




# Use of an “Intelligent Knife” (iknife), Based on the Rapid Evaporative Ionization Mass Spectrometry Technology, for Authenticity Assessment of Pistachio Samples

Francesca Rigano<sup>1</sup> · Sara Stead<sup>2</sup> · Domenica Mangraviti<sup>3</sup> · Renata Jandova<sup>2</sup> · Davy Petit<sup>4</sup> · Nino Marino<sup>5</sup> · Luigi Mondello<sup>1,3,6</sup> 

Received: 12 October 2018 / Accepted: 16 October 2018 / Published online: 19 November 2018

© The Author(s) 2018, corrected publication 2018

## Abstract

The iknife (intelligent knife) is the coupling between rapid evaporative ionization mass spectrometry and an electrosurgical knife. The technique is based on a fast evaporation of a sample, yielding gaseous molecular ions of the major components, thus obtaining a holistic profile, usable as univocal fingerprinting. Since its introduction by Professor Takats of Imperial College of London, it was mainly used for in vivo biological tissues analysis since, differently from desorption ionization methods, it does not require any sample preparation. The aim of the present research is to extend the applicability and the advantages of such a technique to food samples for geography evaluation and authenticity assessment. For this purpose, a database containing MS profiles of authentic samples need to be created. The application here described focus on the profiling of pistachio samples, taking into account that adulterated Sicilian Bronte pistachio can be found on the market, due to its high cost.

**Keywords** Intelligent knife (iknife) · Rapid evaporative ionization mass spectrometry (REIMS) · Authenticity assessment · Geographical traceability · Pistachio · Statistical clustering

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s12161-018-1386-8>) contains supplementary material, which is available to authorized users.

✉ Luigi Mondello  
lmondello@unime.it

<sup>1</sup> Chromaleont S.r.L c/o Dipartimento di Scienze Chimiche, Biologiche, Farmaceutiche ed Ambientali, Polo Annunziata, University of Messina, Viale Annunziata, 98168 Messina, Italy

<sup>2</sup> Waters Corporation, Stamford Avenue, Altrincham Road, Wilmslow SK9 4AX, UK

<sup>3</sup> Dipartimento di Scienze Chimiche, Biologiche, Farmaceutiche ed Ambientali, Polo Annunziata, University of Messina, Viale Annunziata, 98168 Messina, Italy

<sup>4</sup> Waters Corporation, Waters S.A.S, BP 608 En Yvelines Cedex, 78056 Saint-Quentin, France

<sup>5</sup> Pisti – Antichi Saperi dell'Etna S.r.L Viale J.F. Kennedy/Zona artigianale, 12-13 95034, Bronte (CT) Catania, Italy

<sup>6</sup> University Campus Bio-Medico of Rome, Via Álvaro del Portillo 21, 00128 Rome, Italy

## Introduction

The topic of food authenticity and genuineness centralizes special interest for reason related to both economical and sociopolitical factors. From an economical point of view, it is essential to preserve the national producer, but at the same time, to create a competitive, sustainable, and innovative global market. On the other hand, the first important social aspect is the preservation of the consumers, by the provision of safe, nutritious, and high-quality food. Despite the food labeling (Regulation EU 1169/ 2011), which aim to inform consumers what kind of products they are buying and facilitate trade inside and outside Europe, the broad question of “food fraud” is not regulated by a common EU legislation. Such a mechanism often lead to the occurrence of a certain number of cross-border food law violation due to the difficulty for member states’ competent authorities to communicate efficiently with their counterparts in other member states. In this framework, the so-known “General Food Law” (Regulation EC 178/ 2002) set out an overarching and coherent framework for the development of food and feed legislation both at union and

national levels. In particular, the European Commission decided to activate a dedicated network of administrative assistance liaison bodies that would handle specific requests for cross-border cooperation. The dedicated liaison bodies are referred to as “Food Fraud Contact Points” (FFCP). The group of FFCPs is collectively referred to as the “Food Fraud Network” or FFN and aim to activate competent authorities against “food crime,” promoting investigation/prosecution.

Within the network system, the EU Commission Joint Research Centre (JRC) ensures the maximum level of expertise and scientific knowledge in order to apply the best available science. In Italy, the JRC based in Ispra devoted special attention to the topic of food safety and security, aiming to preserve the vast national heritage of typical and unique foods, viz. the PDO, PGI, and DOCG (Protected Designation of Origin, Protected Geographical Indication, and Denomination of Controlled and Guaranteed Origin) trademarks. This initiative calls for the development of advanced and innovative analytical technologies and methods, characterized by high sensitivity, selectivity, throughput, and feasibility. During the last decades, the exponential progresses in capillary gas chromatography (GC) techniques and mass spectrometry (MS) instrumentation made GC and GC-MS the most used tool to detect adulteration or any kind of contamination (plastic compounds, pesticides) in food (Cubero-Leon et al. 2014). Afterward, in the last years, improvements in liquid chromatography (LC) columns and instrumentation and progresses in high resolution MS (HR-MS) incentivized the rapid development of metabolomics mainly based on the analysis of LC-amenable polar and/or large molecule (Wolfender et al. 2013) as food authenticity

markers (polyphenols, amino acids, carotenoids, lipid species). For sake of clarity, the main analytical techniques applied in this field are reported in Table 1 along with their benefits and drawbacks.

Two complementary metabolomics approach can be clearly identified: the first is based on the identification/quantification of one or more group of analytes usually by LC-MS techniques; the second one aim not to identify/quantify each observed molecular species, but to compare a fingerprint generated by the analysis of the whole metabolome. Shotgun MS, vibrational spectroscopy (IR, near-IR (NIR), Raman, and FT Raman), and nuclear magnetic resonance (NMR) fall into the second approach, and take great advantage of very short analysis time because of the absence of a chromatographic separation step and, sometimes, of complex sample preparation procedures. Furthermore, MS, especially HR-MS, being characterized by a high identification power could be employed for the characterization of marker species.

Within this context, a very innovative analytical technique, namely *iknife*, based on a shotgun MS approach could be a valid aid to obtain a fingerprint of different food products. The *iknife*, that means “intelligent knife,” is the term coined in 2013 by Prof. Takats et co-workers (Balog et al. 2013) to describe the coupling between the electrosurgical tool (the knife) and an MS system based on rapid evaporative ionization mechanism (REIMS, rapid evaporative ionization mass spectrometry). Such a technique demonstrated significant potentiality in the clinical medical field, where a certain number of paper about cancer tissue identification has been published (Neidert and Bozinov 2013; Balog et al. 2013; Balog et al. 2015; Alexander et al. 2015; St John et al. 2017), by exploiting the near 100% accuracy in the *in vivo* recognition of cancer cells, in real time (3–4 s analysis time) without

