for the training and test stigmas are presented in Table 1. Once again, the best classification result using PLS-DA was achieved for the samples with KHN origin, where 23 stigmas out of 27 stigmas correctly classified in the test set. It is worth to mention that the highest quality and quantity of Saffron production in Iran belongs to KHN province. Our results reveal that the developed model in this work could correctly identify the origins of samples taken from this province in the prediction set. Furthermore, 11 AZN stigmas (out of 14) and 4 GON stigmas (out of 6) were correctly classified in the test set. However, the FAS class had 11 correctly classified stigmas out of 17, and the ISN class had 4 correctly classified stigmas out of 6 samples in the test set. The discriminatory power of each spectrtal data was again assessed by examining the classification rates (SE, SP, and AC) of the PLS-DA model, as shown in Table 2. Ten-fold cross-validation was performed to validate the PLS-DA model. Analyzing the results for the KHN origin, the AC result of test set provided by PLS-DA was 85.1 %, indicating that the PLS-DA model was robust in distinguishing the samples with KHN origin from the other origins based on the IMS data. The average SE and SP values for the test set were excellent, with values of 72.3 % and 92.5 %, respectively. Similarly, for the samples with AZN origin, the proposed methodology showed good performance for the external test set with SE and SP values of 78.5 % and 95.4 %, respectively. For the GON origin, the SE and SP values for test set were 66.6 % and 90.7 %, respectively. The test set samples with FAS and ISN origins, had the SE and SP values of 64.7 % and 93.3 % for FAS, and 66.6 % and 92.4 % for ISN, respectively. The overall ACs of FAS and ISN origins were 85.4 % and 89.8 %, respectively. It is worth noting that the figures of merit for SP scores were consistently higher than the samples with SE origin, indicating that the PLS-DA model was more effective in rejecting origins that do not truly belong to the designated group. Overall, the performance of the PLS-DA analysis aligned with the initial findings from the PCA exploration and discriminant analysis.

Metabolomics, characterized by the systematic examination of metabolites, serves not only to profile all metabolites in plant-based foods but also to elucidate how these compounds are affected by diverse natural factors. For instance, temperature and light exert significant influence on the expression of genes related to flavonoid biosynthesis, with elevated temperature and increased sunlight correlating with heightened levels of flavonoids. While these findings may be attributed to the long-term effects of climate on the development of local plant varieties, further investigation is warranted to explore the dynamics of

Table 1
The confusion matrices obtained for the training and external test sets using the variable importance in projection (VIP) – partial least squares discriminate analysis (PLS-DA) model.

Real Origins	Calculated Origins (Training Set)					Calculated Origins (Testing Set)				
	AZN	GON	KHN	FAS	ISN	AZN	GON	KHN	FAS	ISN
AZN[a]	24	3	0	2	0	11	2	0	1	0
GON[b]	2	11	1	0	0	1	4	1	0	0
KHN[c]	0	5	57	1	0	0	3	23	1	0
FAS[d]	1	1	2	31	8	1	0	1	11	4
ISN[e]	0	1	1	4	7	0	0	1	1	4

[a]Azerbaijan; [b]Golestan; [c]Khorasan; [d]Fars; [e]Isfahan.

Table 2
Rate (%) of correct authentications provided by VIP-PLS-DA; rate (%) of correctly accepted genuine origins (sensitivity), correctly rejected extraneous origins (specificity), and the recognized origins belonging to the certain origin and correctly rejected origins belonging to different origins (accuracy) for each VIP-PLS-DA class model.

