Journal of Berry Research 7 (2017) 217–229
DOI:10.3233/JBR-170158
IOS Press

LC-MS based metabolomics study of different parts of myrtle berry from Sardinia (Italy)

Gilda D'Urso[a], Giorgia Sarais[b], Carla Lai[b], Cosimo Pizza[a] and Paola Montoro[a,*]
[a]*Department of Pharmacy, University of Salerno, Fisciano (SA), Italy*
[b]*Department of Life and Environmental Sciences, University of Cagliari, Cagliari, Italy*

Received 19 December 2016; accepted 9 March 2017

Abstract.
BACKGROUND: Myrtle berries have had a long history of application in the perfumery, cosmetic, food and pharmaceutical industries as well as being used for the industrial formulation of typical sweet liqueurs. However, no data is currently available on the metabolite composition and distribution in the different parts of the berry.
OBJECTIVE: In the present study a metabolomics approach followed by multivariate data analysis and phytochemical characterization of (poly) phenolic metabolites, using liquid chromatography coupled with high resolution mass spectrometry, was developed to identify novel markers in different parts of myrtle berries and to understand which part of the fruit has the most influence on the metabolomics classification of berries, based on geographic origin of the plant and the cultivars.
RESULTS: By using LC-ESI-Orbitrap-MS analysis, 35 compounds were tentatively identified on the base of their retention time, UV/Vis spectra, MS spectra and MS fragmentation patterns. 19 compounds, pertaining mainly to polyphenol compounds like flavonoids and to a new class of hydrolysable tannins, were detected and identified for the first time in these berries (mainly in seeds).
CONCLUSIONS: By using multivariate statistical analysis, predictive classification models for authenticity and geographical origin, assessment was obtained. With this study, flavonoids and anthocyanins, mainly found in the peel and pulp of the myrtle berry, were recognized as putative marker compounds to assess the geographic origin of these berries.

Keywords: *Myrtus communis*, berry, LC-ESI-Orbitrap-MS, metabolomics, phenolic compounds, multivariate data analysis

1. Introduction

Polyphenols like anthocyanins, ellagitannins, flavonoids conjugates and hydroxycinnamic acid are the most abundant source of health promoting phytochemical compounds in berry fruits [1, 2]. In particular, Ellagitannins, pertain to a group of compounds known as hydrolysable tannins and are polymers and a polyol, which is usually either glucose or quinic acid [3]. They are generally uncommon in fruit and vegetables and are found only in a few berry fruits such as strawberries (*Fragaria ananassa* D.) raspberries (*Rubus idaeus* L.) and blackberries (*Rubus* spp) [4]. For this class of compounds, recent studies have reported several biological properties which make them suitable not only for use in textiles, but also for other applications in cosmetics, medicine, agronomy and phytotherapy [5, 6].

*Corresponding author: Paola Montoro, Department of Pharmacy, University of Salerno, Via Giovanni Paolo II, 84084 Fisciano (SA), Italy. Tel.: +39 089969252; Fax: +39 089969602; E-mail: pmontoro@unisa.it.

These compounds are not equally distributed in both quality and quantity across the different organs of a plant and are subject to quantitative, seasonal and inter-species variations [7].

Some compounds, pertaining to this class, are mostly reported in the leaves of *Myrtus communis* [8, 9].

Myrtus communis belongs to the Myrtaceae family. It is a pleasant annual shrub with dark blue ripe berries, which have been used for some time in the perfumery, cosmetic, food and pharmaceutical industries [10]. In Sardinia (Italy), these berries are mostly used for the industrial formulation of sweet liqueurs, typical of the Sardinia region (Italy) [11]. The species *Myrtus communis* has been the subject of several researches focused on the chemical composition of the essential oil [12, 13] and of the methanol extracts of myrtle berries. The species has been investigated for the presence of flavonoids and anthocyanins [14–17]. Nevertheless, no data are present on the metabolite composition and distribution in the different parts of the berry.

In the present work liquid chromatography, coupled with high resolution mass spectrometry in conjunction with principal components analysis (PCA), was applied to investigate the metabolic composition of different parts of the *Myrtus communis* purple berry. The aim is to understand which part of the fruit has the most influence on the metabolomics classification of berries, based on the geographic origin of the plant.

