LWT - Food Science and Technology 74 (2016) 387–395

Contents lists available at ScienceDirect

LWT - Food Science and Technology

journal homepage: www.elsevier.com/locate/lwt

Characterisation of *Fragaria vesca* fruit from Italy following a metabolomics approach through integrated mass spectrometry techniques

Gilda D'Urso [a], Mariateresa Maldini [b], Giorgio Pintore [b], Luigi d'Aquino [c], Paola Montoro [a, *], Cosimo Pizza [a]

[a] Department of Pharmacy, Università of Salerno, Via Giovanni Paolo II, 84084 Fisciano, Salerno, Italy
[b] Department of Chemistry and Pharmacy, University of Sassari, Via F. Muroni, 23/b, 07100, Sassari, Italy
[c] ENEA, Portici Research Centre, Piazzale E. Fermi, 80055 Portici, Napoli, Italy

ARTICLE INFO

Article history:
Received 8 January 2016
Received in revised form
25 July 2016
Accepted 27 July 2016
Available online 30 July 2016

Keywords:
Wild strawberry
Targeted and *untargeted* analysis
LC-ESI-Orbitrap-MS/MS
LC-ESI-QTrap-MS/MS
Polyphenol

ABSTRACT

The phytochemical content of *Fragaria vesca* L. (wild strawberry) grown in Campania Region (South Italy) was investigated. Wild and cultivated fruit were collected from different geographic areas and from bothautochthonous and non-autochthonous germplasm. Extracts were submitted to untargeted liquid chromatography coupled to high resolution mass spectrometry (LC–ESI-Orbitrap–MS) for metabolite profiling. 39 different phenolic compounds have been detected and identified in the methanolic extract of spontaneous and cultivated strawberries from different origins. Targeted quantitative analyses of selected known metabolites were performed using UHPLC interfaced to Q-Trap mass spectrometer in Multiple Reaction Monitoring mode. Principal Component Analysis (PCA) was applied to both untargeted and targeted results. PCA was applied to the untargeted profiling data indicating that fruits from locations with different pedoclimatic conditions can be discerned. Results from quantitative analysis indicate that autochthonous strawberries display significantly higher levels of anthocyanins, especially cyanidin 3-*O*-glucoside, and other phenolic compounds, like gallic acid and procyanidin B1, compared to the one from non-autochthonous germplasm. The targeted and untargeted metabolite profiling approach used in this work is proposed for the assessment of geographical origin of the berries.

1. Introduction

Fragaria vesca L. (Fam. *Rosaceae*), commonly known as "wild strawberry" or "woodland strawberry", is a diploid (2n = 2x = 14) plant species whose ancestor is believed to be an ancestral genome donor to the octoploid, widely cultivated worldwide, strawberry *Fragaria* X *ananassa* Duch (Sun, Liu, Yang, Slovin, & Chen, 2014). *F. vesca* is considered to be cosmopolite and is very common in the underwood in temperate areas (Yildirim & Turker, 2014).

Under the botanical point of view, the fruit of *F. vesca* is a false fruit, although it is generally referred to as a "berry". The false fruit is the edible part and is very appreciated by consumers for its taste and, more recently, also for its nutraceutical properties. *F. vesca* berries are, in fact, rich in fibres and bioactive compounds such as vitamin C and polyphenols. The most significant health benefits of strawberry are attributed to the (poly)phenolic compounds, represented by phenolic acids, flavonoids (flavonols, flavan-3-ols, anthocyanins, proanthocyanidins) (Del Bubba, Checchini, Chiuminatto, Doumett, Fibbia & Giordanic, 2012; Sun et al., 2014), that are considered as protective micronutrients and whose content is, therefore, assumed as a quality parameter for edible fruit. Antioxidant and/or antiradical activities of spontaneous *F. vesca* berries have been reported in literature, indicating a high activity of spontaneous strawberry fruits (Cheel, Theoduloz, Rodriguez, Caligari, & Schmeda-Hirschmann, 2007; Pellegrini et al., 2003).

Southern Italy is an important producer of *F. vesca* berries and in the Campania Region, particularly in the Alburni and the Alto Sele Areas, where spontaneous strawberries are officially recognized as a traditional food product under the name "Fragolina degli Alburni". In these areas, people traditionally transplant rooted stolons collected from underwood into their gardens, so starting a domestic cultivation of the wild strawberries whose fruits are then

* Corresponding author.
E-mail address: pmontoro@unisa.it (P. Montoro).

http://dx.doi.org/10.1016/j.lwt.2016.07.061
0023-6438/

used in the production of traditional sweets. Unfortunately, the local wild strawberry germplasm blossoms only once per year, thus supplying a fruit yield that is concentrated in a short harvesting period. In order to cope with an increasing demand of wild strawberries from consumers, local farmers are, therefore, encouraged to introduce non-native strawberry germplasm that is able to blossom several times per year, so supplying greater fruit amounts and fresh fruit for a longer period.

Geographical origin of the plant stock is a crucial feature in food quality, since it is commonly recognized that local differences, in addition to agronomic practices, can heavily affect the chemical composition of the yield. In order to emphasize the geographical component in the food quality, Italian institutions established "protected denominations of geographical origin" (IGP) for foods that display local peculiarities, although chemical markers are not always available to discern among food stocks coming from different areas. Recent applications of Mass Spectrometry-based strategies in food-omics include metabolomics studies for food quality and geographical origin assessment (Herrero, Simó, García-Cañas, Ibáñez & Cifuentes, 2012). Moreover, it has also been reported that changes in the level of secondary metabolites are, to some extent, determined by environmental factors and regulated at the genetic level (Ramakrishna & Ravishankar, 2011). In the research focusing on berries, both targeted and untargeted metabolic profiling have so far been used in studying phytochemical diversity and relationships between cultivars (Brown, Murch, & Shipley, 2012; Carvalho, Franceschi, Feller, Palmieri, Wehrens & Martens, 2013; McDougall, Martinussen, & Stewart, 2008). To our knowledge, the relation between the pattern of bioactive compounds of strawberry fruits from different germplasm and/or cultivation areas has not yet been investigated. The aim of this work was to compare the phytochemical content of *F. vesca* fruit collected in different geographical areas growing indifferent conditions and obtained from different germplasm sources using a LC-ESI–Orbitrap–MS metabolic profiling approach. The main focus was to identify differences in the semi-polar metabolite pool, that could be useful in discriminating the geographic origin of wild strawberries.

2. Materials and Methods

2.1. Fruit samples

F. vesca fruits were collected in July 2013 in Petina (Alburni Mountains Area, Campania Region, South Italy) and in Sarno (Sarno River Plain Area, Campania Region, South Italy), from either wild populations, spontaneously growing in the underwood, and cultivated populations, and from either autochthonous and non-autochthonous germplasm. The two sampling locations were about 90 km far from each other, the former located at more than 650 m above the sea level, characterized by clay-rich soils and milder Mediterranean microclimate conditions, the latter located at about 30 m above the sea level, characterized by sandy-rich soils and heavier Mediterranean microclimate conditions. Fruits were collected at the fully ripe stage. Samples, each one consisting of 10 strawberries, were collected in triplicate, thus resulting in 15 biological samples, classified in 5 groups as in the following: SPA – Spontaneous, Petina, Autochthonous; SSA – Spontaneous, Sarno, Autochthonous; CPA – Cultivated Petina, Autochthonous; CPN – Cultivated, Petina, Non-Autochthonous; CSN – Cultivated, Sarno, Non-Autochthonous. Samples CSA (Cultivated Sarno Autochthonous) were not available since autochthonous wild strawberry is currently not cultivated in Sarno River Plain Area.

