SUPPLEMENTARY MATERIAL

Discrimination of almonds (*Prunus dulcis*) geographical origin by minerals and fatty acids profiling

Diana Amorello^a, Santino Orecchio^{a,*}, Andrea Pace^{a,b} and Salvatore Barreca^{a,b,c}

a) Dipartimento di Scienze e Tecnologie Biologiche, Chimiche e Farmaceutiche. Università di Palermo, Viale delle Scienze, Ed. 17, I-90128, Palermo, Italy.

b) Istituto Euro Mediterraneo di Scienza e Tecnologia (IEMEST), Via Emerico Amari 123, 90139 Palermo, Italy

c) Consorzio di Ricerca per lo Sviluppo di Sistemi Innovativi Agroambientali (CORISSIA) Via Libertà 203, 30128 Palermo, Italy

diana.amorello@unipa.it

santino.orecchio@unipa.it

andrea.pace@unipa.it

salvatore.barreca@unipa.it

Abstract

Twenty-one almond samples from three different geographical origins (Sicily, Spain and California) were investigated by determining minerals and fatty acids compositions. Data were used to discriminate by chemometry almond origin by Linear Discriminant Analysis. With respect to previous PCA profiling studies this work provides a simpler analytical protocol for the identification of almonds geographical origin. Classification by using mineral contents data only was correct in 77 % of the samples, while, by using fatty acid profiles, the percentages of samples correctly classified reached 82%. The coupling of mineral contents and fatty acid profiles lead to an increased efficiency of the classification with 87% of samples correctly classified.

Key words: Almond, Minerals, Fatty acids, Geographical origin

* Corresponding author: Tel: +39-091-23897968; E-mail: santino.orecchio@unipa.it

1. Experimental

1.1 Samples

Twenty one almond samples from three different origins were studied. Almond nut samples were obtained from different local market located in Palermo (Italy). Italian cultivations of almonds are located mainly in Sicily in semi-arid climatic conditions, often cultivated without irrigation. Sicilian samples were sampled on Sicilia country, while California and Spanish samples were purchased from market shops. Almond nut were stored in refrigerator at 4 °C and protected from light by aluminium foil to avoid any kind surface and photodegradation of their molecular constituent. In detail, almond samples were composed by: ten from Italy (Sicily), nine from United State of America (California) and two from Spanish (Table S1).

1.2 Reagents

All chemicals were of analytical grade and were purchased from Sigma–Aldrich Corporation (St. Louis, MO) unless specified otherwise. The Fatty Acid Methyl Esters standard (FAME mixture) containing 37 fatty acids and internal standard Octanoic d₁₅ methyl esters were obtained from Supelco, Bellefonte.

Standard solution containing 26 elements was purchased from Ultra Scientific, USA (catalogue number: 191 IQC-026, Lotno. J00022). The diluted standard solutions were prepared daily (Amorello et al. 2015). The reagents used throughout were analytical grade (Carlo Erba, Milano, Italy) and all solutions were prepared in Milli-Q water. Concentrated HNO₃ (Suprapur Carlo Erba, Milano, Italy) was used to mineralize the samples and to acidify the standard solutions.

1.3 Water Content Analysis

About 10 g of sample was dried at 110 °C for one night. The water content was determined by weight loss and was used to correlate all the results with dry weight.

1.4 Determination of fatty acids

Fatty acid composition for almond oils was determined using a modified fatty acid methyl ester method (Llorent-Martínez et al. 2014). The lipids were extracted by automated Soxhlet (Buchi B-811) in *warm mode* by using petroleum ether. To convert triglycerides extracted in to fatty acid methyl esters, 50 µL of almond extract were added at 2 mL of KOH (2 M) in methanol. Capped sample container was placed in a 55 °C water bath for 30 min with occasional shaking. After derivatization procedure, the methyl esters were transferred in hexane solution. The solution was concentrated in a rotary evaporator under reduced pressure at 50 °C, transferred to a GC vial, dried

by using weak nitrogen flow and recomposed by addition of an internal standard solution containing octanoic d₁₅ methyl esters at 20 ppm. The fatty acid methyl esters were analysed by a slb 5 ms column (30 m 0.25 mm i.d. 0.25 mm) (Supelco) installed on a Shimadzu GC-MS 2010 QP Plu and operating in the selected ion mentoring (SIM) mode. Fatty acid methyl ester retention times and mass spectra were checked by using standards solution (Supelco 37 FAME mix cat. N° 47885-U). The most abundant ion was used for quantification and two other ions were additionally used for confirmation. Analysis were done in triplicate.