Metabolite profiling, using liquid chromatography combined with high resolution mass spectrometry (LC-ESI-Orbitrap-MS) has proved to be a powerful tool for discovering changes in metabolite composition in different fruit parts of *Myrtus communis*. Through LC-ESI-Orbitrap-MS analysis and MS/MS experiments, 35 compounds were identified or tentatively identified on the bases of their retention time, UV/Vis absorbance, MS spectra and MS fragmentation patterns and a new class of hydrolysable tannins was identified for the first time in these berries.

2. Materials and methods

2.1. Plant materials

Seeds of two cultivars of *Myrtus communis*, were collected from the geographic area of Sassari and Cagliari (Sardinia, Italia) respectively, and then grown in the same Experimental Station of the University of Sassari located in Oristano (Sardinia, Italy). Three samples of purple berries from plants derived from seeds of the Sassari area (1, 2, 3) and three samples from plants derived from seeds of the Cagliari area (4, 5, 6) were randomly harvested ($\approx$2 kg) during the fruit ripening stage in December 2014. Before analysis, the samples collected were immediately frozen in liquid nitrogen and then freeze dried. The peel and pulp were removed and separated from the seeds. Whole berry, peel, pulp and seeds were used for the analysis.

In total 18 samples were obtained and labeled as follows: 1S, 2S, 3S (seeds from samples developed from original seeds of the Sassari area); 4S, 5S, 6S (seeds from samples developed from original seeds of the Cagliari area); 1P&P, 2P&P, 3P&P (peel and pulp from samples developed from the original seeds of the Sassari area); 4P&P, 5P&P, 6P&P (peel and pulp from samples developed from original seeds of the Cagliari area); 1B, 2B, 3B (whole berries from plants developed by seeds of the Sassari area); 4B, 5B, 6B (whole berries from plants developed by seeds of the Cagliari area). Each sample was analyzed in duplicate (obtaining 36 samples in total).

Different cultivars were not registered (with a specific voucher number), but they were described in several scientific publications where specific names were attributed to them (as required by Italian law), Maria Rita, Barbara, Daniela, Giovanna, Lelia, Tonina [18].

2.2. Sample preparation

Seeds, peel and pulp and whole berries were lightly pulverized following the procedure described by De Vos et al. (2007) [19] with slight modifications: 0.5 g was extracted with 10 ml of a solution comprised of 70% aqueous methanol acidified at 0.1% of formic acid followed by sonication for 15 minutes and centrifugation for

another 15 minutes at 1750 rpm. The supernatant was collected and filtered through 0.45μm filters, and then 100μL of extract was diluted in 900μL of water (of LC-MS grade), for a final concentration of approximately 1 mg/mL.

For the LC-MS analysis 10 μL were used. Each sample was analysed in technical duplicate.

2.3. Reagents and solvents

Formic acid and methanol for extraction were purchased from VWR international PBI S.r.l. (Milano, Italy). Acetonitrile, water and formic acid (all of LC-MS grade) were purchased from Merck (Darmastadt, Germany).

2.4. HPLC-UV/VIS analysis

An Agilent (Palo Alto, CA, USA) 1260 Infinity system consisting of a G1312C binary pump, a G-1328A Rheodyne injector (20 μL injection loop) a G-1379A degasser and a G1314B photodiode array detector were employed to develop the chromatographic method. Analyses were performed using a Waters XSelect CSH C18 (2.1 mm × 150 mm particle size 3.5 μm) column, eluted with water containing 0.1% formic acid (solvent A) and acetonitrile containing 0.1% formic acid (solvent B). A linear gradient program at a flow rate of 0.200 mL/min was used: 0–15 min, from 10 to 20% (B); 15 to 25 min, from 20 to 40% (B); 25 to 35 min from 40 to 60% (B) then to 100% (B) for 5 min and back to 10% (B) for other 5 min. This gradient was used for LC-ESI-Orbitrap-MS analysis described below. Detection was carried out with two wavelengths, 360 nm specific for flavonoids and 520 nm specific for anthocyanins.

