Elsevier Editorial System(tm) for Food Research International

Manuscript Draft

Manuscript Number:

Title: Identification of phenolic markers for saffron authenticity and origin: an untargeted metabolomics approach

Article Type: Research Articles

Keywords: Crocus sativus; food metabolomics; polyphenols; multivariate statistics; food integrity.

Corresponding Author: Professor Luigi Lucini, Ph.D.

Corresponding Author's Institution: Università Cattolica Sacro Cuore

First Author: Biancamaria Senizza

Order of Authors: Biancamaria Senizza; Gabriele Rocchetti; Silvia Ghisoni; Matteo Busconi; Marcelino De Los Mozos Pascual; José Antonio Fernandez; Luigi Lucini, Ph.D.; Marco Trevisan

Abstract: Saffron is a high-quality and expensive spice, widely subjected to adulteration. An UHPLC-ESI/QTOF-MS metabolomic-based approach was therefore used to discriminate adulterated (added with different percentage of other parts of the flower) saffron as well as to trace its geographical origin. Both unsupervised (hierarchical clustering) and supervised OPLS-DA multivariate statistics allowed discriminating authentic styles from styles added of other floral components, as well as PDO vs non PDO saffron samples according to their chemical fingerprints. The markers were then validated through ROC curves. Anthocyanins and glycosidic flavonols were the best markers of the styles' adulteration. However, flavonoids (mainly flavonols and flavones), together with protocatechuic aldehyde and isomeric forms of hydroxybenzoic acid were validated as markers for the discrimination of PDO vs non PDO saffron samples. This work outlines the potential of untargeted metabolomics based on UHPLC-ESI/QTOF mass spectrometry for saffron authenticity and traceability.

Research Data Related to this Submission

Title: Data for: Identification of phenolic markers for saffron authenticity and origin: an untargeted metabolomics approach
Repository: Mendeley Data
https://data.mendeley.com/datasets/r68b8yn64v/draft?a=4ec2bf64-d7ff-4826-bec3-54df565586ce

UNIVERSITÀ CATTOLICA del Sacro Cuore

Dipartimento di Scienze e Tecnologie Alimentari
per una filiera agro-alimentare Sostenibile - DiSTAS

Area di Chimica agraria, alimentare ed ambientale

February 20, 2019

Dear Editor,

We are glad to submit the research article entitled "**Identification of phenolic markers for saffron authenticity and origin: an untargeted metabolomics approach**" to be considered for publication in Food Research International.

In the last years, there has been a growing interest on food authenticity and traceability, with some approaches or markers being proposed. Indeed, frauds and adulteration processes are hampering the food industries having high-quality policies. Among others, saffron represent a very interesting case study because its quality is linked to geographical origin and because of the huge number of frauds/adulterations known for this expensive spice. Nonetheless, the present case study and the approach proposed might be extended to a wide number of other high-quality food products.

To date, most of quality control procedures for saffron are related to spectrophotometric assays (related to its characteristic carotenoids), in spite of the cost of the spice. However, given the diversity of metabolites in a complex food such as saffron, liquid chromatography coupled to untargeted high-resolution mass spectrometry might allow us gaining a comprehensive picture on the compounds present. In our work, we used an untargeted profiling approach, together with multivariate chemometrics, to investigate the relationship/differences between metabolomic signatures as a function of geographical origin. Furthermore, we used the same approach to investigate the most challenging fraud, i.e. the "dilution" of saffron styles using other floral portions (in fact, neither genetics are capable to identify such counterfeits).

Interestingly, supervised statistics (*i.e.*, OPLS-DA) allowed identifying the markers related to origin and authenticity. The following ROC curve allowed validating the very most of the markers proposed. Therefore, the novelty of the work lies in the proposal of a robust and effective approach to support saffron authenticity, potentially having a much wider field of application across different plant foods.

The manuscript was prepared in compliance to the *Guide to Authors;* the co-authors represent the appropriate people having contributed in some way to the design, implementation and/or analysis, interpretation and reporting of the study.

All co-authors have seen a draft copy of the manuscript and agree with its publication, and they declare that there was not any financial/commercial conflict of interest.

The work has not been published previously, either completely, in another form or in part and it is not under consideration for publication elsewhere.

I remain at your disposal for any clarifications pertaining to our submission that might be deemed necessary.

Sincerely,

29122 Piacenza Via Emilia Parmense, 84 - telefono +39 0523 599 245 - fax +39 0523 599 246 - e-mail segreteria.distas-pc@unicatt.it

Luigi Lucini

HIGHLIGHTS

- Authenticity and traceability of saffron was assessed by untargeted metabolomics.
- OPLS-DA multivariate statistics discriminated adulterated saffron samples.
- Anthocyanins and flavonols were strongly affected by adulteration.
- Flavonoids and hydroxybenzoic acids were the best markers of origin.

UHPLC-QTOF phenolic profiling and multivariate statistics to ensure saffron authenticity

Biancamaria Senizza[a], Gabriele Rocchetti[a], Silvia Ghisoni[a], Matteo Busconi[b], Marcelino De Los Mozos Pascual[c], José Antonio Fernandez[d], Luigi Lucini[a*], Marco Trevisan[a]

[a] *Department for sustainable food process, Università Cattolica del Sacro Cuore, Via Emilia Parmense 84, 29122 Piacenza, Italy.*

[b] *Department of sustainable crop production, Università Cattolica del Sacro Cuore, Via Emilia Parmense 84, 29122 Piacenza, Italy.*

[c] *Centro de Investigación Agroforestal de Albaladejito, Instituto Regional de Investigación y Desarrollo Agroalimentario y Forestal, Cuenca, Spain.*

[d] *IDR-Biotechnology and Natural Resources, Universidad de Castilla−La Mancha, Albacete, Spain*

[*] Corresponding author: via Emilia Parmense 84, 29122 Piacenza (Italy). Tel: +39 0523 599156. Email: luigi.lucini@unicatt.it

Abstract

Saffron is a high-quality and expensive spice, widely subjected to adulteration. An UHPLC-ESI/QTOF-MS metabolomic-based approach was therefore used to discriminate adulterated (added with different percentage of other parts of the flower) saffron as well as to trace its geographical origin. Both unsupervised (hierarchical clustering) and supervised OPLS-DA multivariate statistics allowed discriminating authentic styles from styles added of other floral components, as well as PDO *vs* non PDO saffron samples according to their chemical fingerprints. The markers were then validated through ROC curves. Anthocyanins and glycosidic flavonols were the best markers of the styles' adulteration. However, flavonoids (mainly flavonols and flavones), together with protocatechuic aldehyde and isomeric forms of hydroxybenzoic acid were validated as markers for the discrimination of PDO *vs* non PDO saffron samples. This work outlines the potential of untargeted metabolomics based on UHPLC-ESI/QTOF mass spectrometry for saffron authenticity and traceability.