1.5 Determination of mineral contents

A microwave oven (Milestone model MLS-1200 Mega, Milestone Laboratory Systems, Italy) with rotor of high pressure (up to 100 bar) was used to mineralization procedures. About 1000mg of previously homogenized samples were weighted, transferred inside Teflon vessels and mixed with 1 mL of 30% H₂O₂ (Fluka, Milano) and 3 mL of 69 % HNO₃ (Fluka, Milano) Subsequently, the vessel was placed on the microwave turntable to digest the samples (Amorello et al. 2015; Orecchio et al. 2015;). The instrumental conditions used for the microwave digestion were: 1 min at 250 W, 1 min at 0 W, 5 min at 250 W, 5 min a 450 W, 3 min at 600 W and 5 min at 300 W. After digestion, the clear, colourless solution was transferred into a volumetric flask and brought to volume with Milli-Q water ($R > 20 \text{ M}\Omega \text{ cm}^{-1}$) (Merck Millipore).

Minerals (Na, K, Li, Sr, Ca, Mg, Al) content in solution was measured by using an Optima 7000 ICP-OES spectrometer with dual view configuration (Perkin Elmer, Waltham, Ma, USA coupled with WinLab32 software package (Orecchio et al. 2014; Orecchio et al. 2015). Instrumental condition are shown in table S2.

Seven elements were determined in each sample. The inductively coupled plasma optical emission spectrometry (ICP-OES) analysis of trace elements was performed in axial viewing mode. The quantitative analysis was carried out at two different spectral lines for each element (Orecchio, 2013). The data of the elements reported in this paper have been calculated considering the average of the three concentrations obtained at the two wavelengths that, for all analytes, differed by less than 5%.The repeatability of the whole method calculated as relative standard deviation (RSD %) of three independent analysis of the same sample was less than 8%.

1.6 Statistical analysis

Correlations between fatty acids were determined by using Pearson correlation coefficients, which describes the strength of the linear relationship between two quantitative variables, at $P < 0.05$, 0.01 and 0.001, respectively. Pearson correlation coefficients' were calculated using the statistic program

XLSTAT Software (XLSTAT, 2008, Addinsoft, New York, NY). The classification and discrimination of almond samples using fatty acid profiles and metal content analysis were achieved by discrimination analysis using XLSTAT.

Linear discriminant analysis (LDA) is the most frequently used technique (linear and parametric) among the supervised pattern recognition methods (Todeschini, 1992; Silva et al. 2013). LDA is used to find the linear combination of features and the resulting combination may be used as a linear classification. The principle of LDA is based on the determination of linear discrimination functions which brings out clearly the ratio between class variance and reduces the ratio of within-class variance. According to Berrueta and co-workers in LDA, the classes are supposed to follow a multivariate normal distribution and be linearly separated. LDA is also considered, as PCA and the number of principal component factors crucial to the performance of LDA discrimination model. The discriminant analysis was conducted to three clusters (Sicilian, California and Spain) by using three different information types (see supplementary material). In detail, the first LDA analysis was conducted by considering only fatty acid profiles (Figure S2), the second LDA was conducted by considering metal ions content (Figure S3), while the third LDA was conducted coupled fatty acid profiles and mineral content (Figure S4). In all cases, data are projected on the reduced space of the first two discriminant functions evaluated on mineral (Figure S3), fatty acids (Figure S2) and fatty acid coupled with mineral (Figure S5). Figures S2, S3 and S4 show LDA analysis obtained by using the three different data sets. By using the calculated discriminant functions, samples were correctly classified at different percentages. In detail, using mineral contents the 77 % of samples were correctly classified; using fatty acid profiles the percentages of samples correctly classified were about 82%, while coupling mineral contents and fatty acid profiles, the 87% of samples were correctly classified (Tables S5, S6 and S7). Moreover, in the case of using mineral contents and fatty acid profiles, two uncorrected classified samples were placed by LDA analysis in border line, highlighting linear discrimination improvement.