2.5. LC-ESI –orbitrap-MS analysis

LC-ESI-Orbitrap-MS analyses were performed essentially as described by D'Urso et al. (2015) [20].

An HPLC method coupled with a hybrid mass spectrometer, combining a linear trap quadruple (LTQ) and an Orbitrap mass analyser, was developed for the study of the main metabolites characteristic to each part of the berry. Experiments were performed working with a Thermo Scientific liquid chromatography system based on a quaternary Accela 600 pump and an Accela auto sampler, hyphenated with a linear orbitrap hybrid mass spectrometer (LTQ-Orbitrap XL, Thermo Fisher Scientific, Bremen Germany) equipped with electrospray ionization (ESI). Separation was performed on a XSelect CSH C18 (Waters, Milford, MA) column (2.1 mm × 150 mm particle size 3.5 μm). The mobile phase selected after optimization consisted of solvent A (water acidified with 0.1% formic acid) and solvent B (acetonitrile acidified with 0.1% formic acid). The gradient program used is described above in HPLC-UV/Vis analysis. The mass spectrometer was operated in negative ion mode. ESI source parameters were as follows: capillary voltage –12 V; tube lens voltage –121.47; capillary temperature 280°C; Sheath and Auxiliary Gas flow (N_2) 30 and 5, Sweep gas 0 Spray voltage 5. MS spectra were acquired by full range acquisition covering *m/z* 200–1600. A data dependent scan was performed with the aim of obtaining MS/MS experiments, selecting precursor ions corresponding to most intensive peaks in LC-MS analysis. The experiments were also performed in positive ion mode for a quality screening followed by a fragmentation study. ESI source parameters were as follows: capillary voltage 49 V; tube lens voltage 120; capillary temperature 280°C; sheath and Auxiliary Gas flow (N_2) 30 and 5, Sweep gas 0 Spray voltage 5. MS spectra was acquired in the full range acquisition *m/z* 250–1600.

Spectral characteristics fragmentations allowed for the identification of phenolic compounds in addition to specific retention times, and finally, data were compared with literature. Xcalibur software version 2.1 was used for instrument control, data acquisition and data analysis.

2.6. Multivariate data analysis

LC-MS raw data of the 36 samples deriving from LC-ESI-Orbitrap-MS analysis (negative ion mode) were analysed using a platform independent open source software package called MZmine (http://mzmine.sourceforge.net/). Using this toolbox with normalization of total raw signal, we detected 634 peaks. 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 [21, 22]. Pareto scaling was applied before multivariate data analysis.

3. Results and discussion

A preliminary LC-UV analysis, at 360 nm, was focused on the detection of phenolic compounds, mainly flavonoids; an LC-UV analysis, wavelength at 520 nm, was focused on the detection of anthocyanin. Results (data not showed) indicated that both classes of compounds were present in the different berry parts. Therefore, LC-MS experiments were performed in positive and negative ion mode.

3.1. LC-MS metabolomics analysis

Aqueous/methanol extracts of seeds, peel and pulp and whole myrtle berry were analyzed by LC-ESI-Orbitrap-MS.

Base-peak chromatograms in negative ion mode of the different parts of myrtle berries extracts are presented in Fig. 1, while base-peak chromatograms in positive ion mode are presented in Fig. 2. In total 35 compounds were identified or putatively identified based on retention time, accurate mass measurement, fragmentation pattern and comparison with data reported in literature (Table 1). As far as we know 19 of these compounds were identified for the first time in myrtle berries. In the current work, the identified compounds are classified into six groups: hydrolysable tannins; hydroxycinnamic acid; gallomyrtucommulones; flavanols; flavonols; anthocyanins. All of them are present in all parts of the fruit, with variability in intensity.

3.1.1. Hydrolysable tannins

This class of compounds was detected at 360 nm and identified through LC-MS and MS/MS experiments by operating in negative ion mode.

Compound **1**, **2**, **4**, **5**, **13**, **14**, **15**, **16**, **25**, **26**, **21**, **31** were identified as hydrolysable tannins; these compounds were present in all parts of the fruits, but their content was higher in the seeds than in the peel and pulp (see Fig. 1 and 2).