2.2. Reagents and solvents

Standards as procyanidin B1, procyanidin B2, isoquercitrin, (−) epicatechin, (+) catechin, gallic acid, *p*-coumaric acid, phloridzin, cyanidin-3-*O*-glucoside, delphinidin-3-*O*-glucoside, pelargonidin-3-*O*-glucoside, pelargonidin-3-*O*-rutinoside, were purchased from Extrasinthese (Lyon, France). Formic acid and methanol for extraction were purchased from VWR international PBI S. r.l. (Milano, Italy). Acetonitrile, methanol, water and formic acid (LC-MS grade) were purchased from Merck (Darmastadt, Germany).

2.3. Sample preparation

Fruits were extracted immediately after collection to prevent degradation of secondary metabolites. The extraction was performed following the method proposed by Pavlovic et al. 2013 with slight modifications. Briefly, five gram of fresh fruits, snap frozen in liquid nitrogen and ground to a fine powder, were extracted for 24 h with 20 mL of methanol acidified with 1 mL/L of formic acid at room temperature (thermostatically controlled at 18–25 °C), preserving the samples away from light, then 0.2 mL of filtered extracts were dried under N_2 flow and diluted in 1 mL of water acidified with 1 mL/L of formic acid. Extract samples 0.01 mL were used for LC-MS analysis. All samples were analysed in duplicate.

2.4. LC-ESI-Orbitrap-MS analysis

To investigate the main markers specific to each fruit, an LC-ESI-Orbitrap-MS method was developed. Methanol extracts were used for all the analysis. All experiments were performed using a Thermo Scientific liquid chromatography system constituted of a quaternary Accela 600 pump and an Accela auto sampler, connected to a linear Trap-Orbitrap hybrid mass spectrometer (LTQ-Orbitrap XL, Thermo Fisher Scientific, Bremen, Germany) combining linear trap quadrupole (LTQ) and Orbitrap mass analyser with electrospray ionization (ESI). Separation was performed on a XSelect CSH C18 reversed phase column (2.1 × 150 mm 3.5 μm) (Waters, Milford, Massachusetts). The mobile phase consisted of solvent A (water+ 1 mL/L of formic acid) and solvent B (acetonitrile/water (8:2 mL/mL) + 1 mL/L of formic acid). A linear gradient program at a flow rate of 0.200 mL/min was used: 0–50 min, from 0 to 40% (B); 51–56 min, 100% (B); then 0% (B) for 5 min. The mass spectrometer was operated in negative ion mode. ESI source parameters were as follows: capillary voltage −12 V; tube lens voltage −121.47 V; capillary temperature 280 °C; Sheath and Auxiliary Gas flow (N_2) 30 and 5(arbitrary units), Sweep gas 0 (arbitrary units) Spray voltage 5 V. MS spectra were acquired by full range acquisition covering *m/z* 200–1600 (Resolution:30,000). For fragmentation study, a data dependent scan was performed, selecting precursor ions corresponding to most intense peaks in LC-MS analysis. Data were acquired also in positive ion mode following the same chromatographic procedure. ESI source parameters were as follows: capillary voltage 49 V; tube lens voltage 120 V; capillary temperature 280 °C; Sheath and Auxiliary Gas flow (N_2) 30 and 5(arbitrary units), Sweep gas 0 (arbitrary units) Spray voltage 5 V. MS spectra were acquired by full range acquisition covering *m/z* 200–1600. A fragmentation study was applied also in positive mode, by Depending Data Scan, selecting the two most intense ions in the MS profile.

Phenolic compounds were identified according to the corresponding spectral characteristic fragmentation and retention time, by comparison with data from literature. Xcalibur software version 2.1 was used for instrument control, data acquisition and data analysis.

2.5. ESI-QTrap-MS and ESI-QTrap-MS/MS analyses

Full scan ESI-QTrap-MS and collision induced dissociation (CID) ESI-QTrap-MS/MS analyses of standards were performed on an ABSciex (Foster City, CA, USA) API4000 Q-Trap spectrometer. The analytical parameters were optimized by infusing a standard solution of each compound [1 mg/L in methanol/water (1:1 mL/mL)] into the source at a flow rate of 0.01 mL/min. Data were acquired in the positive or negative ion MS and MS/MS modes depending of the standard compound.

2.6. LC–ESI-QTrap-MS and LC–ESI-QTrap-MS/MS analyses

Quantitative on-line UHPLC-ESI-QTrap-MS/MS (MRM) analyses were performed using an UHPLC system interfaced to an ABSciex (Foster City, CA, USA) API4000 Q-Trap instrument in ion trap mode. LC analyses were conducted using a system equipped with a Flexar UHPLC AS system (Perkin-Elmer, USA) consisting of degasser, Flexar FX-10 pump, auto sampler and PE 200 column oven. Samples (0.005 mL) were injected into a XSelect HSS T3 column (Waters, Milford, Massachusetts) (100 × 2.1 mm, 2.5 μm). Mobile phase A was H_2O containing 1 mL/L of formic acid while mobile phase B was acetonitrile containing 1 mL/L of formic acid. Elution for the analyses of anthocyanins was carried out at 41 °C according to the following flow and solvent gradient: 0–4 min, isocratic 0% solvent B and the flow changes from 0.3 mL to 0.35 mL; 4–6 min, linear gradient 0%–12% B and the flow achieves 0.4 mL/min; 6–12 min, linear gradient 12%–20% B and flow constant at 0.4 mL/min; 16–17 min, linear gradient 20%–100% B and flow retrieves to 0.3 mL/min and 0%B. The flow from the chromatography was injected directly into the ESI source. The API 4000 ES source was operated in positive ion mode and the optimized parameters were: CUR: 30; IS: 5500; TEM: 500 °C; GS1: 45; GS2: 50. Elution for the analyses of polyphenols was carried out at flow of 0.3 mL/min according to the following gradient: 0–2 min, isocratic elution at 2% B; 2–10 min, linear gradient 2%–10% B; 10–17 min, linear gradient 10%–20%; 17–30 min, linear gradient 20%–30%, achieving 100%B in 1 min. The column was kept at 42 °C. The flow from the chromatography was injected directly into the ESI source. The API 4000 ES source was operated in negative ion mode and the optimized parameters were: CUR: 30; IS: −4500; T: 500 °C; GS1: 40; GS2: 45. The optimized parameters, fragmentation ions selected for each compound and dwell times are reported in Table 2. Data acquisition and processing were performed using Analyst software 1.6.2 (ABSciex, Foster City, CA, USA).

2.7. Method validation

LC-ESI-QTrap-MS/MS method was validated according to the European Medicines Agency guidelines (EMEA Quality guidelines ICH Q2) relating to the validation of analytical methods in particular precision, specificity, linearity, limit of quantification (LOQ) and limit of detection (LOD). Precision was evaluated at five concentration levels for each compound through triplicate intra-day assays and inter-day assays over 3 days. Specificity was defined as the non-interference by other analytes detected in the region of interest. Linearity was evaluated by correlation values of calibration curves. The limit of quantification (LOQ; equivalent to sensitivity), was estimated by injecting a series of increasingly dilute standard solutions until the signal-to-noise ratio was reduced to 10. The limit of detection (LOD) was estimated by injecting a series of increasingly dilute standard solutions until the signal-to-noise ratio was reduced to 3.

2.8. Principal Component Analysis

For untargeted approach, base peak chromatograms deriving from LC-ESI-Orbitrap-MS analysis (negative ion mode) were evaluated using a platform independent open source software package called MZmine (http://mzmine.sourceforge.net/). Using this toolbox with normalization by total raw signal and excluding noise from LC-MS profiles (Noise level 5.0 E3, all data points below this intensity level were ignored) 170 peaks were detected. After exporting the processed data in tabular format (.cvs file), further analysis of the data matrix were performed by SIMCA P+ software 12.0 (Umetrix AB, Umea Sweden) by Principal Component Analysis (PCA). PCA was performed by applying the peak area obtained from LC/MS analysis (Mari, Montoro, Pizza, & Piacente, 2012; Safer et al., 2011). Pareto scaling was applied before multivariate data analysis. For targeted approach a matrix was obtained starting by quantitative results, constituted by 12 variables and 15 observations. The resulting metabolomics data (not scaled) were processed using SIMCA P+ software 12.0 (Umetrix AB, Umea Sweden) by PCA in order to identify similarities among our samples. Pareto scaling was applied before multivariate data analysis.