Origins	**Training Set**				
	Sensitivity (%)	Specificity (%)	Accuracy (%)	F1	MCC[f]
AZN[a]	82.7	97.2	94.2	0.857	0.821
GON[b]	78.5	92.2	90.9	0.628	0.594
KHN[c]	90.4	94.8	92.8	0.919	0.855
FAS[d]	72.0	93.3	87.2	0.765	0.680
ISN[e]	53.8	93.8	90.2	0.500	0.447
Average	75.5	94.3	91.0	0.734	0.680
Origins	**External Test Set**				
	Sensitivity (%)	Specificity (%)	Accuracy (%)	F1	MCC
AZN	78.5	95.4	91.3	0.814	0.759
GON	66.6	90.7	88.3	0.533	0.482
KHN	85.1	90.9	88.3	0.867	0.763
FAS	64.7	93.3	85.4	0.709	0.619
ISN	66.6	92.4	89.8	0.571	0.521
Average	72.3	92.5	88.6	0.700	0.629

[a]Azerbaijan; [b]Golestan; [c]Khorasan; [d]Fars; [e]Isfahan; [f]Mathew correlation coefficient.

these factors and their relationship with changes in metabolic pathways and growing conditions. It is also worth noting that samples originating from the same area may exhibit distinct metabolic profiles due to slight geographical variations.

For many years, metabolomics has served as a potent tool to observe phenotypic variations in metabolites of plant-based foods from diverse geographic regions. However, presently, the limitation of geographical traceability lies in the absence of a universal technique applicable across all foods. Existing methods are predominantly indirect and often necessitate combination to increase their accuracy. Hence, efforts are directed towards finding a way to combine universality, efficiency and speed. Looking ahead, future consumers are expected to be increasingly conscientious, desiring comprehensive tracking of all aspects pertaining to food consumption. The challenge will be to develop a traceability tool that combines precision and speed (Liang et al., 2024).

In recent years, research on the geographic traceability of plant-based foods has increasingly focused on how climate data (e.g. precipitation, sunlight intensity, daytime temperature) and geographical conditions (e.g. soil characteristics, topography, and altitude) influence plant metabolism. This interest stems from the realizing that natural factors vary by region, prompting researchers to gather multiple batches of samples across consecutive years to mitigate the impact of such fluctuations. Metabolites can be influenced by regional differences in natural factors, resulting in alterations in biosynthetic pathways due to environmental conditions.

4. Conclusions

It has been concluded that the combination of CD^{+}-IMS with MSA techniques, such as PCA or VIP-PLS-DA, has shown significant potential in authentication and classificaiton of the origins of saffron cultivated in Iran. Through MSA, it has been observed that the IMS fingerprint spectra effectively separate the KHN from other origin. Overall, the use of IMS-MSA has proven to be an efficient, rapid, and good analytical method for ensuring the authenticity of saffron stigmas' provenance. We recommend that further studies should be conducted to explore the application of this spectrometric tool, particularly for a geographical origin authentication.

CRediT authorship contribution statement

Nayereh Shamshiri: Investigation, Formal analysis, Data curation. **Reza Fattahi:** Writing – original draft, Data curation. **Ahmad Mani-Varnosfaderani:** Writing – review & editing, Validation, Software, Data curation. **Mohsen Barzegar:** Writing – review & editing, Supervision, Conceptualization. **Mohammad Ali Sahari:** Writing – review & editing, Data curation.

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.

Acknowledgments

This research forms a part of an M.Sc. program supported by Tarbiat Modares University (Tehran, Iran). The authors express their gratitude to Dr. Vahideh Ilbeigi from TOF Tech. Pars Company, Isfahan University of Technology, Isfahan, Iran, for providing technical assistance in this study.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Amirvaresi, A., Nikounezhad, N., Amirahmadi, M., Daraei, B., & Parastar, H. (2021). Comparison of near-infrared (NIR) and mid-infrared (MIR) spectroscopy based on chemometrics for saffron authentication and adulteration detection. *Food Chemistry, 344*, Article 128647. https://doi.org/10.1016/j.foodchem.2020.128647

Amirvaresi, A., Rashidi, M., Kamyar, M., Amirahmadi, M., Daraei, B., & Parastar, H. (2020). Combining multivariate image analysis with high-performance thin-layer chromatography for development of a reliable tool for saffron authentication and adulteration detection. *Journal of Chromatography A, 1628*, Article 461461. https://doi.org/10.1016/j.chroma.2020.461461

Anastasaki, E., Kanakis, C., Pappas, C., Maggi, L., Del Campo, C. P., Carmona, M., Alonso, G. L., & Polissiou, M. G. (2010). Differentiation of saffron from four countries by mid-infrared spectroscopy and multivariate analysis. *European Food Reseach and Technology, 230*, 571–577.