Compounds **1**, **2**, **4**, **5** and **14** were previously identified in leaves of *Myrtus communis* [8, 9].

Compound **1** showed a pseudo molecular ion at *m/z* 481.0617 corresponding to the molecular formula $C_{20}H_{17}O_{14}$ producing in MS/MS a fragment at *m/z* 301 which indicates the release of ellagic acid, thus by comparison of MS data reported in literature the compound was tentatively identified as hexahydroxydiphenoyl (HHDP) hexoside [8].

Compound **2** showed a pseudo molecular ion at *m/z* 331.0664 corresponding to the molecular formula $C_{13}H_{15}O_{10}$, that gives fragment ion at *m/z* 271 $[M\text{-}H\text{-}60]^-$ and 169, that corresponds to the aglycon form due to the loss of a hexosyl moiety (162 Da). This fragmentation pattern agrees with data reported in literature and database [9] and thus the compound was identified as monogalloylhexose; Compound **4** showed a pseudo molecular ion at *m/z* 633.0726 corresponding to the formula $C_{34}H_{21}O_{22}$ and when submitted to fragmentation gave a main product ion at *m/z* 301, which corresponds to the loss of a galloyl hexose unit (332 Da) from the precursor ion. Thus, this compound was characterized as tannin strictinin (galloyl-HHDP hexose), in agree-

Table 1
Identified compounds in the different parts of *Myrtus communis* berries using HPLC-ESI-Orbitrap-MS/MS analysis in negative and positive ion mode

Identity	Molecular Formula	MW	$[M-H]^-$	MS/MS	RT	Reference
Hydrolyzable Tannins						
1 HHDP-hexose	$C_{20}H_{18}O_{14}$	482.0696	481.0617	301/275	2.4	8
2 monogalloylhexose	$C_{13}H_{16}O_{10}$	332.0743	331.0664	313/271/169	2.6	9
4 strictinin (galloyl-HHDP hexose)	$C_{34}H_{22}O_{22}$	634.0806	633.0726	301/421/615	3.6	9*
5 galloylquinic acid	$C_{14}H_{16}O_{10}$	344.0743	343.0664	191/169	3.62	9*
13 tellimagrandin I	$C_{34}H_{26}O_{22}$	786.0915	785.0836	301/483/633	6.4	23
14 punicalin	$C_{34}H_{22}O_{22}$	782.0602	781.0523	N.F.	8.1	9
15 pedunculagin (Bis HHDP hexose)	$C_{34}H_{24}O_{22}$	784.0759	783.0679	481/301	8.4	23
16 casuarictin	$C_{41}H_{28}O_{26}$	936.0868	935.0789	783/633/301	8.7	25
25 castalagin	$C_{41}H_{26}O_{26}$	934.0712	933.0632	N.F.	15.6	26