3. Results and discussion

Fruit sampling was carried out in different areas, from plants grown in different pedoclimatic conditions and from stocks whose origin were either autochthonous and non-autochthonous, in order to investigate if composition differences were due to a different germplasm (autochthonous vs non-autochthonous) grown in similar environmental conditions and locations b) different growing conditions (spontaneous vs cultivated) among fruits grown in the same area all obtained from autochthonous germplasm and c) different geographical locations (Petina vs Sarno) among fruits belonging to the same germplasm and cultivated in similar conditions.

3.1. LC-ESI-Orbitrap-MS qualitative analysis

Results obtained from the compounds identification by LC-ESI-Orbitrap-MS and LC-ESI-Orbitrap-MS/MS are reported in Table 1, Fig. 1 and Fig. 2. Crude extracts from the different fruit samples were analysed by LC-ESI-Orbitrap- MS and LC-ESI-Orbitrap-MS/MS, both in negative and in positive ion mode. The negative LC-MS profile highlighted the presence of a large group of compounds corresponding to the deprotonated molecular ions of different flavonoids (Fig. 2A). The positive LC-MS profile highlighted the presence of anthocyanins (Fig. 2B).

High Resolution MS^n detection (positive and negative ionization modes) was used to obtain information on the structural features and the conjugated forms of phenolic compounds. Identification of phenolic compounds was based on retention times, accurate mass measurements, MS/MS data, exploration on specific metabolites public repository of mass spectral data (Mass Bank) and comparison with data reported in the literature (Del Bubba et al., 2012; Sun et al., 2014). 37 compounds, including anthocyanins, dihydroflavonols and flavonols, flavan-3-ols, proanthocyanidins, and ellagic acid and its derivatives were identified from the different *F. vesca* samples. The molecular formula of the compounds are summarised in Table 1, in which the compounds are reported according to their retention time (RT). LC-ESI-Orbitrap-MS/MS experiments were carried, selecting a sample for each typology, CPA, CSN, CPN, SPA, and SSA (see experimental section for detailed definition of the samples) in order to select and submit these ions to fragmentation experiments using the parameters previously chosen by ESI/MS direct infusion experiments. In addition to

Table 1
Identification of polyphenolic compounds and anthocyanins in *F. vesca* fruits using HPLC-ESI-Orbitrap-MS/MS.

		Molecular formula	MW	$[M-H]^-$	M^+	RT(min)	MSMS	References
Compounds identified in negative ion mode								
1	kaempferolhexoside	$C_{21}H_{20}O_{11}$	448,1005	447,0926		19.09	285	Del Bubba et al. 2012; Sun et al. 2014
2	B-type procyanidin (B1)	$C_{30}H_{26}O_{12}$	578,1424	577,1344		20.59	425,289	Del Bubba et al. 2012; Sun et al. 2014
3	B-type (epi)afzelechin-(epi)catechin	$C_{30}H_{26}O_{11}$	562,1475	561,1395		21.21	543,435,289	Del Bubba et al. 2012; Sun et al. 2014
4	Bis HHDP[a] hexose (pedunculagin)	$C_{34}H_{24}O_{22}$	784,0759	783,0679		21.39	481,301	Del Bubba et al. 2012; Sun et al. 2014
5a/5b	catechin/epicatechin	$C_{15}H_{14}O_6$	290,0790	289,0710		22.06	245,205,179	Del Bubba et al. 2012; Sun et al. 2014
6	*p*-coumaroyl-hexoside	$C_{15}H_{18}O_8$	326,1001	325,0922		22.28	265,187,162,145	Seeram et al., 2006
7	HHDP[a]galloylhexose	$C_{27}H_{22}O_{18}$	634,0816	633,0736		23.76	301,452	Del Bubba et al. 2012; Sun et al. 2014
8	B-type (epi)catechintrimer (procyanidin C1)	$C_{45}H_{38}O_{18}$	866,2058	865,1868		25.26	739,695,577,289	Del Bubba et al. 2012; Sun et al. 2014
9	B-type (epi)afzelechin-(epi)catechin	$C_{30}H_{26}O_{11}$	562,1471	561,1392		26.01	435, 407, 289	Del Bubba et al. 2012; Sun et al. 2014
10	B-type procyanidin (B2)	$C_{30}H_{26}O_{12}$	578,1424	577,1344		27.01	425,289	Del Bubba et al. 2012; Sun et al. 2014
11	dihydromyricetin-hexose	$C_{21}H_{22}O_{13}$	482,1060	481,0981		27.7	463,319,301	Abu-Reidah et al. (2014)
12	galloylhexose-HHDP-gallic acid	$C_{34}H_{26}O_{22}$	786,0915	785,0830		29.35	615,463,301,	Del Bubba et al. 2012; Sun et al. 2014
13	procyanidin B	$C_{30}H_{26}O_{12}$	578,1424	577,1344		30.05	425,289	Del Bubba et al. 2012; Sun et al. 2014
14	tris-galloyl-HHDP-hexose	$C_{41}H_{28}O_{27}$	952,0818	951,0725		31.51	907,783,484,301	Del Bubba et al. 2012; Sun et al. 2014
15	casuarictin	$C_{41}H_{28}O_{26}$	936,0868	935,0789		31.86	633,451,301	Del Bubba et al. 2012; Sun et al. 2014
16	peltatoside	$C_{26}H_{28}O_{16}$	596,1377	595,1297		33.10	463,301	Del Bubba et al. 2012; Sun et al. 2014
17	quercetin-rhamnoside	$C_{21}H_{20}O_{11}$	448,1005	447,0926		34.08	301	Del Bubba et al. 2012; Sun et al. 2014
18	ellagic acid rhamnoside	$C_{20}H_{16}O_{12}$	448,0641	447,0562		34.58	301,300,257	Del Bubba et al. 2012; Sun et al. 2014
19	castalagin/vescalagin isomer	$C_{41}H_{26}O_{26}$	934,0712	933,0632		35.03	915,631,451,301	Del Bubba et al. 2012; Sun et al. 2014
20	methyl ellagic acid glucuronide	$C_{21}H_{16}O_{14}$	492,0540	491,0460		35.47	315,300	Ferreres et al. 2013
21	isoquercitrin	$C_{21}H_{20}O_{12}$	464,0954	463,0875		35.74	301	Del Bubba et al. 2012; Sun et al. 2014
22	ellagic acid	$C_{14}H_6O_8$	302,0062	300,9983		37.12	285,257,229	Del Bubba et al. 2012; Sun et al. 2014
23	phloridzin	$C_{21}H_{24}O_{10}$	436,1369	435.1289		37.90	273	Del Bubba et al. 2012
24	taxifolin 3-alpha-L-arabinofuranoside	$C_{20}H_{20}O_{11}$	436,1005	435,0926		39.05	303,285	Del Bubba et al. 2012; Sun et al. 2014
25	quercetin glucuronide	$C_{21}H_{18}O_{13}$	478,0747	477,0667		39.98	301	Del Bubba et al. 2012; Sun et al. 2014
26	dimethyl ellagic acid pentoside	$C_{21}H_{18}O_{12}$	462,0798	461,0718		40.05	446,315,300	Del Bubba et al. 2012; Sun et al. 2014
27	kaempferolglucuronide	$C_{21}H_{18}O_{12}$	462,0798	461,0877		40.38	451,386, 285	Seeram et al. 2006
28	kaempferolcoumaroylhexoside	$C_{30}H_{26}O_{13}$	594,1373	593,1293		49.9	447,285,255	Del Bubba et al. 2012; Sun et al. 2014
Compounds identified in positive ion mode								
29	cyanidin-3-*O*-glucoside	$C_{21}H_{21}O_{11}$	449,1083		449,1083	18.86	287	Del Bubba et al. 2012
30	pelargonidin-3-*O*-glucoside	$C_{21}H_{21}O_{10}$	433,1134		433,1134	20.6	271	Del Bubba et al. 2012
31	pelargonidin-3-*O*-rutinoside	$C_{27}H_{31}O_{14}$	579,1713		579,1713	20.66	433, 271	Wu & Prior, 2005
32	peonidinglucoside	$C_{22}H_{23}O_{11}$	463,1240		463,1240	21.57	301	Del Bubba et al. 2012
33	cyanidinmalonylglucoside	$C_{24}H_{23}O_{14}$	535,1087		535,1087	25.79	287	Del Bubba et al. 2012
34	pelargonidinmalonylglucoside	$C_{24}H_{23}O_{13}$	519,1138		519,1138	27.81	433, 271	Del Bubba et al. 2012
35	peonidinmalonylglucoside	$C_{25}H_{25}O_{14}$	549,4576		549,4576	28.47	301,463	Del Bubba et al. 2012
36	delphinidin -3-*O*-glucoside	$C_{21}H_{21}O_{12}$	465,1033		465,1033	34.9	303	Wu & Prior, 2005
37	delphinidinmalonylglucoside	$C_{24}H_{23}O_{15}$	551,1036		551,1036	38.01	303	Dugo et al. 2003
Compounds identified using UHPLC system interfaced to an ABSciex (Foster city, CA, USA) API4000 QTrap in negative ion mode								
38	gallic acid	$C_7H_6O_5$	170,02	169.01		3.35	125	Del Bubba et al. 2012
39	*p*-coumaric acid	$C_9H_8O_3$	164,04	163.03		15.15	119	Del Bubba et al. 2012