Keywords: *Crocus sativus*; food metabolomics; polyphenols; multivariate statistics; food integrity.

1. Introduction

Saffron is obtained from the dried stigmas of *Crocus sativus* L. and it is mainly used in the food sector mainly as both spice and food dye because of its particular aromatic properties and color. However, it is reported to possess also pharmacological and therapeutic properties (Gohari, Saeidnia, & Mahmoodabadi, 2013). This spice is successfully cultivated in European countries such as Greece, Spain and Italy, as well as India and Morocco, with Iran being the world's biggest producer and exporter (Zeka *et al*., 2015).

It contains over 150 volatile compounds (Winterhalter & Straubinger, 2000) together with non-volatiles like flavonoids (such as glycosides of kaempferol and quercetin), carotenoids, α- and β-carotenes and isophorones (1-15). The responsible of its peculiar attributes are mainly crocins, picrocrocin and safranal. Crocin, the unique water-soluble carotenoid, and its esters, contribute to red color. In particular, 4-(β-d-glucopyranosyl)-2,6,6-trimethyl-1-cyclohexene-1-carboxaldehyde, a quite polar terpenoid glycoside, called picocrocin, is identified as responsible for saffron bitterness (Valle Garcia-Rodriguez et al., 2014). The characteristic aroma of saffron during drying and preservation is associated to the presence of safranal (Caballero-Ortega, Pereda-Miranda, & Abdullaev, 2007), a compound obtained via hydrolysis and oxidations of picrocrocin. Furthermore, the monoterpene aldehyde 2,6,6-trimethyl-1,3-cyclohexene-1-carboxaldehyde, is included among the main components of the volatile fraction in saffron (Carmona et al., 2005).

In the last years, saffron traceability has become a topic of great interest mainly from an economical point of view, considering that this spice is one of the most expensive agricultural products of the world (Soffritti et al., 2016). In fact, the costs and labor required to plantation and production, together with the huge number of flowers needed to obtain the spice, are the reasons this product is widely subjected to frauds and adulterations. In this regard, the stigma's powder could be easily mixed with different parts of the same plant rather than extraneous plants, coloring substances and synthetic powders. The quality of saffron is certified in the international trade market by the ISO 3632 normative [International standard ISO 3632-2: Saffron (*Crocus sativus* L.)

test methods]. This normative allows to control the quality of saffron by monitoring the chromatographic profiles following UV measurements of polar dyes and pigments (crocins) at 440 nm, 250 nm for picrocrocin and 310 nm for safranal. In addition, the possible presence of some potential toxic colorants can be also evaluated. However, the use of this standard methods has been proved to be unreliable in detect adulterations and frauds, as the standard ISO 3632 is not specific and enable to discriminate authentic and adulterate saffron (Sabatino et al., 2011). The analytics methods proposed to detect illicit addition and adulterations include capillary electrophoresis (Zougagh, Simonet, Rios, Valcarcel, 2005), nuclear magnetic resonance spectroscopy (NMR) (Assimiadis, Tarantilis, & Polissiou, 1998), UV–vis spectrophotometric measurements and high-performance-liquid chromatography. Such approaches are able to detect only a limited number of adulterations because they are based on targeted approaches.

Besides adulterations, it becomes important to assess also the origin of saffron, since its quality is linked to the pedo-climatic and cultivation conditions (Carmona et al., 2005). Indeed, numerous analytical techniques have been used also for classifying saffron on the basis of geographical origin, such as high-performance-liquid chromatography coupled to photometric or mass spectrometric (MS) detection (D'Archivio, Giannitto, Maggi, & Ruggieri, 2016; Guijarro-Díez, Nozal, Marina, & Crego, 2015a). GC–MS methods have been also developed for the characterization and quantitative determination of volatile saffron markers (Bononi, Milella, & Tateo, 2015; Sereshti, Heidari, & Samadi, 2014). Furthermore, H-NMR metabolomic fingerprinting approaches are used to assess the quality of saffron with an unsupervised classification (Yilmaz, Nyberg, Mølgaard, Asili, & Jaroszewski, 2010). Compared with NMR, MS is much more sensitive, enabling to measure low abundance compounds thus providing useful information in the search of new markers. Furthermore, the specificity of MS, ensured by high resolution and/or MS/MS experiments, facilitates the identification of selected markers through the elucidation of their chemical structures. In this regard, metabolomic approaches based on ultra-high-pressure liquid chromatography

coupled to high resolution mass spectrometry are very effective in discriminating the authenticity and adulteration practices in saffron (Guijarro-Díez et al., 2015b).

Therefore, in this work, our goal was to use untargeted metabolomics to ensure the authenticity and traceability of saffron. In more detail, we aimed to investigate the authenticity of saffron in the most difficult adulteration procedure, i.e. when styles are mixed with different parts of the flower (stamen and tepals). Furthermore, we investigated the potential of metabolomics for saffron traceability purposes by including commercial saffron samples together with Italian PDO saffron samples.

2. Materials and methods

2.1. Samples

In Italy, saffron is mainly cultivated near L'Aquila (Piana di Navelli), followed by the Sardinia (Province of Medio Campidano), Tuscany (San Gimignano, Florence Hills and Maremma) and Umbria (Cascia and Città della Pieve) regions. In this regard, saffron from Sardinia, L'Aquila and San Gimignano are three of the European saffron that can claim the PDO certification mark (European Commission, 2015). The PDO Italian saffron samples (namely Navelli, San Gimignano and Sardinia) were kindly provided by local producers within each Consortium for PDO. Besides PDO products, 15 additional commercial non-PDO samples of different brands were acquired in local supermarkets. Finally, Iranian saffron was provided by the Hamadan University. This latter was included considering that Iran is the first saffron producer worldwide.

Finally, twelve pools of saffron tissues (four pools of tepals, four pools of stamens and four pools of stigmas) were obtained from the World Saffron and Crocus Collection (WSCC) located at the Bank of Plant Germplasm of *Centro de Investigación Agroforestal de Albadalejito,* Cuenca (Spain). These latter were used to carry out the adulteration trials. With this aim, samples having different inclusions levels (i.e., 5%, 10%, 30% and 50%) of either stamens or tepals in styles, were

prepared. Such specific adulteration was chosen considering that it represents a major challenge in saffron authenticity, for which also genetic approaches are prone to failure.