26 tellimagrandin II (Trigalloyl HHDP hexose)	$C_{41}H_{30}O_{26}$	938.1025	937.0945	N.F.	17.2	23
21 ellagic acid hexoside	$C_{20}H_{16}O_{13}$	464.0591	463.0511	301	12.3	9*
31 ellagic acid	$C_{14}H_6O_8$	302.0063	300.9983	284/256/185	20.2	26
Gallomyrtucommulones						
35 gallomyrtucommulone C	$C_{27}H_{36}O_{13}$	568.2155	567.2075	331/313/271/169	24.6	9*
Hydroxycinnamic acids						
18 caffeoylhexose	$C_{15}H_{18}O_9$	342.0950	341.0871	161/179	9.87	27
Flavanols						
7 epigallocatechin	$C_{15}H_{14}O_7$	306.0739	305.0660	174/270	4.37	26
22 catechin/epicatechin	$C_{15}H_{14}O_6$	290.0790	289.0711	245/179	13.3	26
Flavonols						
23 myricetin galloylhexoside	$C_{28}H_{24}O_{17}$	632.1013	631.0934	479/317	13.9	9*
24 myricetin hexoside	$C_{21}H_{20}O_{13}$	480.0903	479.0824	317	15.4	30 and 31
27 myricetin pentoside	$C_{20}H_{18}O_{12}$	450.0798	449.0718	317	17.9	30 and 31
28 quercetin galloylhexoside	$C_{28}H_{24}O_{16}$	616.1064	615.0985	463/301	18.2	28*
29 quercetin hexoside	$C_{21}H_{20}O_{12}$	464.0954	463.0875	301	18.6	30and 31
30 myricetin deoxyhexoside	$C_{21}H_{20}O_{12}$	464.0954	463.0875	317	18.7	30 and 31
32 quercetin deoxyhexoside	$C_{21}H_{20}O_{11}$	448.1005	447.0926	301	21.6	30 and 31
33 myricetin galloyl deoxyhexose	$C_{28}H_{24}O_{16}$	616.1064	615.0984	317/463	23.2	9*
34 myricetin hexose deoxyhexose	$C_{27}H_{30}O_{17}$	626.1483	625.1403	479/317	23.3	29*
Anthocyanins			$[\mathbf{M}]^+$			
3 delphinidin hexoside	$C_{21}H_{21}O_{12}^+$	465.1033	465.1034	303	3.0	30 and 31
6 cyanidin hexoside	$C_{21}H_{21}O_{11}^+$	449.1083	449.1083	287	4.2	30 and 31
8 petunidin hexoside	$C_{22}H_{23}O_{12}^+$	479.1189	479.1189	317	4.9	30 and 31
9 delphinidin pentoside	$C_{20}H_{19}O_{11}^+$	435.0925	435.0925	303	5.9	30 and 31
10 peonidin hexoside	$C_{22}H_{23}O_{11}^+$	463.1240	463.1240	301	6.8	30 and 31
11 malvidin hexoside	$C_{23}H_{25}O_{12}^+$	493.1346	493.1346	331	6.9	30 and 31
12 cyanidin pentoside	$C_{20}H_{19}O_{10}^+$	419.0978	419.0978	287	7.6	32
17 petunidin pentoside	$C_{21}H_{21}O_{11}^+$	449.1083	449.1083	317	8.8	30 and 31
19 peonidin pentoside	$C_{21}H_{21}O_{10}^+$	433.1134	433.1134	301	10.45	32
20 malvidin pentoside	$C_{22}H_{23}O_{11}^+$	463.1240	463.1240	331	10.74	30 and 31