[a] HHDP: Hexahydroxydiphenic acid.

Table 2
Concentrations (mg/100 g) obtained for selected phenolic compounds in *F. vesca* fruits (n = 3) using an UHPLC system interfaced to an ABSciex (Foster City, CA, USA) API4000 Q-Trap instrument in ion trap mode. **29** cyanidin-3-*O*-glucoside; **36** delphinidin-3-*O*-glucoside; **30** pelargonidin-3-*O*-glucoside; **31** pelargonidin3-*O*-rutinoside; **5a** (+)-catechin; **5b** (−)-epicatechin; **2** procyanidin B1; **10** procyanidin B2; **21** isoquercitrin; **38** gallic acid; **39** *p*-coumaric acid; **23** phloridzin.

	29	**36**	**30**	**31**[a]	**5a**[a]	**5b**	**2**	**10**	**21**	**38**	**39**	**23**[a]	Total
DP	192	92.8	76.9	98	−93	−93	−99.2	−99.2	−115	−58.9	−56.3	−89	
EP	12	11	10	13	−9	−9	−3	−3	−4	−6	−2	−4	
CE	27.2	26	28.8	40	−21.7	−21.7	−35	−35	−40	−19.7	−20	−22.2	
CXP	20.5	20.2	18.8	18.4	−14.6	−14.6	−17.7	−17.7	−17.5	−4.9	−8	−47.8	
PI	449	465	433	579	289	289	577	577	763	169	163	435	
DI	287	303	271	271	245	245	289	289	301	125	119	273	
CPA	113 ± 9	0.09 ± 0.008	8 ± 0.12	0.03 ± 0.009	0.05 ± 0.004	1.5 ± 0.050	2 ± 0.07	0.04 ± 0.003	3 ± 0.063	10 ± 0.32	0.6 ± 0.006	0.3 ± 0.023	138 ± 0.77
SPA	59 ± 3	0.06 ± 0.005	7 ± 0.13	0.03 ± 0.003	0.62 ± 0.003	3 ± 0.07	4 ± 0.11	0.04 ± 0.005	0.3 ± 0.015	6 ± 0.08	0.6 ± 0.004	0.2 ± 0.01	81 ± 0.31
SSA	25 ± 2	0.04 ± 0.004	7 ± 0.098	0.03 ± 0.001	0.007 ± 0.001	0.4 ± 0.01	0.8 ± 0.02	0.03 ± 0.001	0.08 ± 0.006	12 ± 0.155	0.7 ± 0.005	0.04 ± 0.01	46 ± 0.19
CPN	19 ± 2	0.03 ± 0.005	7 ± 0.028	0.03 ± 0.003	0.030 ± 0.003	1.2 ± 0.02	1.6 ± 0.02	0.02 ± 0.002	0.2 ± 0.006	4 ± 0.159	0.6 ± 0.003	0.2 ± 0.01	33 ± 0.19
CSN	27 ± 3	0.04 ± 0.009	9 ± 0.094	0.05 ± 0.003	0.09 ± 0.005	1.6 ± 0.04	2.5 ± 0.05	0.06 ± 0.004	0.6 ± 0.011	6 ± 0.051	0.6 ± 0.004	0.2 ± 0.017	4 ± 3

CPA Cultivated, Sarno, Authoctonous, SPA Spontaneous, Petina, Authoctonous, SSA Spontaneous, Sarno, Authoctonous, CPN Cultivated, Petina, Non Authoctonous, CSN Cultivated, Sarno, Non Authoctonous.

[a] Mean in mg/100 g fresh weight (n = 3), DP Declustering Potential EP Entrance potential, CE Collision energy, CXP Collision Cell Exit Potential, PI Product ion, DI Daughter ion.

literature, these information helped to identify compounds.

Identified compounds can be classified in five different phenolic classes.

Compounds **1**, **11**, **17**, **21**, **24**, **25**, **27**, **28** were identified as flavonoid derivatives, most of them previously identified in *F. vesca*.

Compound **11** in High Resolution Mass Spectrometry showed a pseudomolecular ion at *m/z* 481,0981, that submitted to fragmentation gave a principal daughter ion at *m/z* 319, corresponding to