Table S1. Sample origins

Cs25	Spanish	Ci3	Italy (Sicily)
Cs2	Spanish	Ci23	Italy (Sicily)
C26	United State of America California	Ci24	Italy (Sicily)
Cc13	United State of America California	Ci11	Italy (Sicily)
Cc1	United State of America California	Ci18	Italy (Sicily)
Cc6	United State of America California	Ci19	Italy (Sicily)
Cc7	United State of America California	Ci16	Italy (Sicily)
Cc9	United State of America California	Ci10	Italy (Sicily)
Cc5	United State of America California	Ci4	Italy (Sicily)
Cc27	United State of America California	Ci12	Italy (Sicily)
C17	United State of America California		

Table S2. - ICP-OES operating conditions

RF power (W)	1300
Sample uptake flow rate (mL min⁻¹)	1.5
Gas flow rates (L/ min⁻¹)	Auxiliary: 0.2; Nebulizer: 0.8; Argon: 15
Viewing mode	radial/axial

Table S3. Fatty acid composition (%) of the analyzed almonds samples

Sample	Oleic	Miristic	Palmitoleic	Palmitic	Linoleic	Linolenic	Oleico	Stearic	Arachidonic
Cs25	67.6	0.10	0.15	4.0	25.2	2.34	67.6	0.56	0.02
Cs2	81.2	0.005	0.23	3.9	13.2	0.01	81.2	1.5	0.01
C26	78.4	0.01	0.10	3.4	17.3	0.14	78.4	0.62	0.01
Cc13	63.8	0.02	0.72	7.7	24.7	0.0050	63.8	2.9	0.11
Cc1	73.9	0.02	0.38	4.5	19.7	0.0050	73.9	1.5	0.05
Cc6	70.2	0.04	0.40	6.1	22.1	0.10	70.2	1.0	0.01
Cc7	64.4	0.0050	0.40	7.4	21.2	0.005	64.4	6.2	0.34
Cc9	60.8	0.05	1.4	7.6	25.4	0.005	60.8	4.6	0.26
Cc5	58.3	0.08	2.3	7.8	27.4	0.12	58.3	3.8	0.20
Cc27	73.9	0.02	0.38	4.5	19.7	0.005	73.9	1.5	0.05
C17	85.7	0.01	0.20	2.9	10.0	0.005	85.7	1.2	0.01
Ci3	62.9	0.03	1.43	7.3	22.2	0.005	62.9	5.7	0.35
Ci23	73.7	0.01	0.31	4.6	18.8	0.04	73.7	2.5	0.02
Ci24	75.6	0.01	0.35	4.4	17.7	0.01	75.6	1.9	0.02
Ci11	52.7	0.02	0.96	8.9	31.1	0.15	52.7	5.84	0.32
Ci18	79.7	0.01	0.25	3.6	13.9	0.07	79.7	2.4	0.03
Ci19	61.5	0.02	0.21	7.22	26.6	0.82	61.5	3.4	0.20
Ci16	67.8	0.01	0.25	4.11	23.5	0.08	67.8	4.1	0.09

Ci10	65.6	0.01	0.82	7.00	21.7	0.01	65.6	4.6	0.24
Ci4	60.1	0.0050	0.52	8.62	26.3	0.16	60.1	4.1	0.22
Ci12	57.9	0.01	0.33	9.57	30.0	0.08	57.9	1.3	0.03