**Table 1** Most employed analytical technique in the field of food authenticity and adulteration

Analytical technique	Benefits	Drawbacks
GC-based approaches	Very high sensitivity High identification power	Limited to volatile and semi-volatile components Need of derivatization or extraction procedures
LC-MS	Suitable for very large compounds (biomolecules) High sensitivity	Poor identification power compared to GC-EI-MS-based techniques
Direct infusion MS	Very short analysis time	Low sensitivity for trace compounds Ion suppression/enhancement effects Unreliable quantification
Vibrational spectroscopy	Rapid and non-destructive analysis Very short analysis time	Not suitable for high moisture samples
NMR	High structural characterization power Very short analysis time	Low sensitivity High maintenance cost
Bulk-IRMS	Direct origin/isotope ratio correlation	High probability of artifacts due to external contamination Difficulty to detect adulteration at trace levels
GC-IRMS (Isotopic Ratio Mass Spectrometry)	Direct origin/isotope ratio correlation	Limited to volatile and semi-volatile components Need of pure compounds (baseline separation, no co-elution)
Sensory analysis	Low cost	Limited to a restricted number of molecules
Physicochemical characterization	Easy to use	Need of complementary analytical techniques

any sample preparation procedure. On the other hand, for the best of author knowledge, only three papers refer to the use of the iknife or REIMS technology for the analysis of food and all focus on animal tissues to discriminate different kind of meat (Balog et al. 2016) or fish (Black et al. 2017) or identify boar taint in pig fat (Verplanken et al. 2017). Certainly, Italy can be the ideal location for the extension of such analytical technique in the food area, taking into account the significant number of PDO, PGI, and DOP trademarks present in the territory. In this work, the novel iknife technique was exploited to obtain a clustering of various pistachio samples coming from different geographical origins. In particular, the main object was the preservation of the Bronte pistachio, a PDO trademark (OJEU 2009/C 130/09) that represents one of the greatest economical source of the Sicily east area; it is cultivated around the Etna volcano in the area of Bronte, where the lava land and climate allow the production of nuts with intense green color and aromatic taste, very appreciated in international markets. Due to the relatively small dimension of the cultivated area, the Italian production is very low compared to Asia and California; nevertheless, the higher quality is recognized in the world. A further purpose of the present work was to determine marker compounds, responsible of the geographical differentiation. To confirm peak identification, conventional chromatographic techniques (GC-MS, LC-MS) were employed and compared to the novel iknife technique. The obtained results demonstrated the capability of the new technology to provide a highly informative holistic profile of a complex matrix.

## Material and Methods

### Reagents

LC-MS grade acetonitrile, methanol, water, and 2-propanol were purchased from Biosolve Chimie (Dieuze, France). Potassium hydroxide, n-hexane reagent grade, and 2-propanol HPLC grade were purchased from Carlo Erba (Milan, Italy). Ammonium formate was obtained from Alfa Aesar GmbH & Co KG (Karlsruhe, Germany). Standard of trionanoin ( $C_9C_9C_9$ ), triundecanoin ( $C_{11}C_{11}C_{11}$ ), tritridecanoin ( $C_{13}C_{13}C_{13}$ ), tripentadecanoin ( $C_{15}C_{15}C_{15}$ ), triheptadecanoin ( $C_{17}C_{17}C_{17}$ ), trionadecanoin ( $C_{19}C_{19}C_{19}$ ), and a C4-C24 even saturated fatty acid methyl ester (FAME) 1000 mg/L hexane solution was purchased from MilliporeSigma (Bellefonte, PA, USA).

### Samples

Samples of unprocessed whole pistachio nuts from different geographical origins and varieties (Bronte, California prime quality, California second quality, Iran normal, Iran perfect green, Turkey mawardi, Turkey red, Turkey perfect green, Greece) were kindly provided by Pisti S.r.L. (Italy, Bronte,

CT). A minimum of five production lots were collected for each variety.

### Sample Preparation for REIMS Analysis

The pistachio nut outer shell and skin was removed by hand. Equal portions of the different production lots were combined and the nuts were subsequently crushed using a domestic coffee grinder device. A conductive paste was prepared via addition of 200  $\mu$ L deionized water per 1 g of crushed nut material. Approximately, 1 g of the pistachio paste was placed on a wet tissue on top of the return electrode plate prior to iknife sampling.

### REIMS Analysis

The iknife hand-held sampling device (Waters Corporation, Wilmslow, UK) was used to apply a localized high-frequency electric current to the surface of each sample, which instantly vaporizes molecules from the latter. It consisted of a monopolar cutting device with a shortened knife blade approximately 6 mm long and was applied in auto-cut mode in combination with a diathermy electrosurgical generator (Erbe VIO 50 C) (Erbe, Tuebingen, Germany) at a power of 30 W. Sampling was carried out for 3 to 5 s and for each sample, technical replicates were analyzed, thus taking into account repeatability of the analysis. Mass spectrometric analysis was performed using a Xevo G2 XS QTOF instrument equipped with a REIMS source comprising a helical coiled ribbon collision surface heated by a constant current power supply set to 4.5 A and 4.2 V (Kanthal D  $1.0 \times 0.1$  mm) (Waters Corporation, Wilmslow, UK). All analysis was performed in REIMS TOF MS sensitivity mode with continuum data acquisition. A matrix of 2-propanol was infused directly into the REIMS source at a constant flow rate of 200  $\mu$ L  $\text{min}^{-1}$  to promote the ionization of lipid species and maintain source cleanliness. The mass resolution was approximately 20,000 FWHM over the mass range of interest. The cone voltage was set at 100 V. MS analysis was performed in negative ionization mode over a mass range of 50–1200  $m/z$  with an (scan) acquisition time of 1 s/scan. Prior to use, the instrument was calibrated using sodium formate in 2-propanol and was infused via the matrix inlet on the REIMS source.

For quality control purposes, the endogenous matrix ion  $m/z$  255.2330, corresponding to the deprotonated molecule of palmitic acid ( $C_{16}H_{31}O_2$ ) was used for internal lock-mass correction. At least 20 replicate measurements were collected for each sample in two different laboratories (Waters-Wilmslow and Chromaleont-Messina) in order to obtain some inter-laboratory repeatability data and explore the possibility to transfer the statistical model between different geographical locations. Furthermore, replicate burns of a QC sample (porcine liver) were collected at the start and end of analytical

batch. The intensity of the base peak ion at  $m/z$  699.497 was recorded and plotted for quality control monitoring. The knife blade was cleaned using methanol after every 15 samples and the 3-m-long transfer tubing and venturi air pump was cleaned in an ultrasonic bath using methanol at the end of each day.

An untargeted MS analysis was performed to discriminate between pistachio nuts grown in different geographical origins (Turkey, Greece, Iran, Sicily, and USA). In the case of pistachio nuts from Turkey and Iran samples, more than one variety was present. The training samples were analyzed in two different laboratories locations on different days to take into account technical variation originating from intra- and inter-laboratory reproducibility. This step thus generated a total of 360 mass spectra (64 spectra for Sicilian Bronte; 32 for Greek normal; 84 for Iranian normal and perfect green; 64 for USA prime and second quality, and 116 for Turkish mawardi, red and perfect green) that were used to train the database. In addition, samples from a different production lot were collected and analyzed in triplicate and used as an independent validation challenge set.

The mass spectrometer was operated in REIMS MS/MS acquisition mode to generate CID fragments for ions identified as discriminatory features during the multivariate statistical analysis. Precursor ions were isolated in the resolving quadrupole region of the instrument. Argon was used as the collision gas and collision energy was optimized on a compound by compound basis. The following MS/MS specific settings were applied, LM and HR resolution 10 and 14.5, respectively, prefilter 2.0 V and ion energy 0.8 V.