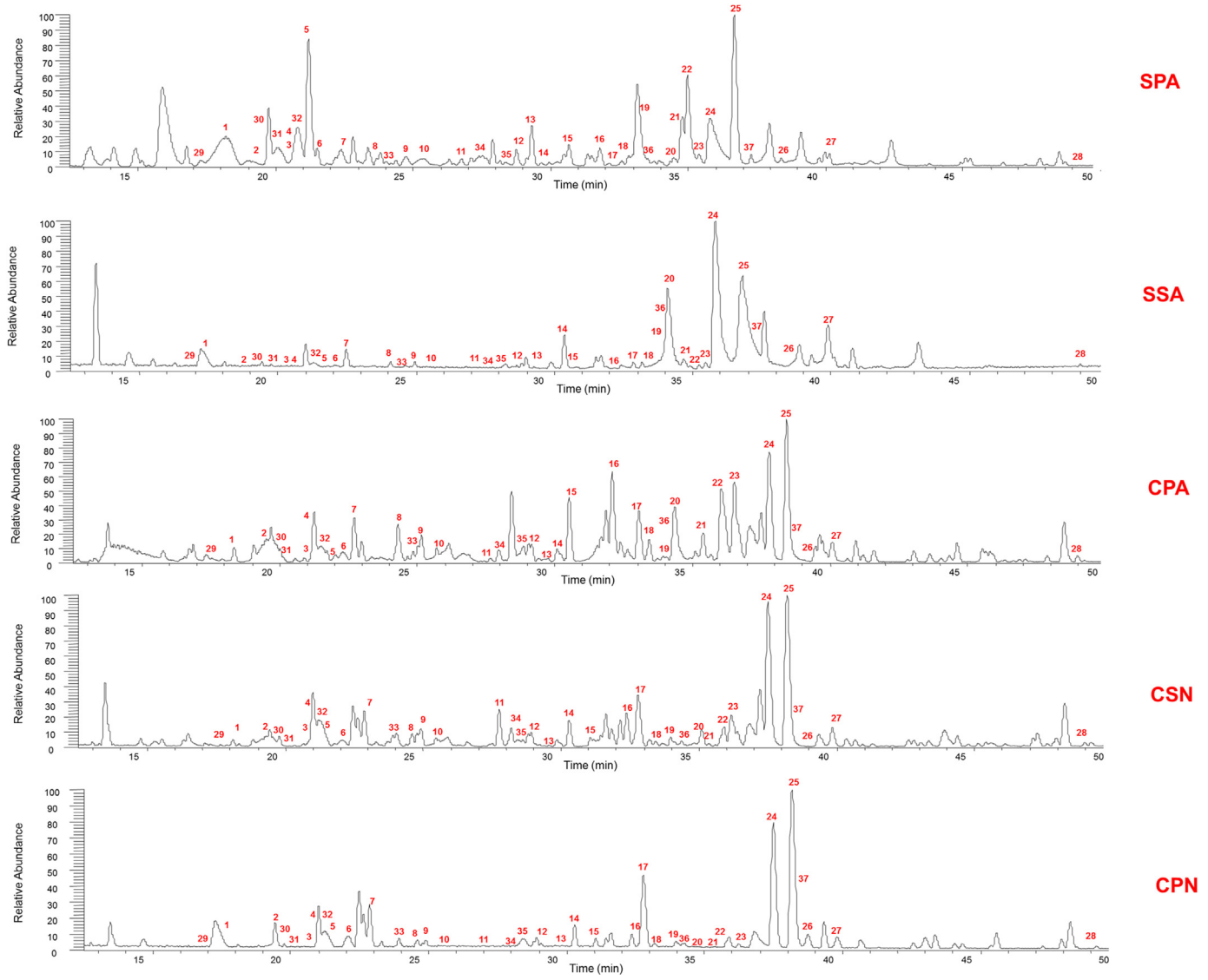

Fig. 1. LC-ESI-Orbitrap-MS Full Scan (200–1600 *m/z*) profiles in negative ion mode of *F. vesca* methanolic extracts, obtained from different samples: SPA Spontaneous, Petina, Autochthonous; SSA Spontaneous, Sarno, Autochthonous; CPA Cultivated, Petina, Autochthonous, CSN Cultivated, Sarno, Non Autochthonous; CPN Cultivated, Petina, Non Autochthonous. Peaks are numbered according to compounds reported in Table 1.

the loss of an hexose unit. A similar fragmentation was previously observed by Abu-Reidah, del Mar Contreras, Arráez-Román', Fernandez-Gutierrez & Segura-Carretero. (2014) for a metabolite detected in *Vicia faba* L (Fam. Fabaceae), and the compound was tentatively identified by High Resolution ESI-MS and High Resolution ESI-MS/MS, as dihydromyricetin hexose, never detected in the genus *Fragaria*.

Compound **27**, presented in High Resolution Mass Spectrometry a pseudomolecular peak at 461.0877, and was identified based on accurate mass and fragmentation. MS/MS spectrum showed a most abundant peak at 285 *m/z*. The same compound was identified in *Fragaria* X *ananassa* analysed in ESI-MS and ESI-MS/MS, by Seeram, Lee, Scheuller, and Heber (2006), as kaempferol-glucuronide, never detected before in *F. vesca*.

Compounds **7**, **12**, **14**, **15**, **16**, **18**, **19**, **20**, **22** and **26** were identified as ellagitannins derivatives, most of them previously identified in *F. vesca*. Compound **20**, was tentatively identified as methyl ellagic acid glucuronide, a compound never detected in this species. It showed a pseudomolecular peak at 491,0460 *m/z*, and was identified based on the exact mass and fragmentation with the use of the mass spectra data base previously reported. MS/MS spectrum showed a most abundant peak at 315 *m/z*, corresponding to the loss of the glucuronide moiety. Nevertheless, the compound **20**, for the first time tentatively identified in *F. vesca*, was described by Mass Spectrometric data by Ferreres, Grosso, Gil-Izquierdo, Valentao & Andrade, (2013), in *Cochlospermum angolensis* Welw.

Compounds **2**, **3**, **5**, **8**, **9**, **10**, **13**, were identified as catechin and proanthocyanidin derivatives, all of them previously identified or tentatively identified in *F. vesca* by Mass Spectrometric analyses (Del Bubba et al., 2012).

Compound **23**, identified as phloridzin, pertaining to the class of dihydrochalcones, was described previously in *F. vesca* (Del Bubba et al. 2012).

Compounds **29**, **30**, **31**, **32**, **33**, **34**, **35**, **36** and **37** were identified as anthocyanin derivatives, most of them previously identified in *F. vesca* (Del Bubba et al., 2012).

Compound **31** showed a pseudomolecular ion at *m/z* 579.1713 amu that submitted to fragmentation gave two principal daughter

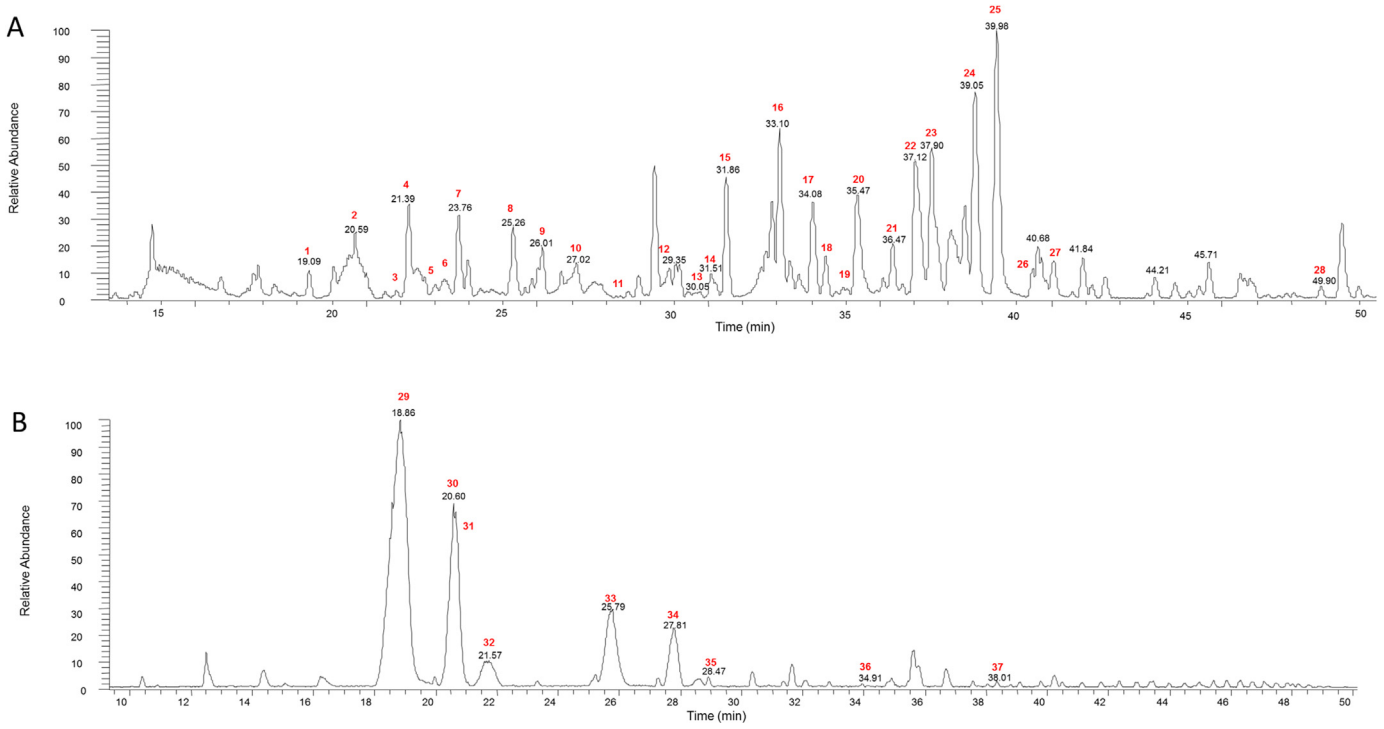

Fig. 2. LC-ESI-Orbitrap-MSFull Scan (200–1600 *m/z*) profile in negative ion mode (A) and positive ion mode (B) of *F. vesca* methanolic extracts obtained from sample CPA Cultivated, Petina, Autochthonous. Peaks are numbered according to compounds reported in Table 1.

ion at *m/z* 433, and *m/z* 271 corresponding to the loss of a pentose unite and a following loss of an hexose unit. This fragmentation pattern was described by Wu & Prior, in 2005 for pelargonidin-rutinoside. The compound was then identified as pelargonidin glucoside, never detected in *F. vesca*.