Bassbasi, M., De Luca, M., Ioele, G., Oussama, A., & Ragno, G. (2014). Prediction of the geographical origin of butters by partial least square discriminant analysis (PLS-DA) applied to infrared spectroscopy (FTIR) data. *Journal of Food Composition and Analysis, 33*, 210–215. https://doi.org/10.1016/j.jfca.2013.11.010

Biancolillo, A., Foschi, M., & D'Archivio, A. A. (2020). Geographical classification of Italian saffron (Crocus sativus L.) by multi-block treatments of UV-Vis and IR spectroscopic data. *Molecules, 25*. https://doi.org/10.3390/molecules25102332

Chen, T., Qi, X., Chen, M., & Chen, B. (2019). Gas chromatography-ion mobility spectrometry detection of odor fingerprint as markers of rapeseed oil refined grade. *Journal of Analytical Methods and Chemistry, 2019*. https://doi.org/10.1155/2019/3163204

Chen, T., Qi, X., Lu, D., & Chen, B. (2019). Gas chromatography-ion mobility spectrometric classification of vegetable oils based on digital image processing. *Journal of Food Measurement and Characterization, 13*, 1973–1979. https://doi.org/10.1007/s11694-019-00116-5

D'Archivio, A. A., Giannitto, A., Maggi, M. A., & Ruggieri, F. (2016). Geographical classification of Italian saffron (Crocus sativus L.) based on chemical constituents determined by high-performance liquid-chromatography and by using linear discriminant analysis. *Food Chemistry, 212*, 110–116. https://doi.org/10.1016/j.foodchem.2016.05.149

Darzi, M., & Tabrizchi, M. (2013). An NO+ reactant ion source for ion mobility spectrometry. *International Journal for Ion Mobility Spectrometry, 16*, 275–280. https://doi.org/10.1007/s12127-013-0137-1

Farag, M. A., Hegazi, N., Dokhalahy, E., & Khattab, A. R. (2020). Chemometrics based GC-MS aroma profiling for revealing freshness, origin and roasting indices in saffron spice and its adulteration. *Food Chemistry, 331*, Article 127358. https://doi.org/10.1016/j.foodchem.2020.127358

Fattahi, R., Mani-varnosfaderani, A., Barzegar, M., & Sahari, M. A. (2022). Rapid metabolites fingerprinting by ion mobility spectrometry : A novel evaluation method for bio-adulteration of saffron (Crocus sativus L.). *Industrial Crops and Products, 188*, Article 115707. https://doi.org/10.1016/j.indcrop.2022.115707

Fattahi, R., Mani-varnosfaderani, A., Barzegar, M., & Sahari, M. A. (2023). An ion mobility spectrometry-chemometrics combination approach for assessing adulteration in saffron (Crocus sativus L.) with synthetic colorants. *Industrial Crops and Products, 193*, Article 116161. https://doi.org/10.1016/j.indcrop.2022.116161

García-Nicolás, M., Arroyo-Manzanares, N., Arce, L., Hernández-Córdoba, M., & Viñas, P. (2020). Headspace gas chromatography coupled to mass spectrometry and ion mobility spectrometry: Classification of virgin olive oils as a study case. *Foods, 9*. https://doi.org/10.3390/foods9091288

Garrido-Delgado, R., Eugenia Muñoz-Pérez, M., & Arce, L. (2018). Detection of adulteration in extra virgin olive oils by using UV-IMS and chemometric analysis. *Food Control, 85*, 292–299. https://doi.org/10.1016/j.foodcont.2017.10.012

Gazerani, S., Mohammadi Sani, A., & Tajalli, F. (2013). Effect of solvent extraction on qualitative parameters of saffron edible extract. *Researche and Reveiws in BioSciences, 7*, 15–27.