## Chemometric Data Analysis

Multivariate statistical software package LiveID™ (Waters Corporation, Wilmslow, UK) was used as a model builder and recognition tool. In order to generate models from the untargeted profiling REIMS TOF MS data acquired in MassLynx v. 4.1 (Waters Corporation, Wilmslow, UK), the following data treatment steps were performed: lock-mass correction applied using the endogenous matrix ion at  $m/z$  255.2330; all spectra contained within each “burn event” termed the region of interest (ROI) were combined to form a single continuum spectrum; adaptive background subtraction (ABS) algorithm was applied to reduce the chemical background in the combined spectra; data resampling (binning to 0.1 Da) was performed to reduce the data dimensionality; and the resulting spectrum was normalized using the TIC. For principal component analysis (PCA), the data was centered using the mean value of the entire data set. For linear discriminant analysis (LDA), the data was centered using the mean values of each model class. In either type, the mean for each  $m/z$  bin is subtracted from the values of that bin. Other than normalization and centering, no additional manipulation was performed (for example, scaling).

Following data pretreatment steps, a PCA/LDA model was calculated. Firstly, an unsupervised PCA (singular-value decomposition algorithm) transform is applied to the spectral data calculating the scores and loadings; a supervised LDA transform is then applied to the scores calculated by the PCA transform. LDA is a transform that maximizes the inter-class variance, while minimizing the intra-class variance, resulting in a projection where examples from the same class are projected close to each other and, at the same time, the class centers (means) are as far apart as possible. Although it is not a true regularization technique, PCA-LDA is found to reduce the chance of overfitting that may occur with a pure LDA model.

During the recognition step, the model transforms spectra acquired from test samples into the associated model space, after which, a classifier decides to which class (if any) the spectra belongs. The model classifier uses a multivariate normal distribution (MVN) for each model class. The MVN produce a likelihood measure for each class, and Bayes' rule was then applied to derive posterior probabilities.

In silico, fivefold stratified validation was performed to determine the predictive accuracy of the model. The model-building dataset was divided in five partitions (fivefold), each of which contains a representative proportion of each class within it (stratified). Four partitions (80%) of dataset are used to build a model under the same conditions as the original model. This model is used to predict the classifications of the one partition (20%) of the training set that was left out. The cycle was repeated iteratively 5 times and each partition was predicted once by a model trained from the other four. The output of the validation details the total number of correct and incorrect classifications, as well as the number of outliers. Outliers were calculated according to the Mahalanobis distance to the nearest class center (Mahalanobis 1936).

If this distance is greater than the outlier threshold, the sample is considered an outlier. Following iterations of model optimization, an independent validation step was performed using a sample series not included in the training set.

Additional and complementary statistical analyses were performed using MetaboAnalyst 3.0 (Xia et al. 2015; Xia and Wishart 2016), Progenesis Q1 (Non-linear Dynamics, Newcastle, UK), EZInfo, and SIMCA-P (Umetrics Sartorius Stedim Biotech, Sweden) to determine candidate biomarkers and tentative assignments for preliminary pathway analysis.

## Lipid Extraction for Chromatographic Analyses

Crushed Bronte pistachio nuts (2 g) underwent lipid extraction by using 6 mL of a hexane/2-propanol 3:2 (v:v) mixture at room temperature under vigorous stirring for 1 h. The homogenate was filtered and the residue was washed twice with 4 mL of the same extraction mixture. The filtrate was dried by using a rotary evaporator.



## GC-MS Analysis of Total Fatty Acid Composition of Bronte Pistachio

The total fatty acid composition was investigated after conversion of intact lipids into fatty acid methyl esters (FAMES), by methanolic solution of potassium hydroxide at room temperature: 1 mL of hexane and 0.1 mL of methanolic potassium hydroxide solution 2 N were added to 50 mg of the lipid extract; the mixture was shaken for 30 s and the upper hexane phase was injected into a GCMS-QP2010 (Shimadzu, Milan, Italy) equipped with a split-splitless injector and an AOC-20i auto-sampler. Separation was achieved by using a SLB-IL60i (30 m  $\times$  0.25 mm id. 0.20- $\mu$ m film thickness) column (MilliporeSigma) under the following temperature program: from 50 to 280 °C at 3.0 °C/min. Injector was kept at 280 °C; injection volume was 0.2  $\mu$ L (split 1:100). Helium was used as the carrier gas at 30 cm/s linear velocity and a pressure of 27.7 KPa. MS parameters were as follows: mass range 40–550 amu with a scan interval of 0.20 s; ion source and interface temperatures were 200 °C and 220 °C, respectively. The GCMSsolution software version (Shimadzu) was used for data collection and handling. The C4–C24 even carbon saturated FAME standard solution was used for linear retention index (LRI) calculation to support identification. Then, peak assignment was based on a double filter (MS similarity over 85% and LRI  $\pm$  10) compared to lipids library (Shimadzu Europe, Duisburg, Germany).

## LC-MS Analysis of Non-Polar Lipids

The lipid extract was injected into a Nexera X2 system coupled to a LCMS-2020 mass spectrometer (Shimadzu, Kyoto, Japan). Separation was carried out on a Titan C18 100  $\times$  2.1 mm (L  $\times$  ID), 1.9  $\mu$ m dp column (MilliporeSigma), operated at a 400  $\mu$ L/min flow rate, by using acetonitrile (a) and 2-propanol (b) as mobile phase under gradient conditions: 0–52.5 min, 0–50% B. Oven temperature was set at 35 °C and the injection volume was 2  $\mu$ L. MS detection was performed by using APCI interface in positive ionization mode under the following conditions: the mass range was 300–1250  $m/z$ , with an event time of 0.5 s; desolvation line, heat block, and interface temperature were kept at 300 °C, 300 °C, and 350 °C, respectively; nebulizer and drying gas (N<sub>2</sub>) flow were set at 3 L/min and 5 L/min, respectively.

The recently introduced LRI filter was used as additional identification tool, by using a C<sub>9</sub>C<sub>9</sub>C<sub>9</sub>–C<sub>19</sub>C<sub>19</sub>C<sub>19</sub> odd carbon saturated triacylglycerol (TG) standard solution (1000 mg/L) as reference homolog series (Rigano et al. 2018). The analyses were acquired by using LabSolutions software version 5.91 (Shimadzu) and processed by using the Chromatoplus Spectra software (Chromaleont, Messina, Italy).