2.2. Extraction and UHPLC-QTOF-MS analysis

Polyphenols were extracted in triplicates from 10 mg of either different organs (namely styles, stamens and tepals) or commercials samples, using an hydroalcoholic solution consisting in methanol 80% acidified with 1% formic acid. A homogenizer-assisted extraction was applied, by using an Ultra-turrax (Ika T10, Staufen, Germany) for 5 min. Samples were transferred in Eppendorf tubes and then centrifuged for 10 min at 20 °C, at 7000 *x g*. After centrifugation, supernatants were filtered using 0.22 μm cellulose syringe filters directly into amber vials for analysis.

The screening of phytochemicals in the extracts was carried out as previously described by Ben Mohamed and co-authors (2018), with small modifications. Briefly, analysis was carried out through ultra-high-performance liquid chromatography coupled to electrospray quadrupole-time-of-flight mass spectrometry (UHPLC-ESI/QTOF). The mass spectrometer (G6550 mass spectrometer; from Agilent Technologies, Santa Clara, CA, USA) operated in positive ionization (ESI+) and scan mode, acquiring accurate masses in the 100–1000 m/z range. The chromatographic separation conditions were optimized in previous works (Blasi et al., 2018). Raw features were processed by using the software Profinder B.07 (Agilent Technologies), based on the "find-by-formula" algorithm. In particular, features identification was recursively carried out exploiting both accurate mass and isotopic profiles. A custom database obtained combining polyphenols (Phenol-Explorer 3.6; http://phenol-explorer.eu/) integrated by some of the most important compounds characterizing saffron (namely crocetin, picrocrocin, safranal and zeaxanthin) was used as a reference for annotation purposes, with a 5-ppm tolerance for mass accuracy. The following data processing (Ben Mohamed et al., 2018) allowed to retain only those compounds identified within 100% of

replications in at least one treatment. The dataset obtained was then used for statistics and chemometrics.

2.3. Statistical analysis

Normalization of metabolomics-based data was done using the Agilent Mass Profiler Professional B.12.06 software, as previously reported (Rocchetti et al., 2018). A hierarchical clustering (HCA) was created in order to group samples according to intrinsic similarities. Afterwards, supervised orthogonal partial least squares discriminant analysis (OPLS-DA) was carried out using SIMCA 13 software (Umetrics, Malmo, Sweden). OPLS-DA model was cross validated and inspected for outliers, as described in a previous work (Rocchetti et al., 2018). Thereafter, model parameters (R^2Y and Q^2Y) were recorded and misclassification tables generated. The variables importance in projection (VIP) was then used to select those compounds possessing the highest discrimination potential (VIP score > 1) in the predictive model.

Finally, to validate the potential markers outlined by the VIP approach, Receiver Operating Characteristics (ROC) curves were performed, using the SPSS Statistics software (v.25.0) (Xia et al., 2013). The area under the ROC curve (AUC) was inspected in order to evaluate the global performance of each VIP marker.

3. Results and discussion

3.1 UHPLC-QTOF-MS discrimination of styles adulterated with different percentage of tepals and stamens

Overall, saffron quality is linked to the concentration of three main constituents, i.e. crocins, picrocrocin and safranal, that are abundant into styles. In this regard, picrocrocin has been widely considered as the best authenticity biomarker of saffron (Alonso, Zalacain, & Carmona, 2012). In our experimental conditions, both the unsupervised HCA and the supervised OPLS-DA multivariate

statistical approaches allowed a clear differentiation between styles and styles with inclusions. Interestingly, the HCA allowed classifying different parts of the saffron flower according to their metabolomic profile and, through randomly comparison between profiles, to evaluate differences and similarities. The unsupervised HCA resulted in two main groups: the first cluster included all stigma samples, while the second cluster consisted of stigmas added with tepals and stamens (supplementary material). In fact, the heat map showed that stigma samples possess a group of compounds completely absent in the counterfeit counterparts, and these differences are already evident at 5% of inclusion. Afterwards, in order to investigate the contribution of each group of metabolites for discrimination purposes, the supervised orthogonal projection to latent structures discriminant analysis (OPLS-DA) was carried out. Indeed, multivariate analysis of metabolomics-based data is usually performed by applying both supervised (e.g. OPLS-DA) and unsupervised approach (e.g. HCA) (Worley & Powers, 2013). In this regard, OPLS-DA is also able to effectively separate the variation not directly correlated with Y in X matrix (i.e., orthogonal signal correction), considering only the Y-predictive variation (Galindo-Prieto, Eriksson, & Trygg, 2014). Consistently with the unsupervised cluster analysis, the OPLS-DA class prediction model allowed discriminating the different treatments. The corresponding score plot (Figure 1) showed a clear differentiation among the group of counterfeited samples (added with either stamens or tepals) and the authentic ones, made only by styles. Therefore, the chemical fingerprints gained from phenolic compounds and saffron-related metabolites showed a high discrimination potential with regards to saffron adulteration.

Afterwards, the VIP approach was used in order to evaluate the variables importance in projection of the OPLS-DA model. In particular, the VIP selection method was particularly effective for obtaining those variables having the highest discrimination potential into the OPLS score plot. In this regard, the VIP approach identified 77 compounds able to differentiate the authentic saffron from the adulterated ones. The most important metabolites were finally reported in Table 1 together with their individual VIP score (> 1), standard error, Log fold-change, up/down

regulation and ROC AUC values. It is important to underline that the VIP approach following OPLS-DA highlighted the presence of flavonoids (36% of the markers), with 11 flavonols and 7 anthocyanins that were found to be all up regulated into the adulterated samples, then proving that these subclasses of compounds are particularly affected by the misleading practices. Furthermore, lignans accounted for the 16% of the VIP markers (i.e., 12 compounds) and they were all down regulated in authentic samples. Interestingly, zeaxanthin together with the monoterpene glycoside picrocrocin were both down regulated into the adulterated samples, thus suggesting a possible "dilution effect" of these compounds into the styles due to the different inclusion levels of stamens and petals. Finally, phenolic acids accounted for the 13% of the VIP markers (above all hydroxycinnamics) and they were found to be characteristic of the styles because of an overall down regulation following adulteration (Table 1), while the 28% of the remaining VIP markers consisted in other down regulated compounds, such as lower-molecular-weight polyphenols including tyrosols and phenolic terpenes.