ISO. (2010). *International Standard ISO 3632–2: SA (Crocus sativus L.) Test Methods*. Geneva: The International Organization for Standardization.

Jafari, F., Emami, S. A., Javadi, B., Salmasi, Z., Tayarani-Najjaran, M., & Tayarani-Najaran, Z. (2022). Inhibitory effect of saffron, crocin, crocetin, and safranal against adipocyte differentiation in human adipose-derived stem cells. *Journal of Ethnopharmacology, 294*, Article 115340.

Karasu, S., Bayram, Y., Ozkan, K., & Sagdic, O. (2019). Extraction optimization crocin pigments of saffron (Crocus sativus) using response surface methodology and determination stability of crocin microcapsules. *Journal of Food Measurement and Characterization, 13*, 1515–1523. https://doi.org/10.1007/s11694-019-00067-x

Khalesi, M., Sheikh-Zeinoddin, M., & Tabrizchi, M. (2011). Determination of ochratoxin A in licorice root using inverse ion mobility spectrometry. *Talanta, 83*, 988–993. https://doi.org/10.1016/j.talanta.2010.11.004

Li, J., Hua, J., Dong, C., Yang, Y., Deng, Y., Wang, J., Jiang, Y., Yuan, H., & Zhou, Q. (2020). Real-time fingerprinting of the dynamics of green tea volatiles by ion mobility spectrometry for aroma assessment and discrimination. *LWT- Food Science and Technology, 131*, Article 109751. https://doi.org/10.1016/j.lwt.2020.109751

Li, Y., Zhao, W., Qian, M., Wen, Z., Bai, W., Zeng, X., & Dong, H. (2024). Recent advances in the authentication (geographical origins, varieties and aging time) of tangerine peel (*Citri reticulatae* pericarpium): A review. *Food Chemistry, 138531*. https://doi.org/10.1016/j.foodchem.2024.138531

Liang, L., Li, Y., Mao, X., & Wang, Y. (2024). Metabolomics applications for plant-based foods origin tracing, cultivars identification and processing: Feasibility and future aspects. *Food Chemistry, 449*, Article 139227. https://doi.org/10.1016/j.foodchem.2024.139227

Liao, S., & Liang, X. (2020). Rapid measurement of phthalic acid ester environmental hormones using ion mobility spectrometry. *International Journal for Ion Mobility Spectrometry, 23*, 97–103. https://doi.org/10.1007/s12127-020-00268-x

Liu, L., Xia, L., Wu, C., Qu, F., Li, G., Sun, Z., & You, J. (2016). Zirconium (IV)-based metal organic framework (UiO-67) as efficient sorbent in dispersive solid phase extraction of plant growth regulator from fruits coupled with HPLC fluorescence detection. *Talanta, 154*, 23–30. https://doi.org/10.1016/j.talanta.2016.03.038

Mao, X., Xiao, W., Wan, Y., Li, Z., Luo, D., & Yang, H. (2021). Dispersive solid-phase extraction using microporous metal-organic framework UiO-66: Improving the matrix compounds removal for assaying pesticide residues in organic and conventional vegetables. *Food Chemistry, 345*, Article 128807. https://doi.org/10.1016/j.foodchem.2020.128807

Masi, E., Taiti, C., Heimler, D., Vignolini, P., Romani, A., & Mancuso, S. (2016). PTR-TOF-MS and HPLC analysis in the characterization of saffron (Crocus sativus L.) from Italy and Iran. *Food Chemistry, 192*, 75–81. https://doi.org/10.1016/j.foodchem.2015.06.090

Moras, B., Pouchieu, C., Gaudout, D., Anchisi, A., Saupin, X., & Jame, P. (2022). Authentication of Iranian saffron (Crocus sativus) using stable isotopes δ13C and δ2H and metabolites quantification. *Molecules, 27*(20), 6801. https://doi.org/10.3390/molecules27206801