## HILIC-MS Analysis of Polar Lipids

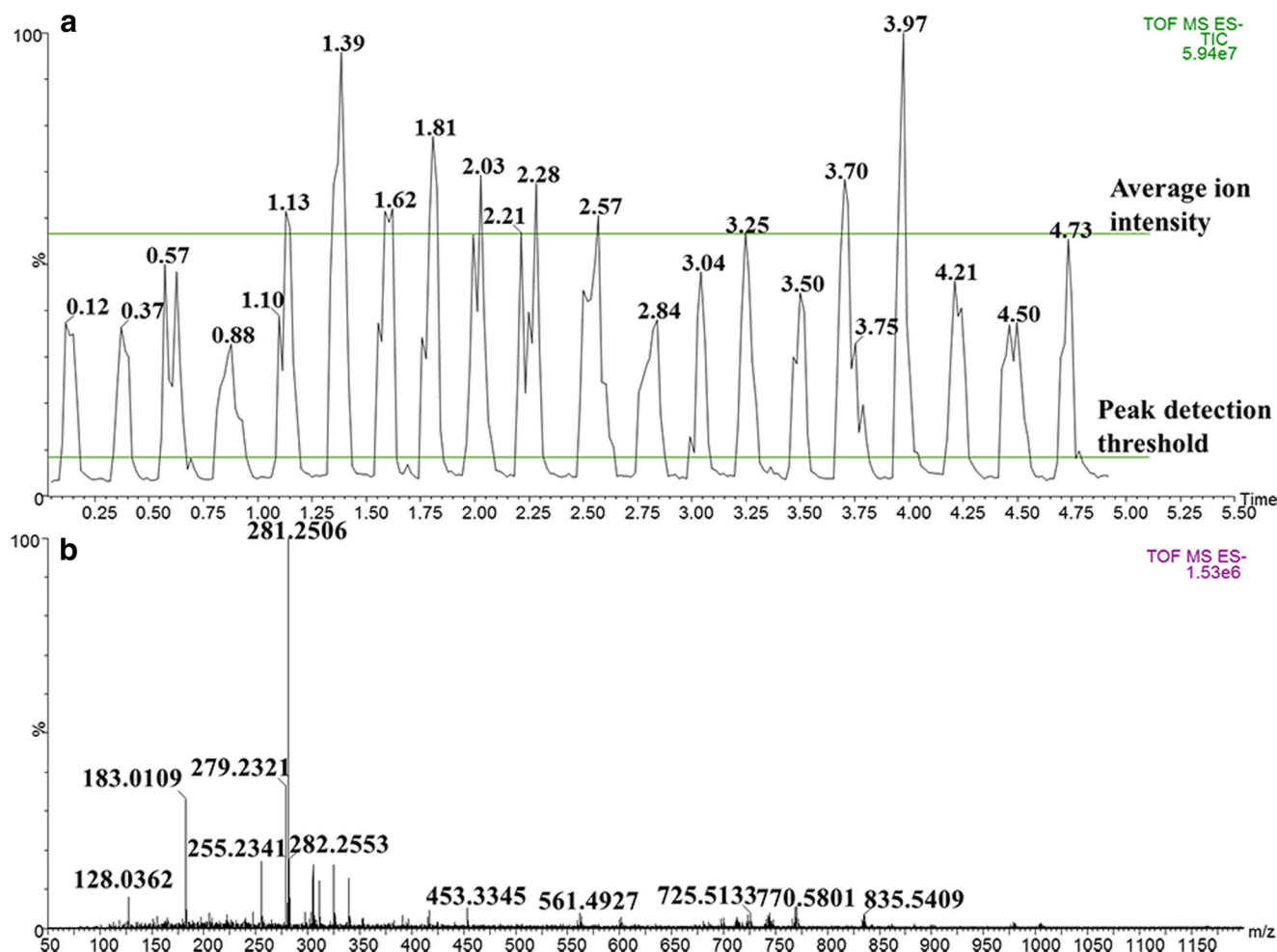
The lipid extract was injected into a Nexera X2 system coupled to a LCMS-2020 mass spectrometer (Shimadzu, Kyoto, Japan). Separation was carried out on an Ascentis Express HILIC 150  $\times$  2.1 mm (L  $\times$  ID), 2.7  $\mu$ m dp column (MilliporeSigma), operated at a 300  $\mu$ L/min flow rate, by using acetonitrile:10 mM ammonium formate 95:5 v:v (a) and acetonitrile:methanol:10 mM ammonium formate 55:35:10 v:v:v (b) as mobile phase under gradient conditions: 0–5 min, 0% B, 5–25 min, 0–40% B, 25–50 min, and 50–70% B. Oven temperature was set at 30 °C and the injection volume was 5  $\mu$ L. MS detection was performed by using ESI interface in both positive and ionization mode under the following conditions: the mass range was 200–1200  $m/z$ , with an event time of 0.3 s; desolvation line and heat block temperatures were both kept at 230 °C; nebulizer and drying gas (N<sub>2</sub>) flow were set at 1.5 L/min and 5 L/min, respectively.

Data acquisition and processing were carried out by using LabSolutions software version 5.91 (Shimadzu).

## Results and Discussion

### Mass Spectral Data

Pistachio nuts were widely studied in literature, mainly due to their beneficial effects on human health (Ryan et al. 2006; Tomaino et al. 2010; Mandalari et al. 2013; Chung et al. 2013; Song et al. 2018). Particularly, it is well known that almost 50% of their chemical composition is constituted by lipids, essentially TGs, as well as it is well-known that the most abundant FAs are oleic, linoleic, and palmitic acid (Ryan et al. 2006; Chung et al. 2013; Song et al. 2018). Since the majority of applications of REIMS technique deal with the untargeted profile of the lipid fraction, such as FAs and glycerophospholipids (PLs) in negative ionization mode (Schafer et al. 2009; Balog et al. 2010; Balog et al. 2013; Strittmatter et al. 2014; Alexander et al. 2015; Balog et al. 2015; Golf et al. 2015a; Golf et al. 2015b; Balog et al. 2016; Strittmatter et al. 2016; Black et al. 2017; St John et al. 2017; Verplanken et al. 2017), the comparison of the lipid profile between nine pistachio varieties was exploited also in this study for authenticity assessment of the geographical origin. Figure 1a reports the typical total ion current (TIC) chromatogram generated by 20 subsequent cuts of a pistachio sample in order to insert in the statistical model a sufficient number of spectra replicates, if possible, acquired by different operators and on different crushed nuts. The average ion intensity was around 35 million for all cuts with a standard deviation of  $\pm$  10 million, corresponding to a relative standard deviation percentage of less than 30%, thus revealing a satisfactory repeatability of manual cuts. The typical MS spectrum in negative



**Fig. 1** **a** Typical total ion count chromatogram showing 20 replicate iKnife sampling events taken from Sicilian Bronte pistachio nut paste; **b** combined mass spectrum (ten scans), accurate mass corrected to the ion at

255.23  $m/z$  and background subtracted, acquired using REIMS ionization in negative polarity over 50–1200  $m/z$  range

ionization mode is shown in Fig. 1b; the large mass range investigated (50–1200  $m/z$ ) allowed to detect not only FAs, PLs, and TGs, but also small molecules such as amino acids and phenolic acids. Actually, the mass range was slightly reduced to build the statistical model, since *in silico* validation test using the novel LiveID software demonstrated that the highest correctness score could be achieved by considering the mass range 200–1000  $m/z$ , including both FAs and PLs, while excluding smaller molecules. Moreover, in order to improve spectra quality, a 6-million peak detection threshold was set.