*Found in leaves of *Myrtus communis*. MW: Molecular weight; N.F. Not Fragmented.

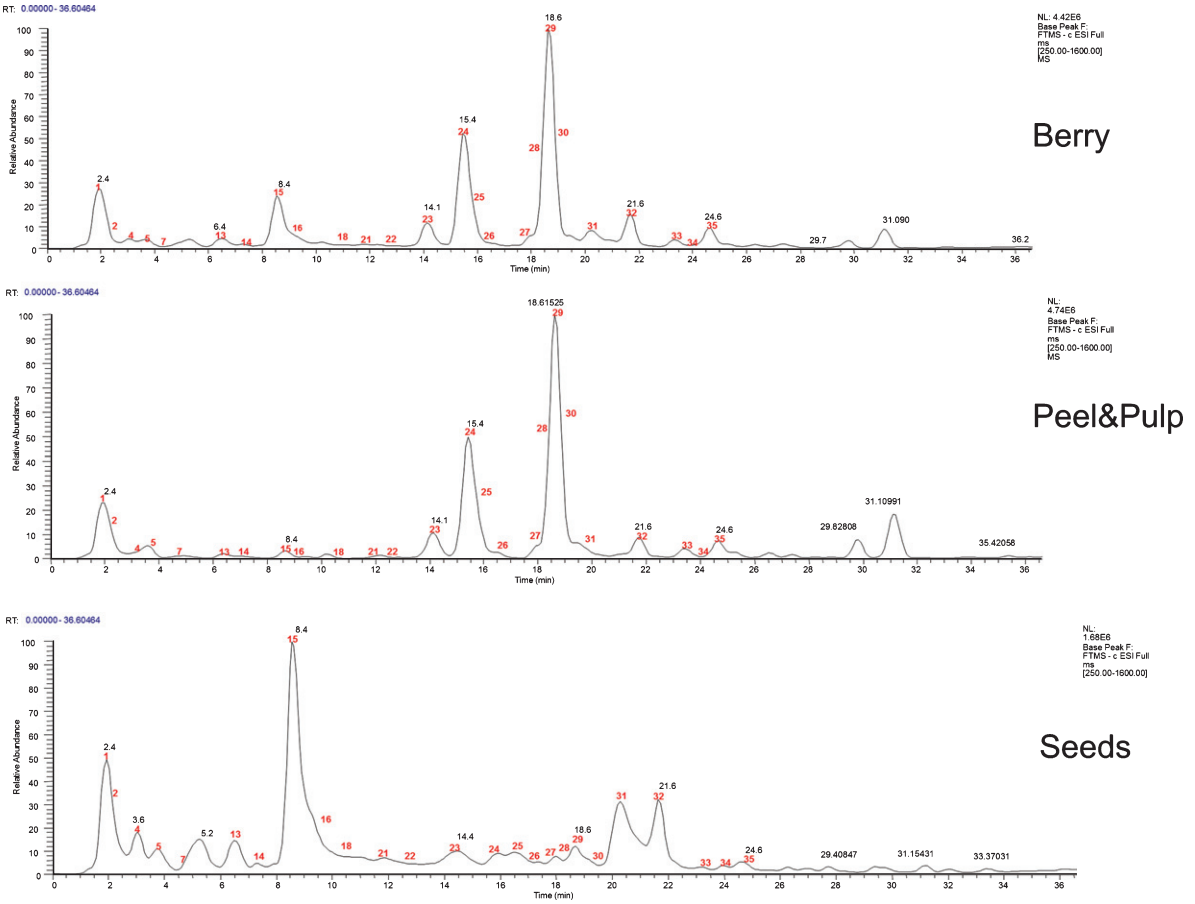

Fig. 1. Base peak chromatograms in negative ion mode of whole berry, Peel & pulp and seeds of *Myrtus communis* methanol extracts.

ment with the data reported in the literature [9]. Compound **5** showed a pseudo molecular ion at *m/z* 343.0664 corresponding to the formula $C_{14}H_{15}O_{10}$ that in MS/MS gave two principal product ions at 191 and 169, thus the compound was identified as galloylquinic acid, the same fragmentation pattern was reported by Taamalli et al. (2014) in the leaves of *Myrtus communis*. Compound **14** showed a pseudo molecular ion at *m/z* 781.0523 corresponding to the formula $C_{34}H_{21}O_{22}$, thus in comparison with database (Knapsack; mass bank) and data reported in literature [8], it was tentatively identified as punicalin. Compounds **13**, **15**, **16**, **25** and **26** were for the first time identified in the species of *Myrtus communis*.

Compound **13**, showed a pseudo molecular ion at *m/z* 785.0836 corresponding to the molecular formula $C_{34}H_{25}O_{22}$ that submitted to fragmentation gave a principal product ion at *m/z* 301 (loss of digalloylhexose), 483(loss of HHDP), can be identified as digalloylhexose, presumably tellimagrandin I; in fact a similar fragmentation was previously observed for this compound and it was tentatively identified by Boulekbache-Makhlouf et al. (2013) [23] in leaves of *Eucalyptus globulus* (family Myrtaceae). Compound **15** showed a pseudomolecular ion at *m/z* 783.0679 corresponding to the molecular formula $C_{34}H_{23}O_{22}$, whose MS/MS spectrum is reported as an example in Fig. 3. The pseudo molecular ion at *m/z* 783.0679 produced in MS/MS fragment ions at *m/z* 301 (ellagic acid; [M- 482]$^-$, loss of HHDP hexose) at *m/z* 481 (deprotonated HHDP hexose; [M- 302]$^-$, loss of HHDP) and a minor peak at 613(probably due to the loss of gallic acid [M-152]$^-$ and rearrangement