Looking at the flavonols proposed by our VIP approach, it is important to underline that the adulteration practice allowed us to observe an increase of both diglycosidic forms of quercetin and kaempferol derivatives (Table 1). In fact, among the 11 flavonols outlined by VIP, we found isomeric forms of both the above-mentioned flavonols (Table 1). In this regard, the glucosyltransferase UGT707B1, isolated from stigmas and tepals of *Crocus sativus*, could be responsible of the trends observed. In fact, this enzyme has been involved in the synthesis of both kaempferol and quercetin sophorosides (Trapero *et al*., 2012), some of the flavonols outlined as VIP markers. Notably, glycosidic forms of kaempferol (i.e. kaempferol 3-*O*-glucoside, kaempferol 3-*O*-sophoroside, kaempferol 3,7-*O*-diglucoside, kaempferol 3,7,4′-*O*-triglucoside, kaempferol 3-*O*-sophoroside-7-*O*-glucoside) have been recently proposed as saffron authenticity markers (Guijarro-Díez *et al.*, 2015). These authors used LC-QTOF-MS followed by chemometric methods to compare ten high-quality saffron samples (category I) with other ten suspected of adulteration. The adequacy of kaempferol glycosides as markers for saffron authenticity were confirmed by another

additional study based on the addition of different percentages (0% to 100%) of gardenia extracts to saffron styles (Guijarro-Díez *et al.*, 2017). In particular, kaempferol glucosides were found to decrease linearly with the corresponding inclusion of gardenia extracts.

Moreover, the tepals of saffron possess a high content of total flavonoids, that was reported to be higher than styles (Jadouali *et al*., 2018). Therefore, the previous findings corroborate our metabolomic results, considering that the most of flavonoids (mostly flavonols) were found to be up regulated into the adulterated saffron. In our experimental conditions, another interesting trend could be noticed for the phenolic subclass of anthocyanins. In fact, the anthocyanins outlined by VIP were found to be all up-regulated into the adulterated saffron samples. Interestingly, it is widely recognized that anthocyanins are a group of visible plant pigments that impart color to flowers, fruits, and other plant organs.

Finally, the potential of these marker compounds to distinguish adulterated *vs* authentic samples was assessed by using receivers operating characteristic (ROC) curves and evaluating their AUC values. ROC curves have been proposed for the evaluation of the sensitivity and specificity based on the false positive/negative rate (Xia et al., 2013). In fact, ROC AUC can be calculated for the robustness of the statistical analysis, being the measure of how well a VIP marker can distinguish between two groups. Looking at our results, most of the markers were characterized by AUC values ranging from 0.9 and 1, thus confirming their importance for discrimination purposes (Table 2). Interestingly, the VIP markers not validated by the ROC curve approach were less then 15%.

3.2. UHPLC-QTOF-MS discrimination of PDO and non-PDO samples

Three Italian PDO saffron products were compared to commercial samples and with an authentic Iranian sample, being this latter the first producer country worldwide. The unsupervised HCA produced from the fold-change-based heat map resulted in three groups: the first group consisted of

four commercials saffron, the second one included the PDO Italian saffron samples and the third one the remaining commercials samples (supplementary material). Afterwards, supervised OPLS-DA was applied in order to predict the variability of the different samples using geographical origin as class membership criterium. The OPLS-DA score plot (Figure 2) showed a high degree of discrimination among groups of samples, and the separation between each geographical group was evident. In particular, PDO Italian saffron possessed a rather different secondary metabolites profile when compared to the other samples. In fact, the other commercial saffron were grouped very close into the OPLS-DA score plot, thus suggesting less distictive phenolic profiles and relatively closer to the samples from Iran. In our experimental conditions, the samples clustered clearly with more than adequate the cross-validation parameters in the OPLS-DA model, being $R^2Y = 0.93$ and $Q^2Y = 0.81$ with a significant CV-ANOVA ($p = 1.03\ 10^{-27}$ for regression). Permutation test cross validation (N=100) could exclude overfitting. On these bases, the model parameters proved that the separation between groups, based on the metabolites annotated, was real and effective.

The following VIP approach was used to identify the best markers of the separation obtained. For this second OPLS model, we selected those markers having a VIP score > 1, as provided in Table 2. Overall, 28 phenolic compounds (classified according to the corresponding class and sub-class) explained the most of variation into the predictive model. In order to examine the potential of these metabolites for food traceability, especially distinguishing the origin of saffron samples, receivers operating characteristic (ROC) curves with AUC (area under the curve) values were produced and are provided in Table 2. According to our results, each marker proposed was characterized by AUC values from 60-100% (Table 2), although those having an AUC > 80-90% can be considered the best classifiers, as reported in literature (Xia et al., 2013). Most of the markers able to discriminate PDO *vs* non-PDO saffron samples were flavonoids (9 compounds) belonging to flavonols and flavones. Moreover the sub-class of phenolic acids (10 compounds) including protocatechuic aldehyde and isomeric forms of hydroxybenzoic acid, was characterized by both high VIP scores and AUC values (0.84). Additionally, other classes possessed a high

discrimination potential, such as lignans and other polyphenols, with sesamol outlined as a good classifier (i.e. AUC > 0.80).

The markers proposed to distinguish PDO *vs* non-PDO samples can be considered very important for discriminating high quality saffron. In this regard, different geographical origins, harvesting conditions and dehydration procedures are able to modify the quality parameters of saffron (Del Campo et al., 2010). For example, some previous studies demonstrated that PDO samples were characterized above all by higher amounts of picrocrocins and crocins, two of the primary saffron quality components, while the commercial ones were mainly abundant in fatty acids. (Cagliani, Culeddu, Chessa, & Consonni, 2015). Interestingly, D'Archivio and co-authors (2016) showed that saffron cultivated in Sardinia (Italy) differs from those produced in central Italy for the content of crocins and other minor metabolites, while Anastasaki et al. (2009) discriminated saffron from different geographical origin by using the volatile compounds profile. Besides, also saffron processing plays a key role for the quality of the product; for example, the drying process is able to produce some trasformation products that could affect its final characteristics (Rubert, Lacina, Zachariasova, & Hajslova, 2016). Nowadays, there is an increasing interest from both producers and consumers towards high-quality food products; in this context, saffron is widely subjected to adulteration or frauds because of its cost. For this reason, the potential of targeted/untargeted high-resolution mass spectrometric approaches (e.g. UHPLC-QTOF-MS) to identify a wide set of compounds related to both geographical origin and authenticity is becoming worthwhile.

4. Conclusions

An untargeted metabolomic approach based on ultra-high-pressure liquid chromatography coupled to quadrupole time-of-flight mass spectrometer (UHPLC/QTOF-MS), followed by multivariate statistics was carried out to discriminate authenticity and traceability of saffron according to their chemical fingerprints. In more detail, phenolics and saffron characteristic

compounds were considered. Interestingly, saffron has demonstrated a diversified, distinctive and complex phenolic profile. Such phenolic diversity could be used for authenticity and traceability purposes. Both the unsupervised cluster analysis and the supervised orthogonal projections to latent structures discriminant analysis (OPLS-DA) allowed discriminating both adulterated and PDO *vs* non-PDO saffron samples. In particular, the adulteration of styles with tepals and stamens was well characterized starting from an inclusion level of 5%. Furthermore, the combination of untargeted MS analysis and chemometrics allowed to discern italian PDO from non-PDO saffron samples. Our approach could be exploited for both traceability and authenticity purposes, considering that the identification of markers to support high-quality products is advisable.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Creation and management of the World Saffron and Crocus Collection (WSCC) has been supported mainly by the research projects: 1) CROCUSBANK: Genetic Resources of Saffron and Allies (018 AGRI GEN RES 870/2004AGRI GEN RES 018 Action) with the financial support of the European Commission and 2) National Program for Conservation and Utilization of Plant Genetic Resources for Agriculture and Food (Project INIA RFP2014-00012) with the financial support of the Spanish government.