Compound **36** showed a pseudomolecular ion at *m/z* 465.1033 that in MS/MS gave a principal daughter ion at *m/z* 303, corresponding to the aglycon. This fragmentation was described by ESI-MS/MS data by Wu & Prior in 2005 for delphinidin glucoside. Thus the compound, never detected in *F. vesca*, was identified as delphinidin glucoside.

Compound **37**, showed a pseudomolecular ion at *m/z* 551.1036 that in MS/MS gave a principal daughter ion at *m/z* 303, corresponding to the aglycon. The same fragmentation was described by ESI-MS/MS data by Dugo, Mondello, Morabito & Dugo in 2003 in red orange juice, for a compounds identified as delphinidin-malonyl-glucoside, never detected in *F. vesca*.

3.2. Multivariate data analysis on qualitative data

Preliminary qualitative investigation on SPA, SSA, CPA,CSN,CPN fruits showed a different flavonoids distribution (Fig. 1).

For the analysis of the acquired dataset with multivariate methods, LC-ESI-Orbitrap-MS chromatograms were pre-processed using MZmine to compensate variations in retention time and *m/z* value between the chromatographic runs. The pre-processed chromatograms were exported as a peak list table, with rows representing the individual samples (30 samples: 15 biological samples in technical duplicates), and columns representing integrated and normalised peak areas. Moreover, these data were used through an approach of untargeted analysis, treated with an unsupervised Multi Variated Data Analysis (PCA). Samples with different geographical origin took place in a different area of the plot.

PCA was performed by applying the peak areas of each peak present in the LC-MS dataset (excluding the noise as reported in section 2.8 of Materials and Methods), and a matrix was obtained by using these areas (variables), while the columns of the matrix were the different analysed samples. The resulted score scatter plot is reported in Fig. 3 (A and B). The first component explained the 48.6% of variance while the second the 18.5%. The choice of principal components was done on the basis of the fitting (R2X) and predictive (Q2X) values for the PCA.

A good discrimination was obtained between the five classes of samples, combining positive and negative ion mode results, representative of the variability of the complex datasets obtained for the samples under investigation.

Fig. 3 reports the score scatter plot colored according to spontaneous or cultivated specimens (Panel A), displaying a clear differentiation of the samples based on this parameter, while the same scatter plot in panel B is colored according to the classification of samples based on spontaneous or cultivated combined with the geographic area where samples were collected and the germplasm.

By this visualization mode the germplasm origin appears as an important discriminator for samples, because the samples are grouped and localised in different areas of the plot.

3.3. LC-ESI-QTrap-MS/MSquantitative analysis

The results obtained from the targeted analysis by LC-ESI-QTrap-MS of *F. vesca* fruit by quantification of selected marker compounds (based on commercial available standards) and PCA are reported in Table 2, Fig. 4. In order to obtain accurate data concerning the amounts of phenolic compounds in the fruit, a quantitative LC-ESI-QTrap-MS/MS method was developed. For this purpose, preliminary ESI-QTrap-MS/MS spectra were recorded following the direct introduction of standards into the ESI source of a Mass Spectrometry instrument equipped with a triple quadruple

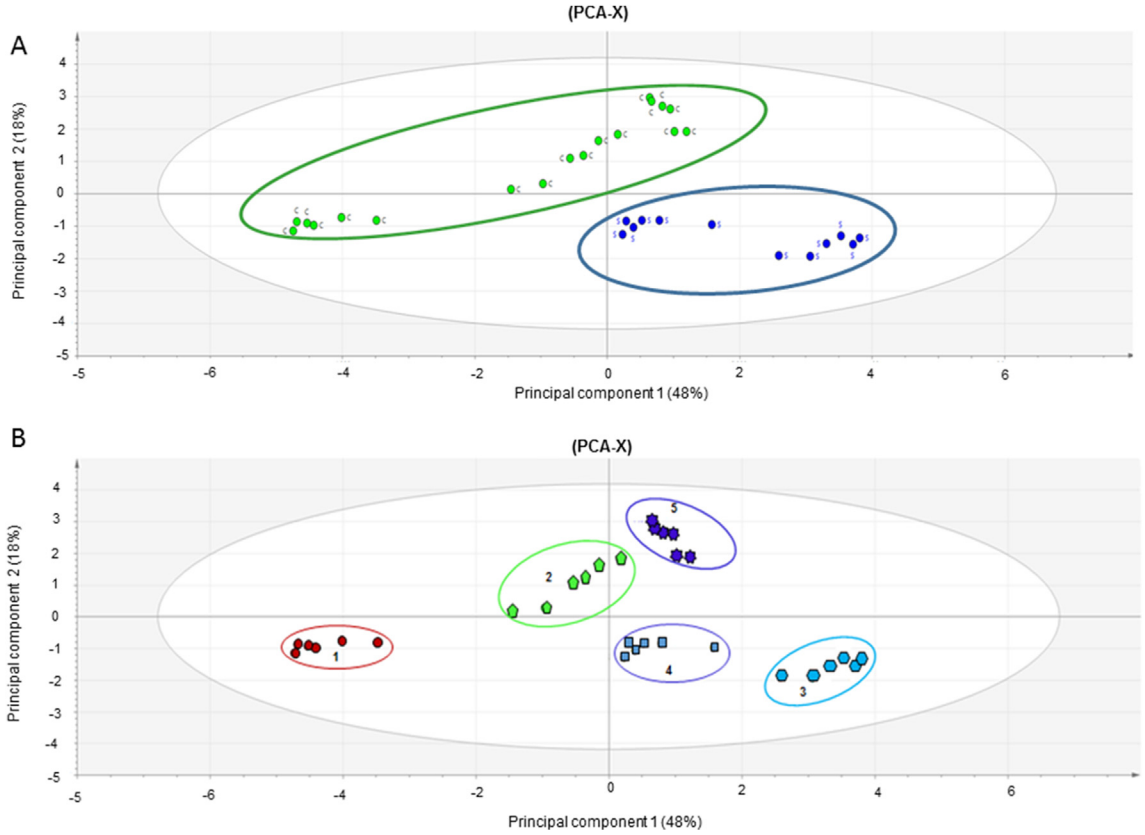

Fig. 3. Principal Component Analysis, score scatter plot, of *F. vesca* samples obtained by untargeted LC-ESI-Orbitrap-MS analysis. Panel A: colored according spontaneous (S) versus cultivated (C) species. Panel B: colored according samples classification (1) Cultivated Petina Autochthonous (CPA); (2) Cultivated Sarno Non Autochthonous (CSN); (3) Spontaneous Sarno Autochthonous (SSA); (4) Spontaneous Petina Autochthonous (SPA); (5) Cultivated Petina Non Autochthonous (CPN).

analyser. The transitions observed during ESI/MS/MS experiments, were used to develop a selective and sensitive LC-ESI-QTrap-MS/MS method by using the technique of Multiple Reaction Monitoring (MRM). The MRM method therefore involved the precursor/product transitions described in Table 2, that shows in addition mass spectral parameters for the standard compounds.