Table S4. Mineral composition (mg/Kg d.w.) of the analyzed almonds samples

Sample	Na	K	Li	Ca	Sr	Mg	Al
Cs25	62	208	0.68	154	0.22	197	11
Cs2	403	2703	2.5	1584	2.7	2269	10
C26	279	334	2.2	1825	2.1	2594	11
Cc13	466	1941	0.49	1598	2.2	2659	15
Cc1	209	2008	2.1	1987	2.7	2438	16
Cc6	423	1696	5.9	1813	2.5	2940	11
Cc7	228	2054	2.1	1633	2.7	3047	25
Cc9	260	1976	5.6	1582	2.4	2587	13
Cc5	440	1528	7.7	1895	2.0	3170	13
Cc27	785	979	3.3	3830	6.3	5286	22
C17	611	6236	11	3478	5.3	4735	44
Ci3	407	2652	3.7	1853	9.0	2184	27
Ci23	1494	3612	11	5686	9.1	4609	125
Ci24	467	2179	1.6	1501	1.6	2820	120
Ci11	367	2538	2.6	2404	2.3	2422	10
Ci18	608	1611	5.7	3407	4.6	2371	40
Ci19	203	773	2.2	741	1.6	663	11
Ci16	379	1229	6.4	1354	6.0	1145	23
Ci10	326	2226	5.3	1317	1.7	2228	10
Ci4	298	2562	1.7	1988	2.5	2783	10
Ci12	308	2484	8.3	2387	3.0	2512	11

Table S5. Statistical data about LDA obtained using mineral contents

	California	Sicily	Spanish	Summ
California	7	2	0	9
	33,33%	9,52%	0,00%	42,85%
Sicily	2	8	0	10
	9,52%	38.09%	0,00%	47,61%
Spanish	0	1	1	2
	0,00%	4.76%	4.76%	9,52%
Sum	9	11	1	21
	42,85%	52,32%	4,76%	100,00%
Error rate				22,73

Table S6. Statistical data about LDA obtained using fatty acids

	California	Sicily	Spanish	Sum
California	8	1	0	9
	38.09%	4.76%	0,00%	42,85%
Sicily	2	8	0	10
	9,09%	36,36%	0,00%	47,61%
Spanish	1	0	1	2
	4.76,00%	0,00%	4,76%	9,52%
Sum	11	9	1	21
	52,32%	42.85%	4,76%	100,00%
Error rate				18,18

Table S7. Statistical data about LDA obtained using fatty acids and mineral contents

	California	Sicily	Spanish	Sum
California	8	1	0	9
	38.09%	4.76%	0,00%	42,85%
Sicily	0	10	0	10
	0,00%	47,61%	0,00%	47,61%
Spanish	0	1	1	2
	0,00%	4,76%	4,76%	9,52%
Sum	10	10	1	21
	47,61%	47.61%	4,76%	100,00%
Error rate				13,73

Figure S1 - Relationship between the oleic and linoleic acids

Figure S2 – LDA analysis performed by using fatty acids content

Figure S3 – LDA analysis performed by using metals content

Figure S4 – LDA analysis performed by using metals and fatty acids content

2. References

- Amorello D, Barreca S, Orecchio S, Ferro S. 2015. Platinum in indoor settled dust matter (homes and cars). *Microchemical.* 123:76–83.
- Berrueta LA, Alonso-Salce RM, Heberger K. 2007. Supervised pattern recognition in food analysis *Journal of Chromatography A.* 1158:196–214.
- Llorent-Martínez EJ, Domínguez-Vidal A, Rubio-Domene R, Pascual-Reguera MI, Ruiz-Medina A, Ayora-Cañada MJ, 2014. Identification of lipidic binding media in plasterwork decorations from the Alhambra using GC–MS and chemometrics: Influence of pigments and aging. *Microchemical.* 115:11–18.
- Orecchio S. 2013. Microanalytical characterization of decorations in handmade ancient floor tiles using inductively coupled plasma optical emission spectrometry (ICP-OES). *Microchemical.* 108:137-150.
- Orecchio S, Amorello D, Raso M, Barreca S, Lino C, Di Gaudio F. 2014. Determination of trace elements in gluten-free food for celiac people by ICP-MS, *Microchemical.* 116: 163–172.
- Orecchio S, Culotta L 2015. Assessment of quality of air in Palermo by chemical (ICP-OES) and cytological analyses on leaves of *Eucalyptus camaldulensis*. *Environmental Science and Pollution Research.* 22: 1891–1905.
- Silva CS, Borba F.de Souza Lins, Pimentel MF, Pontes MJ Coelho, Honorato R S., Pasquini C. 2013. Classification of blue pen ink using infrared spectroscopy and linear discriminant analysis. *Microchemical.* 109: 122–127.
- Todeschini R. 1992. Linear discriminant classification tree: A user-driven multicriteria classification method. *Chemometrics and Intelligent Laboratory Systems* 16:25-35.