Nescatelli, R., Carradori, S., Marini, F., Caponigro, V., Bucci, R., De Monte, C., Mollica, A., Mannina, L., Ceruso, M., Supuran, C. T., & Secci, D. (2017). Geographical characterization by MAE-HPLC and NIR methodologies and carbonic anhydrase inhibition of saffron components. *Food Chemistry, 221*, 855–863. https://doi.org/10.1016/j.foodchem.2016.11.086

Reale, S., Biancolillo, A., Foschi, M., Di Donato, F., Di Censo, E., & D'Archivio, A. A. (2023). Geographical discrimination of Italian carrot (Daucus carota L.) varieties: A comparison between ATR FT-IR fingerprinting and HS-SPME/GC-MS volatile profiling. *Food Control, 146*, Article 109508. https://doi.org/10.1016/j.foodcont.2022.109508

Senizza, B., Rocchetti, G., Ghisoni, S., Busconi, M., De Los Mozos Pascual, M., Fernandez, J. A., Lucini, L., & Trevisan, M. (2019). Identification of phenolic markers for saffron authenticity and origin: An untargeted metabolomics approach. *Food Research International, 126*, Article 108584. https://doi.org/10.1016/j.foodres.2019.108584

Song, J., Shao, Y., Yan, Y., Li, X., Peng, J., & Guo, L. (2021). Characterization of volatile profiles of three colored quinoas based on GC-IMS and PCA. *LWT- Food Science and Technology, 146*, Article 111292. https://doi.org/10.1016/j.lwt.2021.111292

Tabrizchi, M., Khayamian, T., & Taj, N. (2000). Design and optimization of a corona discharge ionization source for ion mobility spectrometry. *Review of Scientific Instruments, 71*, 2321–2328. https://doi.org/10.1063/1.1150618

Tabtabaei, S., D'Archivio, A. A., Maggi, M. A., Brutus, M., Bajracharya, D. H., Konakbayeva, D., Soleimani, A., Brim, H., & Ashktorab, H. (2019). Geographical classification of Iranian and Italian saffron sources based on HPLC analysis and UV-Vis spectra of aqueous extracts. *European Food Research and Technology, 245*, 2435–2446. https://doi.org/10.1007/s00217-019-03352-7

Uarrota, V. G., Moresco, R., Coelho, B., Nunes, E. D. C., Peruch, L. A. M., Neubert, E. D. O., Rocha, M., & Maraschin, M. (2014). Metabolomics combined with chemometric tools (PCA, HCA, PLS-DA and SVM) for screening cassava (Manihot esculenta Crantz) roots during postharvest physiological deterioration. *Food Chemistry, 161*, 67–78. https://doi.org/10.1016/j.foodchem.2014.03.110

Wakefield, J., McComb, K., Ehtesham, E., Van Hale, R., Barr, D., Hoogewerff, J., & Frew, R. (2019). Chemical profiling of saffron for authentication of origin. *Food Control, 106*. https://doi.org/10.1016/j.foodcont.2019.06.025

Zhang, X., Dai, Z., Fan, X., Liu, M., Ma, J., Shang, W., Liu, J., Strappe, P., Blanchard, C., & Zhou, Z. (2020). A study on volatile metabolites screening by HS-SPME-GC-MS and HS-GC-IMS for discrimination and characterization of white and yellowed rice. *Cereal Chemistry, 97*, 496–504. https://doi.org/10.1002/cche.10264

Zhao, L., Zhang, H., Huang, F., Liu, H., Wang, T., & Zhang, C. (2023). Authenticating Tibetan pork in China by tracing the species and geographical features based on stable isotopic and multi-elemental fingerprints. *Food Control, 145*, Article 109411. https://doi.org/10.1016/j.foodcont.2022.109411

Zhong, P., Wei, X., Li, X., Wei, X., Wu, S., Huang, W., & Lei, H. (2022). Untargeted metabolomics by liquid chromatography-mass spectrometry for food authentication: A review. *Comprehensive Reviews in Food Science and Food Safety, 21*(3), 2455–2488. https://doi.org/10.1111/1541-4337.12938