Table 2 reports also the results for quantitative analysis of phenolic compounds in the *F. vesca* samples used in this work. Each of the 15 samples was analysed in triplicate, and the results reported in table are the average values of the three analyses.

The content of phenolic compounds shows the same order of magnitude in the different samples, with the higher total phenolic values showed by cultivated fruits, from the area of Alburni from local seeds. In addition for any compound there are significant differences. Cyanidin-3-*O*-glucoside for example, appears to be present in high quantity in the samples collected in the Alburni Area.

3.4. Method validation

LC-ESI-QTrap-MS/MS method was validated according to the EMEA guidelines. Calibration curves were obtained by plotting the area of External Standard against the known concentration of each compound; each concentration of standard solutions was analysed in triplicate. A good linearity (correlation coefficients (r) from 0.994 to 0.999) was achieved for all analytes in the concentration range.

The limit of detection (LOD) and the limit of quantification (LOQ) for each target standard compound were determined, under the optimised conditions, by the serial dilution of a standard solution until the signal-to-noise ratios (S/N) were 3:1 and 10:1, respectively. The LOD for each analyte varied from 0.001 to 0.012 g/L and LOQ from 0.004 to 0.09 g/L, indicating that the developed method exhibited good sensitivity. The results for each compound are given in Table 3.

Three aliquots of each same sample were analysed within the same day, and another three aliquots of the same sample were analysed during three consecutive days, one for each day. Percentage relative standard deviation (RSD) was used to express precision of the method (Table 3).

Recovery experiments were performed with the optimized parameters to evaluate the extraction efficiency and the developed analytical method. Standard solutions at three different concentration levels (high, middle and low) were added in a known amount of sample and analysed by LC-ESI-QTrap-MS/MS and then triplicate experiments were performed at each level. Within the same day, the recovery (%) ranged from 94.6% to 104.7%, thus demonstrating good recovery and precision.

3.5. Multivariate data analysis on quantitative data

In order to better understand as variation in quantity of each metabolite could classify the samples and in order to validate the PCA findings described above, a targeted multivariate approach was applied to quantitative data.

The Principal Component Analysis was applied to the matrix obtained, with the quantitative content of each marker compound in each type of samples (15 samples, average values of triplicates). Fig. 4A shows the 2D projection plot (score scatter plot) of the *F. vesca* samples. The first component explains the 46.65% of variance while the second the 15.38%. As for the previous PCA analysis,

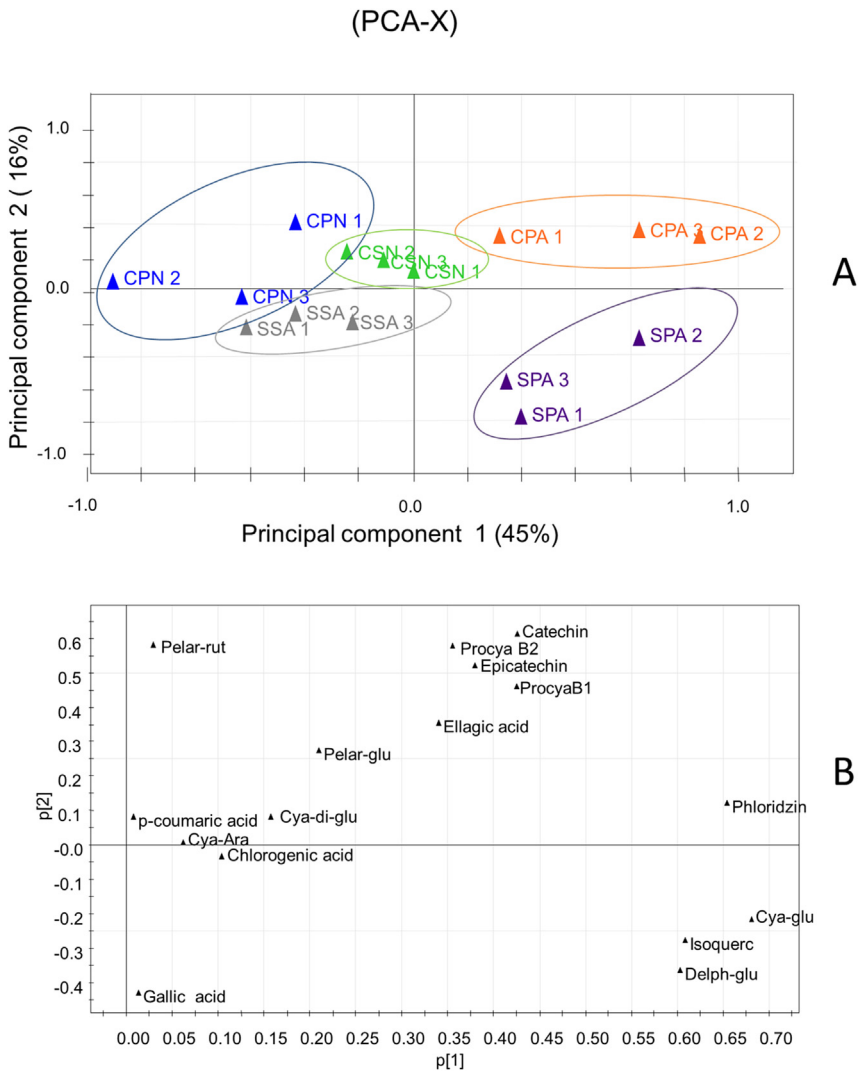

Fig. 4. Principal Component Analysis of *F. vesca* samples obtained by targeted quantitative analysis data. Panel A: score scatter plot; Panel B: loading scatter plot. Cultivated Petina Autochthonous (CPA); Cultivated Sarno Non Autochthonous (CSN); Spontaneous Sarno Autochthonous (SSA); Spontaneous Petina Autochthonous (SPA); Cultivated Petina Non Autochthonous (CPN).

Table 3
Validation data for the quantitative UPLC-MS/MS method developed by using an UHPLC system interfaced to an ABSciex (Foster City, CA, USA) API4000 Q-Trap instrument in ion trap mode to analyse possible marker compounds in *F. vesca* samples.

		r^2	Regression equation	LOD	LOQ	Precision RSD (%)	
						Intraday	Interday
				(mg/ml)	(mg/ml)	(n = 3)	(n = 3)
29	cyanidin-3-*O*-glucoside	0.994	y = 6.14e4x−2.03e4	0.003	0.010	3.42	6.72
36	delphinidin-3-*O*-glucoside	0.9982	y = 2.01e6x+5.46e4	0.012	0.050	2.96	6.76
30	pelargonidin-3-*O*-glucoside	0.9962	y = 3.96e6x+1.26e6	0.001	0.004	4.74	5.82
31	pelargonidin-3-*O*-rutinoside	0.9991	y = 1.92e6x−722	0.004	0.009	4.31	5.56
5a	(+) catechin	0.9999	y = 6.9e5x+7.63e3	0.004	0.008	2.32	3.56
5b	(−) epicatechin	0.9999	y = 6.41e5x+9.44e4	0.003	0.008	1.72	2.06
2	procyanidin B1	0.9991	y = 4.41e5x−1.23e4	0.002	0.005	0.47	1.26
10	procyanidin B2	0.9993	y = 4.31e5x+3.17e3	0.003	0.007	5.90	6.50
21	isoquercetin	0.9986	y = 2.33e6x +3.3e4	0.001	0.003	2.03	3.15
38	gallic acid	0.9982	y = 6.18e4x+1.64e4	0.030	0.090	3.43	3.56
39	*p*-coumaric acid	0.9968	y = 1.34e6x−3.66e5	0.005	0.009	1.01	1.97
23	phloridzin	0.9993	y = 1.97e5x+951	0.003	0.006	4.02	5.23

also in this case the choice of principal components was done on the basis of the fitting (R2X) and predictive (Q2X) values for the PCA model.

The Score Scatter Plot gives an evidence that there are confined cluster areas in the 2D diagram, linked to geographical origin.