Acknowledgements

The authors wish to thank the “Romeo ed Enrica Invernizzi” foundation (Milan, Italy) for kindly supporting the metabolomic facility and Hajar Salehi, Hamadan University, for providing authentic Iranian saffron.

References

Alonso, G. L., Zalacain, A., & Carmona, M. Saffron. K.V. Peter (Ed.), Handbook of herbs and spices, vol. 1, Woodhead Publishing, Cambridge (2012), pp. 469-498, 2nd edition.

Anastasaki, E., Kanakis, C., Pappas, C., Maggi, L., Del Campo, C.P., Carmona, M., Alonso, G. L., & Polissiou, M. G. (2009). Geographical differentiation of saffron by GC–MS/FID and chemometrics. *European Food Research and Technology,* 229(6), 899-905.

Assimiadis, M.K., Tarantilis, P.A., & Polissiou, M.G. (1998). UV-vis FT-Raman and 1H NMR spectroscopies of cis–trans carotenoids from saffron (*Crocus sativus* L.). *Applied Spectroscopy*, 52, 519–522.

Ben Mohamed, M., Rocchetti, G., Montesano, D., Ali, S. B., Guasmi, F., Grati-Kamoun, N., & Lucini, L. (2018). Discrimination of Tunisian and Italian extra-virgin olive oils according to their phenolic and sterolic fingerprints. *Food Research International*, 106, 920-927.

Blasi, F., Rocchetti, G., Montesano, D., Lucini, L., Chiodelli, G., Ghisoni, S., Baccolo, G., Simonetti, M.S., & Cossignani, L. (2018). Changes in extra-virgin olive oil added with *Lycium barbarum* L. carotenoids during frying: Chemical analyses and metabolomic approach. *Food Research International*, 105, 507-516.

Bononi, M., Milella, P., & Tateo, F. (2015). Gas chromatography of safranal as preferable method for the commercial grading of saffron (*Crocus sativus* L.). *Food Chemistry*, 176, 17–21.

Caballero-Ortega, H., Pereda-mirada., R., & Abdullaev, F.I. (2007). HPLC quantification of major active components from 11 different saffron (*Crocus sativus* L.) sources, 100, 1126-1131.

Cagliani, R. L., Culeddu, N., Chessa, M., & Consonni, R. (2015). NMR investigations for a quality

assessment of Italian PDO saffron (*Crocus sativus* L.). *Food control*, 50, 342-348.

Carmona, M., Zalacain, A., Pardo, J. E., López, E., Alvarruiz, A., & Alonso, G. L. (2005). Influence of different drying and aging conditions on saffron constituents. *Journal of Agricultural and Food Chemistry*, 53, 3974–3979.

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*, 112, 110-116.

Del Campo, C. P., Carmona, M., Maggi, L., Kanakis, C. D., Anastasaki, E. G., Tarantilis P. A., Polissiou, M. G., & Alonso, G. M., (2010). Effects of mild temperature conditions during dehydration procedures on saffron quality parameters. *Journal of the science of Food and Agriculture, 90(4),* **https://doi.org/10.1002/jsfa.3877**

Galindo-Prieto, B., Eriksson, L., & Trygg, J. (2014). Variable influence on projection (VIP) for orthogonal projections to latent structures (OPLS). *Journal of chemometrics*, 28(8), **https://doi.org/10.1002/cem.2627**

Gohari, A. R., Saeidnia, S., & Mahmoodabadi, M. K. (2013). An overview on saffron, phytochemicals, and medicinal properties. *Pharmacognosy Review*, 7(13), 61-66.

Guijarro-Díez, M., Nozal, L., Marina, M. L., & Crego, A. L. (2015). Metabolomic fingerprinting of saffron by LC/MS: novel authenticity markers. *Analytical and Bioanalytical Chemistry*, 407, 7197-7213.

Guijarro-Díez, Castro-Puyana, M., Crego, A. L., & Marina, M. L. (2017). A novel method for the quality control of saffron through the simultaneous analysis of authenticity and adulteration markers by liquid chromatography-(quadrupole-time of flight)-mass spectrometry. *Food*

Chemistry, 228, 403-410.

Jaduali, S. M., Atifi, H., Mamouni, R., Majourhat, K., Bouzoubaâ, Z., Laknifli, A., & Faouzi, A. (2018). Chemical characterization and antioxidant compounds of flower parts of Moroccan *Crocus sativus* L. *Journal of the Saudi Society of Agricultural Sciences*, https://doi.org/10.1016/j.jssas.2018.03.007.

Rocchetti, G., Lucini, L., Gallo, A., Masoero, F., Trevisan, M., & Giuberti, G. (2018). Untargeted metabolomics reveals differences in chemical fingerprints between PDO and non-PDO Grana Padano cheeses. *Food Research International*, 113, 407-413.

Rubert, J., Lacina, O., Zachariasova, M., & Hajslova, J. (2016). Saffron authentication based on liquid chromatography high resolution tandem mass spectrometry and multivariate data analysis. *Food chemistry*, 204, 201-209.

Sereshti, H., Heidari, R., & Samadi, S. (2014). Determination of volatile components of saffron by optimised ultrasound-assisted extraction in tandem with dispersive liquid-liquid microextraction followed by gas chromatography– mass spectrometry. *Food Chemistry*, 143, 499–505.

Soffritti, G., Busconi, M., Sánchez, R. A., Thiercelin, J., Polissiou, M., Roldán, M., & Fernández, J. A. (2016). Genetic and Epigenetic Approaches for the Possible Detection of Adulteration and Auto-Adulteration in Saffron (*Crocus sativus* L.) *Spice. Molecules*, 21(3), 343.

Trapero, A., Ahrazem, O., Rubio-Moraga, A., Jimeno, M. L., Gómez, M. D., & Gómez-Goméz, L. (2012). Characterization of a glucosyltransferase enzyme involved in the formation of kaempferol and quercetin sophorosides in *Crocus sativus*. *Plant Physiology*, 159(4), 1335-1354.

Valle Garcia-Rodriguez, M., Serrano-Diaz, J., Tarantilis, P.A., López-Córcoles, H., Carmona, M., & Alonso, G.L. (2014). Determination of saffron quality by high-performance liquid chromatography. *Journal of Agricultural and Food Chemistry*. 62 (32), 8068–8074.