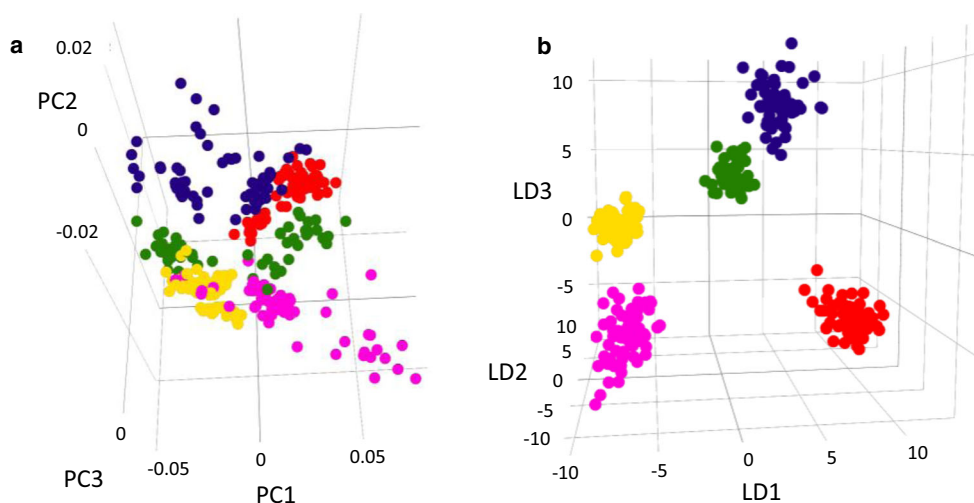
### Model Building and Validation 1: Classification According to the Geographical Origin

The results from the *in silico* stratified fivefold validation of the geographical origin of pistachio nut model, iteratively generated using different fitting parameters using the LiveID™ software, were reported in Table 1-S, while Fig. 1-S contains the confusion matrix generated using the *in silico* validation

tool in LiveID for the parameters selected as optimum. Only 1 out of 360 spectra was wrongly identified by the model, thus demonstrated the accuracy of approximately 100% in the identification of unknowns. Moreover the high identification power is ensured by the fact that at least three replicates of tentative identifications should be taken into account for the experimental validation of the model.

The reliability of the model is strongly influenced by the possibility to have access to really authentic samples, since a “false” labeled sample could totally corrupt the model, hampering any right identification.

The tridimensional visualization of the validated model, both PCA and PCA/LDA score plots, is reported in Fig. 2a, b. It enables only the classification/clusterization according to the geographical origin, not providing any information about quality (California prime and second quality) or maturation degree (different kind of Turkey and Iran nuts). Instead, the analyses on all nine pistachio varieties were used to build a big 5-class model, where each class corresponds to each geography.



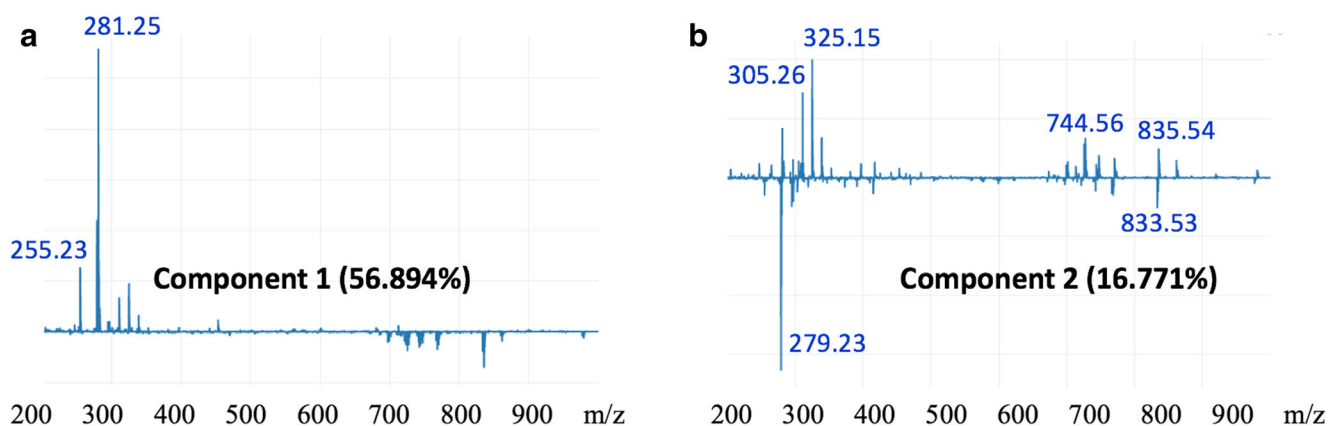
**Fig. 2** Optimized geographical origin model 3D PCA (a) and PCA/LDA (b) scores plots, generated using LiveIDTM, using parameters reported in Fig. 1-S from REIMS negative polarity ToF-MS inter-laboratory data from three different pistachio lots, from different geographical origins

Three production lots and analyses performed by different operators in different analytical laboratories were used for model building, thus performing for the first time since the introduction of this novel technique in the food field an inter-laboratory study, maximizing the analytical variability. The model was subsequently challenged using 12 different samples not included in the model training set originating from a different production lot, analyzed in triplicate using the real-time recognizer functionality. The overall outcome was determined from the mean of the three replicates (Table 2-S). Individual results having a % confidence factor of less than 80% were excluded from the overall outcome. On the basis of the independent validation, all samples were correctly classified leading to an overall correct assignment rate of 100%. Based on this validation set a 0% false positive rate and false negative rate was estimated. Such preliminary performances

are very satisfactory, especially considering that different varieties have been combined.

The same results can be highlighted from the playback recognition performed by the LiveID software for all the geographies (Fig. 2-S, A–E).

Since the LiveID software is a novel software, developed ad hoc for REIMS applications, one of the most employed software for statistical data processing, namely MetaboAnalyst by means of OPLS-DA data analysis algorithm, was also used for comparison purposes. The principal component analysis and OrthoPLS-DA scores plot can be visualized in Fig. 3-S A and B, showing already in the unsupervised analysis (Fig. 3-S A) a clear trace of subdivision in five elliptical regions according to the origin. Such a result confirmed the ability of the studied ionization strategy to provide specific fingerprints of pistachio nuts in relation with geographical/climatic region.



**Fig. 3** Loading plots generated by the LiveID software showing the  $m/z$  features responsible for the discrimination in the region between 200 and 1000  $m/z$  in the first (a) and (b) in the second PCA component

**Table 2** Tentative identifications for a selection of the discriminatory features for Sicilian Bronte pistachio highlighted as significant following multivariate statistical analysis (ANOVA  $p > 0.01$ , fold change  $\geq 2$ )

Significant ion ( $m/z$ )	Tentative assignment	Main class	Elemental composition	Ion detected
255.2328	Palmitic acid (C16:0)	Fatty acids and conjugates	C <sub>16</sub> H <sub>32</sub> O <sub>2</sub>	[M-H] <sup>-</sup>
281.2488	Oleic acid (C18:1n9)	Fatty acids and conjugates	C <sub>18</sub> H <sub>34</sub> O <sub>2</sub>	[M-H] <sup>-</sup>
283.2621	Stearic acid (C18:0)	Fatty acids and conjugates	C <sub>18</sub> H <sub>36</sub> O <sub>2</sub>	[M-H] <sup>-</sup>
279.2321	Linoleic acid (C18:2n6)	Fatty acids and conjugates	C <sub>18</sub> H <sub>32</sub> O <sub>2</sub>	[M-H] <sup>-</sup>
309.1729	Gadoleic acid (C20:1n9)	Fatty acids and conjugates	C <sub>20</sub> H <sub>34</sub> O <sub>2</sub>	[M-H] <sup>-</sup>
367.3579	Lignoceric acid (C24:0)	Fatty acids and conjugates	C <sub>24</sub> H <sub>48</sub> O <sub>2</sub>	[M-H] <sup>-</sup>
722.4880	PG (C16: 0/C16:0)	Phosphatidylglycerols	C <sub>38</sub> H <sub>74</sub> O <sub>10</sub> P	[M-H] <sup>-</sup>
833.5264	PI (C16:0/C18:2)	Phosphatidylinositols	C <sub>43</sub> H <sub>79</sub> O <sub>13</sub> P	[M-H] <sup>-</sup>
835.5418	PI (C16:0/C18:1)	Glycerophosphoinositols	C <sub>43</sub> H <sub>81</sub> O <sub>13</sub> P	[M-H] <sup>-</sup>
881.7622	TG (C18:1/C18:1/C18:2)	Triacylglycerols	C <sub>57</sub> H <sub>102</sub> O <sub>6</sub>	[M-H] <sup>-</sup>
883.7815	TG (C18:1/C18:1/C18:1)	Triacylglycerols	C <sub>57</sub> H <sub>104</sub> O <sub>6</sub>	[M-H] <sup>-</sup>