The Alburni germplasm has more importance in classification with respect to the cultivation area. The spontaneous samples from Alburni, and the cultivated in Alburni samples by original seeds from the same area, are restricted in the right part of the plot. The lower square is occupied mainly by spontaneous samples and the higher square by cultivated samples, but generally at the right area respect to the y axis are collocated all the samples with a typical and certified origin from the Alburni Area.

In order to evaluate the influence of each variable on the classification of the samples the loading plot obtained for the same dataset was then studied, and it is shown in Fig. 4B: Interestingly, the loading plot shows the *m/z* values corresponding, for each region of the 2D space, to the peak observed in the specific samples. In particular variables relevant to the differentiation of the samples and to their placement in a specific area of the plot, can be highlighted.

A host of polyphenols seems to have a primary role in the discrimination of samples and we can see them in the lower right square of the loading plot of the PCA model obtained (Fig. 4B). Derivatives of cyanidin (i.e. cyanidin-3-*O*-glucoside) appear to be over expressed metabolites in samples coming from the Alburni area. Other metabolites that characterize these germplasm, seems to be another anthocyan compound, delphinidin-3-*O*-glucoside, and a flavonol, isoquercetin. This can be informative and useful as it allows to rapidly identify the geographical origin of the plants to attribute the typical label to the fruit.

4. Conclusions

The aim of this work was to investigate the feasibility to discriminate wild strawberries from different geographic origins, grown in different conditions and from different germplasm on the basis of their phytochemicals by applying untargeted and targeted liquid chromatography-mass spectrometry techniques combined with PCA.39 different phenolic compounds have been detected in the methanolic extract of spontaneous and cultivated strawberries from different origins. Spontaneous strawberries from the Alburni Area produced a higher amount of anthocyanins when compared to the ones from other geographical origin. Our results highlight the potential efficacy of untargeted metabolite profiling in the assessment of the geographical origin of the berries. This is a very important issue because fruit value is largely based on the origin authenticity, especially for products with local peculiarity. MS based protocols (targeted and untargeted) combined with multivariate statistical analysis is then an important tool for food quality evaluation based on authenticity of the products, and principally in the way to identify specific biomarkers. Based on the statement that MS instruments are expensive and LC-MS are not routine techniques, such approaches can be a useful way to identify specific origin and authenticity markers, that, when defined, can be analysed with cheaper and easier techniques in common routine laboratories for food quality control.

References

Abu-Reidah, I. M., del Mar Contreras, M., Arráez-Román', D., Fernández-Gutiérrez, A., & Segura-Carretero, A. (2014). UHPLC-ESI-QTOF-MS-based metabolic profiling of *Vicia faba* L. (Fabaceae) seeds as a key strategy for characterization in foodomics. *Electrophoresis, 35*, 1571–1581.

Brown, P. N., Murch, S. J., & Shipley, P. (2012). .Phytochemical diversity of cranberry (*Vaccinium macrocarpon* Aiton) cultivars by anthocyanin determination and metabolomic profiling with chemometricanalysis. *Journal of Agricultural and Food Chemistry, 60*, 261–271.

Carvalho, E., Franceschi, P., Feller, A., Palmieri, L., Wehrens, R., & Martens, S. (2013). A targeted metabolomics approach to understand differences in flavonoid biosynthesis in red and yellow raspberries. *Plant Physiology and Biochemistry, 72*, 79–86.

Cheel, J., Theoduloz, C., Rodriguez, J. A., Caligari, P. D. S., & Schmeda-Hirschmann, G. (2007). Free radical scavenging activity and phenolic content in achenes and thalamus from Fragariachiloensis spp. chiloensis, *F. vesca* and *F. x ananassa* cv. Chandler. *Food Chemistry, 102*, 36–44.

Del Bubba, M., Checchini, L., Chiuminatto, U., Doumett, S., Fibbia, D., & Giordanic, E. (2012). Liquid chromatographic/electrospray ionization tandem mass spectrometric study of polyphenolic composition of four cultivars of *Fragaria vesca* L. berries and their comparative evaluation. *Journal of Mass Spectrometry, 47*, 1207–1220.

Dugo, P., Mondello, L., Morabito, L., & Dugo, G. (2013). Characterization of the anthocyanin fraction of Sicilian blood orange juice by Micro-HPLC-ESI/MS. *Journal of Agricultural and Food Chemisty, 51*, 1173–1176.

EMEA Quality guidelines: Validation of analytical procedures: Text and methodology (ICH Q2). Available at: http://www.emea.europa.eu/pdfs/human/ich/038195en.pdf.

Ferreres, F., Grosso, C., Gil-Izquierdo, A., Valentao, P., & Andrade, P. B. (2013). Ellagic acid and derivatives from *Cochlospermum angolensis* Welw. Extracts: HPLC-DAD-ESI/MS^n profiling, quantification and in vitro antidepressant, anticholinesterase and antioxidant activities. *Phytochemical Analysis, 24*, 534–540.

Herrero, M., Simó, C., García-Cañas, V., Ibáñez, E., & Cifuentes, A. (2012). Foodomics: MS-based strategies in modern food science and nutrition. *Mass Spectrometry Reviews, 31*, 49–69.

Mari, A., Montoro, P., Pizza, C., & Piacente, S. (2012). Liquid chromatography tandem mass spectrometry determination of chemical markers and principal component analysis of *Vitexagnus-castus* L. fruits (Verbenaceae) and derived food supplements. *Journal of Pharmaceutical and Biomedical Analysis, 70*, 224–230.

McDougall, G., Martinussen, I., & Stewart, D. (2008). Towards fruitful metabolomics: High throughput analyses of polyphenol composition in berries using direct infusion mass spectrometry. *Journal of Chromatography B, 871*, 362–369.

Pavlovic, A. V., Dabic, D. C., Momirovic, N. M., Dojcinovic, B. P., Milojkovic-Opsenica, D. M., Tesic, Z. Lj., et al. (2013). Chemical composition of two different extracts of berries harvested in Serbia. *Journal of Agricultural and Food Chemistry, 61*(17), 4188–4194.

Pellegrini, N., Serafini, M., Colombi, B., Del Rio, D., Salvatore, S., & Bianchi, M. (2003). Total antioxidant capacity of plant foods, beverages and oils consumed in Italy assessed by three different in vitro assays. *Journal of Nutrition, 133*, 2812–2819.

Ramakrishna, A., & Ravishankar, G. A. (2011). Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signaling&Behavior, 6*(11), 1720–1731.

Safer, S., Serhat, Cicek, S., Pieri, V., Schwaiger, S., Schneider, P., … Stuppner, H. (2011). Metabolic fingerprinting of *Leontopodium* species (Asteraceae) by means of 1H NMR and HPLC–ESI-MS. *Phytochemistry, 72*, 1379–1389.

Seeram, N. P., Lee, R., Scheuller, H. S., & Heber, D. (2006). .Identification of phenolic compounds in strawberries by liquid chromatography electrospray ionization mass spectroscopy. *Food Chemistry, 97*, 1–11.

Sun, J., Liu, X., Yang, T., Slovin, J., & Chen, P. (2014). Profiling polyphenols of two diploid strawberry (*Fragaria vesca*) inbred lines using UHPLC-$HRMS^n$. *Food Chemistry, 146*, 289–298.

Wu, X., & Prior, R. L. (2005). Identification and characterization of anthocyanins by high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry in common foods in the United States: Vegetables, nuts, and grains. *Journal of Agricultural and Food Chemistry, 53*, 3101–3113.

Yildirim, A. B., & Turker, A. U. (2014). Effects of regeneration enhancers on micropropagation of *Fragaria vesca* L. and phenolic content comparison of field-grown and in vitro-grown plant materials by liquid chromatography-electrospray tandem mass spectrometry (LC–ESI-MS/MS). *Scientia Horticulturae, 169*, 169–178.