Worley, B., & Powers, R. (2013). Multivariate analysis in metabolomics. *Current Metabolomics*, 1(1), 92-107.

Winterhalter, P., & Straubinger, M. (2000). Saffron- renewed interest in an ancient spice. *Food reviews international*, 16:1, 39-59.

Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics*, 9(2), 280-299.

Yilmaz, A., Nyberg, N. T., Mølgaard, P., Asili, J., & Jaroszewski, J. W. (2010). ^{1}H NMR metabolic fingerprinting of saffron extracts. *Metabolomics*, 6, 511–517.

Zeka, K., Ruparelia, K.C., Continenza, M.A., Stagos, D., Vegliò, F., Arroo, R.R., 2015. Petals of *Crocus sativus* L: as a potential source of the antioxidants crocin and kaempferol. *Fitoterapia*, 107, 128–134.

Zougagh, M., Simonet, B.M., Rios, A., Valcarcel, M., 2005. Use of nonaqueous capillary electrophoresis for the quality control of commercial saffron samples. *Journal of Chromatography A*, 1085, 293–298.

Figure captions

Figure 1. Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) score plot for authentic *vs* adulterated saffron samples. R^2Y and Q^2Y predictive parameters are also reported.

Figure 2. Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) score plot for PDO *vs* non-PDO saffron samples. R^2Y and Q^2Y predictive parameters are also reported.

Table captions

Table 1. Markers having the strongest discrimination potential between authentic saffron and its counterfeit counterparts. Compounds were gained through UHPLC-ESI/QTOF-MS untargeted metabolomics and selected by OPLS-DA discriminant analysis followed by VIP (variables of importance in projection). Different compounds are grouped in functional classes and provided together with VIP score, fold-change analysis, and ROC AUC values.

Table 2. Markers having the strongest discrimination potential between PDO saffron *vs* non-PDO saffron samples. Compounds were gained through UHPLC-ESI/QTOF-MS untargeted metabolomics and selected by OPLS-DA discriminant analysis followed by VIP (variables of importance in projection). Different compounds are grouped in functional classes and provided together with VIP score, fold-change analysis, and ROC AUC values.

Supplementary material

Supplementary table 1. Dataset of identified compounds when considering both adulteration and traceability, with individual abundances and composite spectra (mass-abundance combinations). Supplementary Fig. 1. Besides, cross-validation parameters of both OPLS-DA models built (i.e.

permutation test, Hotelling's T2 and CV-ANOVA) are also provided.

Supplementary figure 1. Non-averaged unsupervised cluster analysis on the untargeted profile of authentic vs adulterated saffron samples (similarity: Euclidean; linkage rule: Ward). Compound's intensity was used to build up heat map, on the basis of which the clusters were generated.

Supplementary figure 2. Non-averaged unsupervised cluster analysis on the untargeted profile of PDO vs non-PDO saffron samples (similarity: Euclidean; linkage rule: Ward). Compound's intensity was used to build up heat map, on the basis of which the clusters were generated.