### Candidate Biomarkers for Bronte Pistachio Nuts

After model building, the corresponding loading plots for the main principal components were carefully investigated to reveal significant ions responsible for sample differentiation. In particular, Fig. 3a, b shows the loading plots in the first two components, contributing to the discrimination with a percentage of about 57% and 17%, respectively. The base peaks in both cases correspond mainly to FA species, including palmitic and oleic acid ( $m/z$  255.23 and  $m/z$  281.25, respectively) in the positive direction of the first component, linoleic acid ( $m/z$  279.23) in the negative direction of the second component.

As alternative/complementary software, Progenesis QI and EZInfo were used to show the significant features ( $m/z$ ) responsible to the discrimination between Sicilian Bronte pistachio and all the other geographies (Fig. 4-S A–D). Such loading plots allowed to find selected biomarkers for Bronte pistachio nuts, according to one of the main purposes of the present work to preserve the precious Sicilian variety. In order for an ion to be relevant a cutoff value of  $p > 0.01$  (by ANOVA test, fold change  $\geq 2$ ) was established. Also, in this case, the novel LiveID software provided very similar results with respect to more consolidated software packages. The main fragment ions revealed from both software were selected as candidate biomarkers and are reported in Table 2.

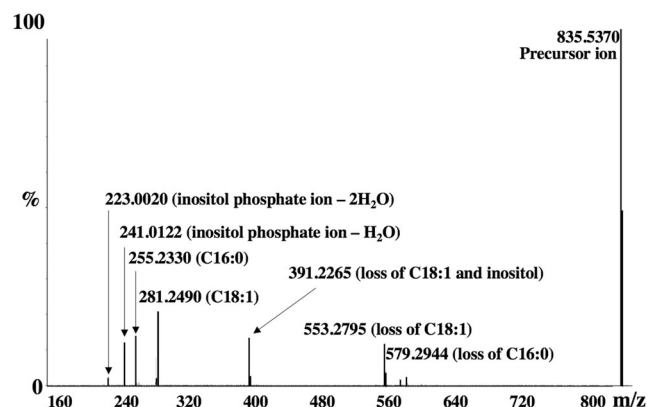
MetaboAnalyst software was also employed to identify the features responsible for the discrimination between pistachio nuts from different geographical origins, obtaining many  $m/z$  values already highlighted by using both LiveID and Progenesis QI software. Figure 5-S represents a hierarchical cluster analysis heat map, generated using MetaboAnalyst 3.0 software, summarizing the features responsible for the discrimination between pistachio nuts from all five different geographical origins (ANOVA  $p$  value  $\leq 0.05$  max fold change  $\geq 2$  and minimum CV  $\geq 30$ ). All these compounds have been already assigned as candidate biomarkers after data

processing by LiveID and Progenesis QI software (Table 2), thus representing a further confirmation of the proper applicability of the novel software to REIMS analyses.

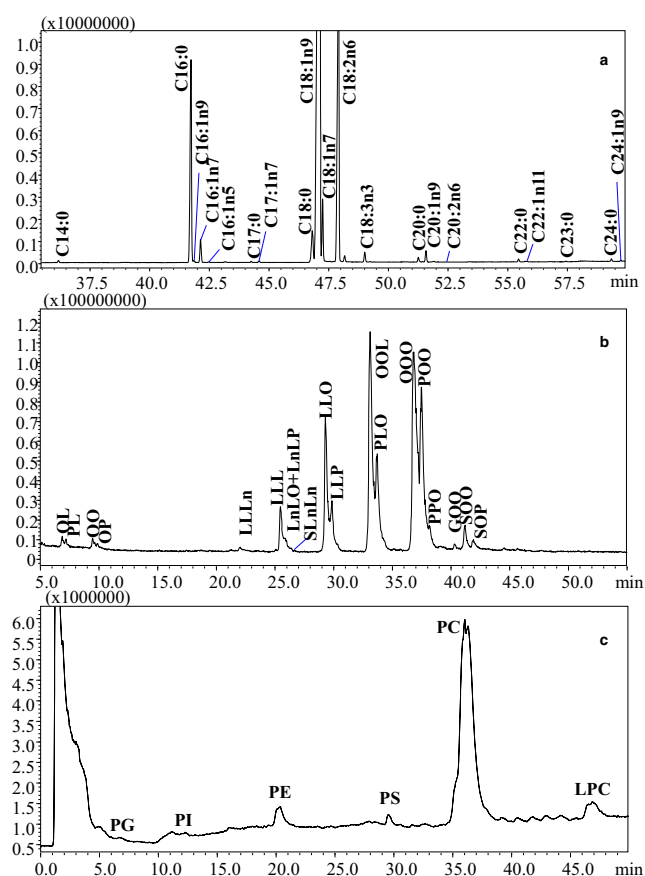
Identification of PLs and TAGs was confirmed by MS/MS experiments. Particularly MS/MS spectra highlighted the right PL class, FA combination and positional isomer. As an example, Fig. 4 reports the MS/MS spectra of phospholipidinositol PI (C16:0/C18:1); the losses of inositol polar head, C18:1 and C16:0 FA, indicates the PI class and the single-FA combination, while fragment ratio and the presence of an intense  $m/z$  391.2265 value, corresponding to the loss of inositol and C18:1, and lead to the correct positional isomer, where C18:1 occupies the sn1 PL position.

In addition, a conventional GC-MS analysis was performed for the elucidation of the total FA composition, particularly to confirm the presence of FAs normally not reported in previous pistachio works (Ryan et al. 2006; Chung et al. 2013), such as C24:0. The total FA profile is reported in Fig. 5a, along with peak identification.

Similarly, LC-MS techniques were employed for intact lipid analysis. In particular, non-polar lipids (essentially TGs and

**Fig. 4** MS/MS spectrum of PI (C16:0/C18:1) along with fragment elucidation





**Fig. 5** **a** GC-MS profile of FAMES obtained from Bronte pistachio nuts, **b** reversed phase LC-MS profile of acylglycerols in the lipid extract of Bronte pistachio nuts, **c** HILIC-MS profile of PLs in the lipid extract of Bronte pistachio nuts. Fatty acid legend: P, palmitic acid C16:0; Ln, linolenic acid C18:3n3; L, linoleic acid C18:2n6; O, oleic acid C18:1n9; S, stearic acid; G, gadoleic acid C20:1n9; phospholipid legend: PG, phosphatidylglycerol; PI, phosphatidylinositol; PE, phosphatidylethanolamine; PC, phosphatidylcholine; LPC, lysoPC

diacylglycerols DGs) were separated and identified by non-aqueous reversed phase LC-MS (Fig. 5b), while PLs were analyzed by HILIC-MS (Fig. 5c). Identified lipids were reported in Table 3-S and Table 4-S along with observed  $m/z$  values.