Figure 1

Click here to download high resolution image

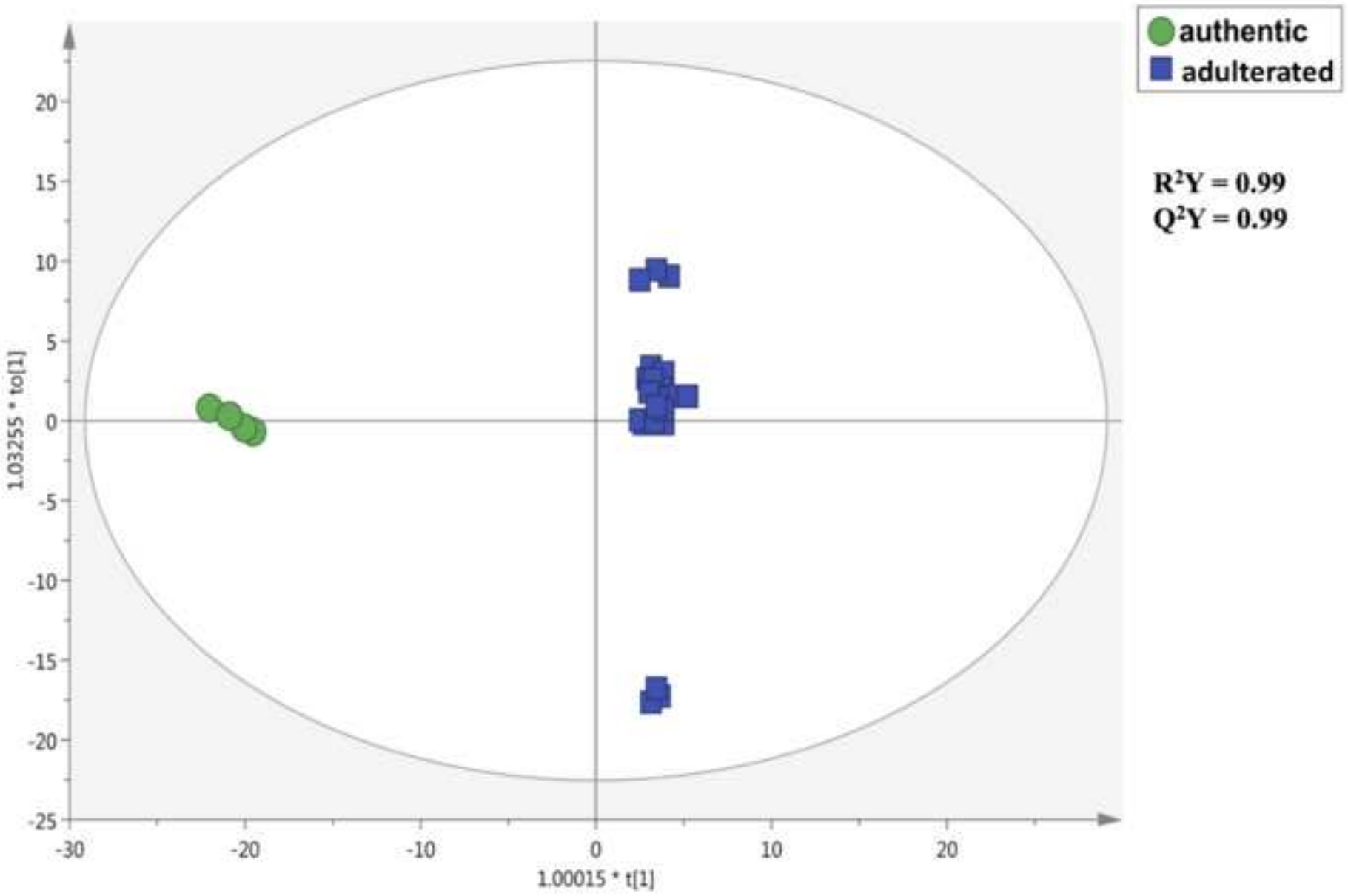

Figure 2

Click here to download high resolution image

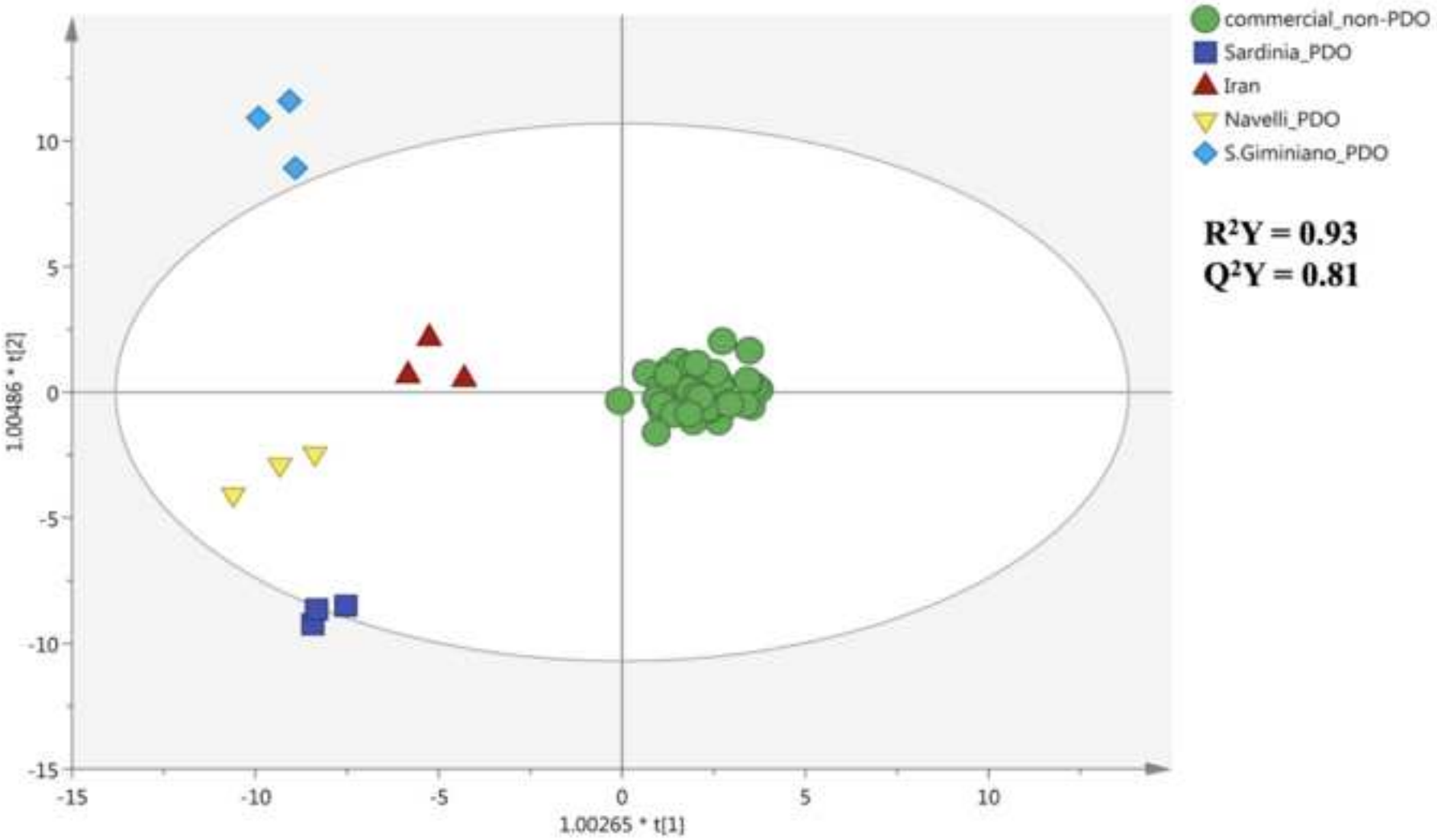

Table 1

Class	Subclass	Marker	VIP score	LogFC	Regulation [Adulterated vs Authentic]	ROC AUC
Flavonoids	Anthocyanins	Pelargonidin	1.24 ± 0.15	4.48	Up	1
		Pelargonidin 3,5-*O*-diglucoside	1.14 ± 0.34	3.11	Up	1
		Delphinidin 3-*O*-(6"-*p*-coumaroyl-glucoside)/ Cyanidin 3-*O*-(6"-caffeoyl-glucoside)	1.09 ± 0.41	5.01	Up	1 1
		Cyanidin 3,5-*O*-diglucoside	1.08 ± 0.42	5.34	Up	1
		Cyanidin 3-O-sophoroside	1.08 ± 0.42	5.34	Up	1
	Flavonols	Kaempferol 3-*O*-sophoroside 7-*O*-glucoside/ Quercetin 3-*O*-glucosyl-rhamnosyl-glucoside/ Quercetin 3-*O*-glucosyl-rhamnosyl-galactoside/ Kaempferol 3,7,4'-O-triglucoside/	1.21 ± 0.18	2.83	Up	1
		Quercetin 3-*O*-galactoside 7-*O*-rhamnoside/ Kaempferol 3-*O*-sophoroside/ Quercetin 3-*O*-rutinoside/ Quercetin 3-O-rhamnosyl-galactoside/ Kaempferol 3,7-O-diglucoside	1.14 ± 0.37	5.76	Up	1
		Quercetin 3-O-(6"-acetyl-galactoside) 7-O-rhamnoside	1.00 ± 0.53	3.23	Up	1
		Spinacetin 3-O-glucosyl-(1-6)-glucoside	1.00 ± 0.54	2.65	Up	1
	Flavanones	Isoxanthohumol	1.46 ± 0.37	- 33.55	Down	1
		Neohesperidin/ Hesperidin	1.26 ± 1.05	- 32.12	Down	ns
		Poncirin	1.22 ± 1.24	-36.63	Down	1
		Didymin	1.22 ± 1.24	-36.61	Down	1
		Eriodictyol 7-O-glucoside	1.08 ± 0.24	2.12	Up	1
	Flavones	Sinensetin/Tangeretin	1.47 ± 0.34	- 39.53	Down	1
		Apigenin 7-O-(6"-malonyl-apiosyl-glucoside)	1.11 ± 0.41	3.40	Up	1
	Chalcones	Xanthohumol	1.46 ± 0.37	-33.55	Down	1
	Dihydrochalcones	Phloretin 2'-O-xylosyl-glucoside	1.38 ± 0.66	- 33.85	Down	1
	Dihydroflavonols	Dihydroquercetin 3-O-rhamnoside	1.08 ± 0.24	2.12	Up	1
Lignans	-	Lariciresinol/Cyclolariciresinol	1.48 ± 0.27	-38.43	Down	1
		Episesamin/Sesamin	1.47 ± 0.33	-39.61	Down	1

		7-Oxomatairesinol	1.47 ± 0.34	-39.53	Down	1
		Sesamolinol	1.47 ± 0.33	-39.52	Down	1
		Todolactol A	1.47 ± 0.37	-42.88	Down	1
		Pinoresinol/Matairesinol	1.47 ± 0.37	-42.88	Down	1
		7-Hydroxysecoisolariciresinol	1.24 ± 0.18	4.02	Up	1
		Anhydro-secoisolariciresinol	1.06 ± 0.23	-14.46	Down	0.96
		Lariciresinol-sesquilignan	1.24 ± 0.18	4.02	Up	1
		Arctigenin	1.00 ± 0.95	-25.27	Down	ns
Monoterpene glycosides	-	Picrocrocin	1.47 ± 0.34	-39.30	Down	1
Carotenoids	-	Zeaxanthin	1.18 ± 1.24	-39.12	Down	1
Phenolics acids	Hydroxycinnamics	Ferulic acid 4-O-glucoside	1.47 ± 0.36	-35.65	Down	1
		Feruloyl glucose	1.47 ± 0.36	-35.65	Down	1
		Sinapine	1.44 ± 0.58	-36.06	Down	1
		24-Methylcholestanol ferulate	1.26 ± 1.09	-33.54	Down	1
		p-Coumaroyl malic acid	1.22 ± 0.29	6.29	Up	1
		Stigmastanol ferulate	1.10 ± 1.45	-35.49	Down	1
		24-Methyllathosterol ferulate	1.10 ± 1.34	-29.70	Down	ns
	Hydroxyphenylacetics	Homoveratric acid	1.20 ± 1.24	-29.21	Down	ns
		1,2-Diferuloylgentiobiose	1.45 ± 0.45	-35.63	Down	1
		1,4-Naphtoquinone	1.10 ± 1.43	-33.73	Down	ns
Other compounds	Tyrosols	3,4-DHPEA-AC	1.20 ± 1.24	-29.21	Down	ns
		Hydroxytyrosol	1.01 ± 0.50	-5.47	Down	0.94
	Curcuminoids	Demethoxycurcumin	1.47 ± 0.29	-34.63	Down	1
		Bisdemethoxycurcumin	1.24 ± 0.16	2.12	Up	1
	Methoxyphenols	3/4-Methylcatechol	1.45 ± 0.16	3.96	Up	1
		Guaiacol	1.45 ± 0.16	3.96	Up	1
	Alkyphenols	5-Nonadecylresorcinol	1.42 ± 0.68	-33.41	Down	1
		5-Heptadecylresorcinol	1.40 ± 0.29	-10.40	Down	1
		5-Heneicosenylresorcinol	1.23 ± 1.09	-31.77	Down	ns
		5-Tricosylresorcinol	1.21 ± 1.14	-36.39	Down	1
		5-Tricosenylresorcinol	1.20 ± 1.21	-31.34	Down	ns

	5-Heneicosylresorcinol	1.05 ± 1.51	-30.50	Down	ns
	4-Ethylphenol	1.02 ± 0.28	-14.77	Down	0.92
Naphtoquinones	Juglone	1.32 ± 0.28	-35.89	Down	1
Phenolic terpenes	Thymol	1.14 ± 0.16	-3.12	Down	0.97
	Carvacrol	1.14 ± 0.16	-3.12	Down	0.97
	Epirosmanol/Rosmanol	1.13 ± 0.21	-9.80	Down	0.97
	Rosmadial	1.06 ± 0.23	-14.46	Down	0.96
Hydroxyphenylpropenes	Acetyl eugenol	1.11 ± 0.34	-9.94	Down	1
Alkylmethoxyphenols	4-Ethylguaiacol	1.02 ± 0.19	-8.08	Down	0.93
Other	Catechol	1.30 ± 0.99	-35.32	Down	1

Table 2

Phenolic class	Phenolic subclass	Marker non-PDO *vs* PDO	VIP score	ROC AUC
Flavonoids	Anthocyanins	Pelargonidin 3-*O*-(6"-succinyl-glucoside)	1.10 ± 0.23	0.89
	Flavanones	Isoxanthohumol	1.12 ± 0.38	0.97
	Flavones	Nobiletin	1.32 ± 0.13	0.69
		Jaceosidin	1.29 ± 0.24	0.94
		6-Hydroxyluteolin	1.29 ± 0.22	0.92
	Flavonols	3-Methoxysinensetin	1.32 ± 0.13	0.69
		3,7-Dimethylquercetin	1.29 ± 0.24	0.94
		Quercetin	1.28 ± 0.22	0.92
	Isoflavonoids	6"-*O*-Malonylglycitin	1.16 ± 0.19	0.93
Phenolic acids	Hydroxybenzaldehydes	Protocatechuic aldehyde	1.43 ± 0.24	0.84
		4-Hydroxybenzaldehyde	1.07 ± 0.20	0.78
		Vanillin	1.05 ± 0.18	0.85
	Hydroxybenzoic acids	2/3/4-Hydroxybenzoic acids	1.43 ± 0.25	0.84
		Benzoic acid	1.07 ± 0.20	0.78
	Hydroxycinnamic acids	Sinapine	1.36 ± 0.24	0.95
		p-Coumaroyl malic acid	1.28 ± 0.22	0.92
		p-Coumaric acid	1.00 ± 0.36	0.61
		Cinnamoyl glucose	1.00 ± 0.12	0.81
	Hydroxyphenylacetic acids	4-Hydroxyphenylacetic acid	1.05 ± 0.18	0.85
Lignans	-	Sesamol	1.43 ± 0.24	0.84
		Arctigenin	1.25 ± 0.14	0.72
		Trachelogenin	1.19 ± 0.15	0.93
		Medioresinol	1.19 ± 0.15	0.93
Other polyphenols	Alkylphenols	5-Heptadecylresorcinol	1.01 ± 0.16	0.81
	Tyrosols	Hydroxytyrosol 4-*O*-glucoside	1.01 ± 0.15	0.82
	Phenolic glycosides	Phlorin	1.09 ± 0.24	0.67

Supplementary table 1

Click here to download Supplementary material for online publication only: ST1.xlsx

*Graphical Abstract

Click here to download high resolution image