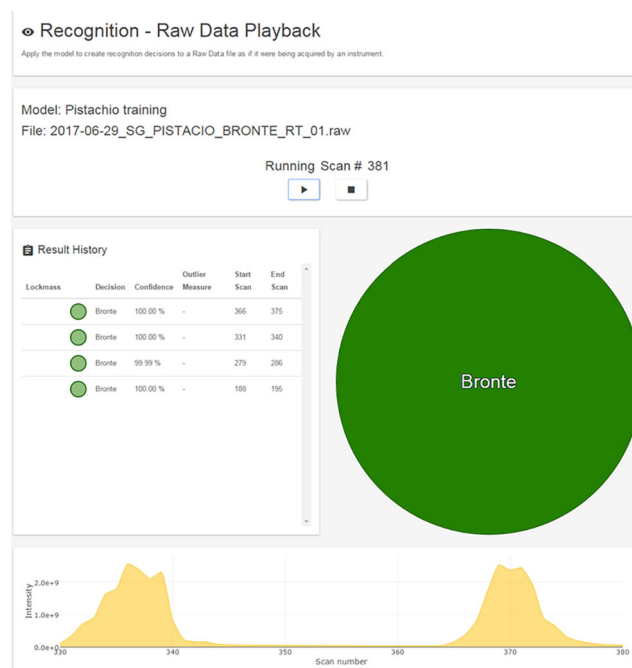
For a more complete overview, the pistachio nut REIMS spectrum in positive polarity was also acquired to focus on the TG profile (Fig. 6-S). The REIMS (+) spectrum highlighted the presence of TG species, but also DGs and monoacylglycerols (MGs) were detected. Consequently, due to the deep similarity between iKnife and conventional LC-MS and GC-MS results, it is possible to assess that the new technique can be successfully exploited as a new powerful shotgun lipidomics tool.

Interestingly, the LC-MS elucidation of intact lipids pinpointed that candidate biomarkers for Bronte pistachio matches with the most abundant TG species and with two of the less representative PLs, namely two phosphatidylinositols (PIs), while the most abundant phosphatidylcholines (PCs), followed by phosphatidylethanolamines (PEs), and LysoPCs (LPCs) are not included in the discriminant features.

## Model Transferability

The inter-laboratory study presented in this work offered the possibility, for the first time in the state of the art of the present analytical approach, to transfer the model entirely built in a given analytical laboratory to another laboratory, located in a totally different geographical region. Different ambient conditions, as well as different quality of solvents and gas employed, could affect the ionization pattern and, consequently, the fingerprinting of a given sample. Therefore, it is extremely important to verify if the spectra of the same sample acquired in the two laboratories are significantly different from a statistical standpoint.

This point was assessed not only validating a model including both laboratories spectral data (see previous section), but also by using in the second laboratory, the model built in the first one for the identification of a set of validation samples of all the geographies. As an example, Fig. 6 reports the correct identification of Bronte pistachio nuts analyzed in Messina, according to a model built in Wilmslow. The recognition occurs for more subsequent cuts with an accuracy of around 100%, by setting a maximum standard deviation of 30 units, meaning that two samples are recognized “identical” if they differ of less than this threshold (in this case the samples are not statistically different). In effect, the spectra comparison reported in Fig. 7-S A and B, highlights slight differences in the profiling, but this is not the main object of this novel approach, aiming to create a database of samples, rather than a spectral library.

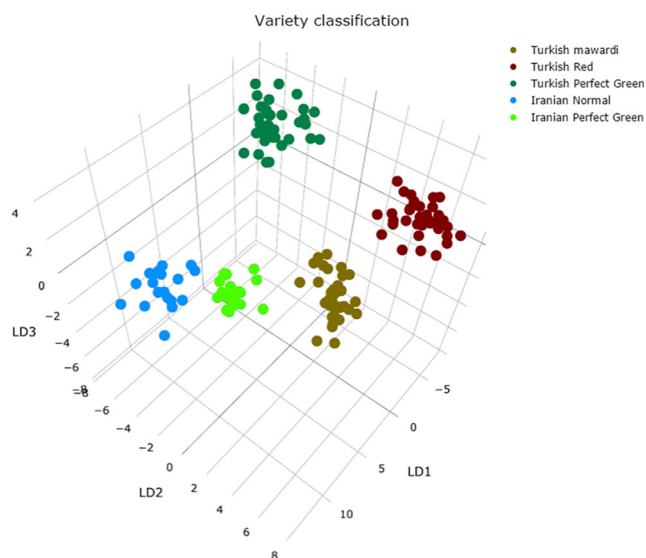


**Fig. 6** Identification by the LiveID software of Sicilian Bronte pistachio nuts analyzed in Messina, according to a model built in Wilmslow

The model transferability here demonstrated represents a valid and powerful starting point for the establishment of the novel technique in the food application field, since it should be possible to build a model in an analytical laboratory and use it in another laboratory for authenticity evaluation and for the safeguard of PDO, PGI, or DOCG products. In other words, a research laboratory can build a suitable model for discriminating a specific product, and the same model could be employed from the cross-border investigative authorities to detect fraudulent practices.

## Model Building and Validation 2: Classification According to the Variety

A two-tier classification system was devised whereby a second model was trained to recognize the specific variety in the case of the Iranian and Turkish pistachio nuts following the initial classification of geographical origin (Fig. 7). The same iterative approach described above was used for the second model optimization. The following parameters were selected for the variety classification model:  $m/z$  200–1000, 20 PCA dimensions followed by four LDA dimensions. The predictive classification accuracy was calculated around 98% via stratified fivefold in silico validation, which presented only three “false suspicious” measurement out of 147 analyses, viz. a 2% false positive rate. Such preliminary results are very encouraging, especially considering the relatively low number of samples involved.



**Fig. 7** a 3D PCA/LDA scores plots for sub-classification of pistachio nut variety generated in LiveID using 20 PCA dimensions and 4 LDA dimension generated from the training set of 147 spectra (34 Turkish mawardi, 35 Turkish perfect green, 35 Turkish red, 20 Iranian normal, 23 Iranian perfect green) across a mass range of 200–1000  $m/z$  with a binning parameter of 0.1 Da

Sub-classification was subsequently performed using the variety model for samples classified as Iranian or Turkish.

Table 5-S reports the experimental validation results obtained on samples not employed for model building. A correct sample assignment (for three replicates) was achieved for all the varieties and production lots, in both analytical laboratories and considering analyses carried out by different operators. As couple of examples, Fig. 8-S A and B provides the correct recognition of both Turkish and Iranian samples.

## Conclusions

The technique described in the present research demonstrated very high potentiality in the differentiation of pistachio nuts coming from different geographical origins. The main goal with respect to other food applications of the iKnife technique, where different animal species or genders have been successfully compared, was that exactly the same botanical species was analyzed for all the geographies. As a consequence, the present work could be the starting point of a more extensive project that aims to preserve the finest and most famous “Made in Italy” products.

Moreover, some discussions with small local producers suggest to create a similar model for pistachio-derived products, such as cream, ice-cream, and pesto, since the probability to detect fraudulent actions significantly increase with respect to pure pistachio nuts. In addition, because of the high probability to find mixtures of different pistachio in the final derived products, it will be very helpful to include in the database mixture of pistachio nuts at known percentage ratios.

Moreover, an inter-laboratory study was performed for the first time, confirming the potentiality of the new technique to be easy-to-use by any operator, even not expert, after the building of a reliable database in an analytical laboratory.

Finally, the comparison with conventional chromatographic techniques made the identification of candidate biomarkers very reliable, conversely from the tentative assignment of previous works.

## Compliance with Ethical Standards

**Conflict of Interest** Francesca Rigano declares that she has no conflict of interest. Sara Stead declares that she has no conflict of interest. Domenica Mangraviti declares that she has no conflict of interest. Renata Jandova declares that she has no conflict of interest. Davy Petit declares that he has no conflict of interest. Luigi Mondello declares that he has no conflict of interest.

**Ethical Approval** This article does not contain any studies with human participants or animals performed by any of the authors.

**Informed Consent** Informed consent is not applicable for this study.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Alexander J, Gildea L, Balog J et al (2015) A novel methodology for in vivo endoscopic phenotyping of colorectal cancer based on real-time analysis of the mucosal lipidome: a prospective observational study of the iKnife. *Surg Endosc* 31(3):1361–1370
- Balog J, Szaniszló T, Schaefer K-C, Denes J, Lopata A, Godorhazy L, Szalay D, Balogh L, Sasi-Szabo L, Toth M, Takats Z (2010) Identification of biological tissues by rapid evaporative ionization mass spectrometry. *Anal Chem* 82:7343–7350
- Balog J, Sasi-Szabó L, Kinross J, Lewis MR, Muirhead LJ, Veselkov K, Mirnezami R, Dezső B, Damjanovich L, Darzi A, Nicholson JK, Takats Z (2013) Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. *Sci Transl Med* 5(194):194ra93
- Balog J, Kumar S, Alexander J, Golf O, Huang J, Wiggins T, Abbassi-Ghadi N, Enyedi A, Kacska S, Kinross J, Hanna GB, Nicholson JK, Takats Z (2015) In vivo endoscopic tissue identification by rapid evaporative ionization mass spectrometry (REIMS). *Angew Chem Int Ed* 54:11059–11062
- Balog J, Perenyi D, Guallar-Hoyas C, Egri A, Pringle SD, Stead S, Chevallier OP, Elliott CT, Takats Z (2016) Identification of the species of origin for meat products by rapid evaporative ionization mass spectrometry. *J Agric Food Chem* 64:4793–4800
- Black C, Chevallier OP, Haughey SA, Balog J, Stead S, Pringle SD, Riina MV, Martucci F, Acutis PL, Morris M, Nikolopoulos DS, Takats Z, Elliott CT (2017) A real time metabolomic profiling approach to detecting fish fraud using rapid evaporative ionisation mass spectrometry. *Metabolomics* 13:153–165
- Chung KH, Shin K, Hwang HJ, Choi K-S (2013) Chemical composition of nuts and seeds sold in Korea. *Nutr Res Pract* 7(2):82–88
- Cubero-Leon E, Peñalver R, Maquet A (2014) Review on metabolomics for food authentication. *Food Res Int* 60:95–107
- Golf O, Muirhead LJ, Speller A, Balog J, Abbassi-Ghadi N, Kumar S, Mróz A, Veselkov K, Takats Z (2015a) XMS: cross-platform normalization method for multimodal mass spectrometric tissue profiling. *J Am Soc Mass Spectrom* 26:44–54
- Golf O, Strittmatter N, Karancsi T, Pringle SD, Speller AVM, Mroz A, Kinross JM, Abbassi-Ghadi N, Jones EA, Takats Z (2015b) Rapid evaporative ionization mass spectrometry imaging platform for direct mapping from bulk tissue and bacterial growth media. *Anal Chem* 87:2527–2534
- Mahalanobis PC (1936) On the generalised distance in statistics. *Proc Natl Inst Sci India* 12:49–55
- Mandalari G, Bisignano C, Filocamo A, Chessa S, Sarò M, Torre G, Faulks RM, Dugo P (2013) Bioaccessibility of pistachio polyphenols, xanthophylls, and tocopherols during simulated human digestion. *Nutrition* 29:338–344
- Neidert MC, Bozinov O (2013) Mass spectrometry-based intraoperative tissue identification in neurosurgery. *World Neurosurg* 80:683–684
- Official Journal European Union 2009/C 130/09: Publication of an application pursuant to Article 6(2) of Council Regulation (EC) No 510/2006 on the protection of geographical indications and designations of origin for agricultural products and foodstuffs
- Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002
- Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011
- Rigano F, Oteri M, Russo M, Dugo P, Mondello L (2018) Proposal of a linear retention index system for improving identification reliability of triacylglycerol profiles in lipid samples by liquid chromatography methods. *Anal Chem* 90:3313–3320
- Ryan E, Galvin K, O'Connor TP, Maguire AR, O'Brien NM (2006) Fatty acid profile, tocopherol, squalene and phytosterol content of Brazil, pecan, pine, pistachio and cashew nuts. *Int J Food Sci Nutr* 57:19–228
- Schäfer K-C, Denes J, Albrecht K et al (2009) In vivo, in situ tissue analysis using rapid evaporative ionization mass spectrometry. *Angew Chem Int Ed Engl* 48:8240–8242
- Song S, Cheong L-Z, Wang H, Man QQ, Pang SJ, Li YQ, Ren B, Wang Z, Zhang J (2018) Characterization of phospholipid profiles in six kinds of nut using HILIC-ESI-IT-TOF-MS system. *Food Chem* 240:1171–1178
- St John ER, Balog J, McKenzie JS et al (2017) Rapid evaporative ionisation mass spectrometry of electrosurgical vapours for the identification of breast pathology: towards an intelligent knife for breast cancer surgery. *Breast Cancer Res* 19:59–72
- Strittmatter N, Rebec M, Jones EA, Golf O, Abdolrasouli A, Balog J, Behrends V, Veselkov KA, Takats Z (2014) Characterization and identification of clinically relevant microorganisms using rapid evaporative ionization mass spectrometry. *Anal Chem* 86:6555–6562
- Strittmatter N, Lovrics A, Sessler J, McKenzie JS, Bodai Z, Doria ML, Kucsma N, Szakacs G, Takats Z (2016) Shotgun lipidomic profiling of the NCI60 cell line panel using rapid evaporative ionization mass spectrometry. *Anal Chem* 88:7507–7514
- Tomaino A, Martorana M, Arcoraci T, Monteleone D, Giovinozzzo C, Saija A (2010) Antioxidant activity and phenolic profile of pistachio (*Pistacia vera* L., variety Bronte) seeds and skins. *Biochimie* 92:1115–1122
- Verplanken K, Stead S, Jandova R, Poucke CV, Claereboudt J, Bussche JV, Saeger SD, Takats Z, Wauters J, Vanhaecke L (2017) Rapid evaporative ionization mass spectrometry for high-throughput screening in food analysis: the case of boar taint. *Talanta* 169:30–36
- Wolfender J-L, Rudaz S, Choi YH, Kim HK (2013) Plant metabolomics: from holistic data to relevant biomarkers. *Curr Med Chem* 20(8):1056–1090
- Xia J, Sinelnikov IV, Han B, Wishart David S (2015) MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic Acids Res*, 43(W1):W251–257
- Xia J, Wishart DS (2016) Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. *Curr. Protoc. Bioinform.* 55:14.10.1–14.10.91