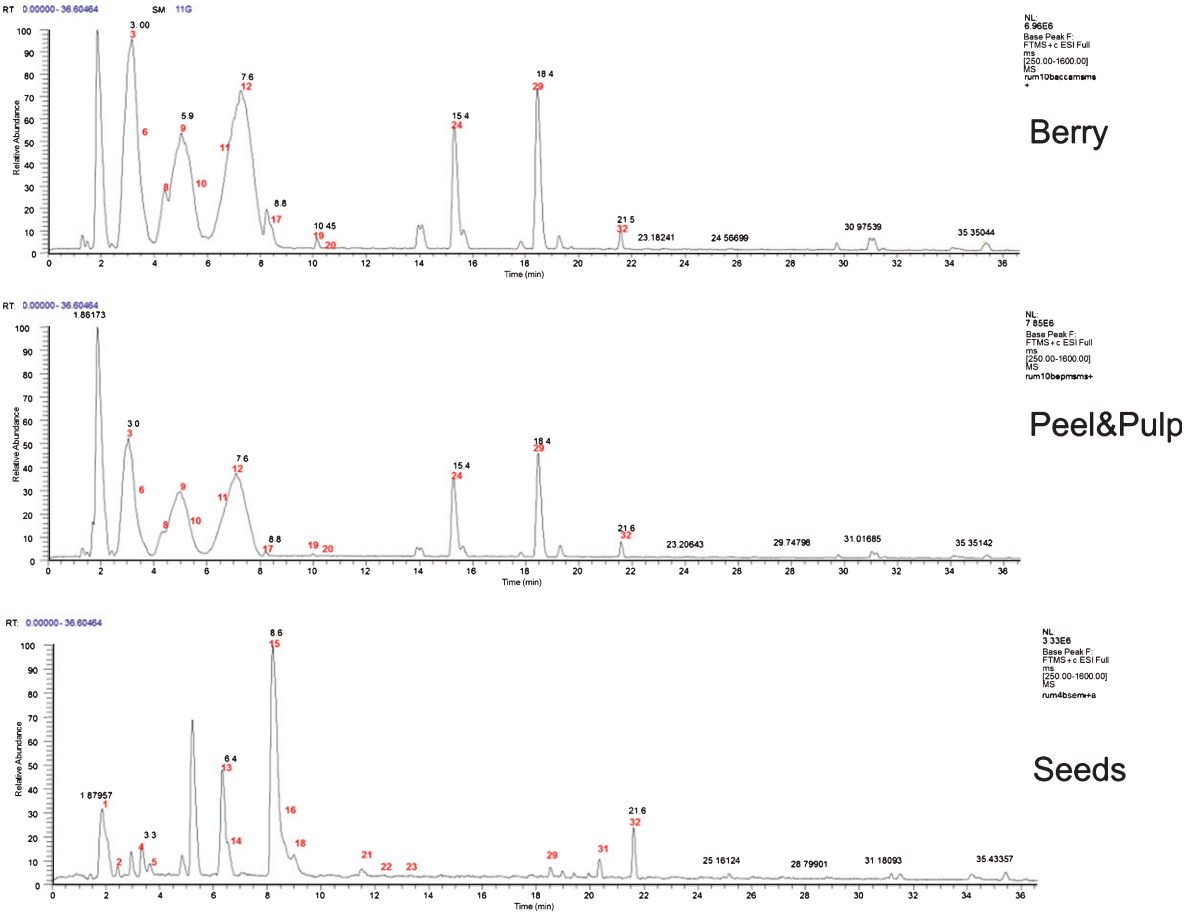

Fig. 2. Base peak chromatograms in positive ion mode of whole berry, Peel & pulp and seeds of *Myrtus communis* methanol extracts.

of the resulting ion), thus it was identified as di-HHDP hexose, presumably pedunculagin; the fragmentation pattern was previously reported by Simirgiotis et al. (2013) [24] and a similar fragmentation was observed by Boulekbache-Makhlouf et al. (2013) [23] in leaves of *Eucalyptus globulus* (family Myrtaceae). Compound **16** showed a pseudo molecular ion at *m/z* 935.0789 corresponding to the molecular formula $C_{41}H_{27}O_{26}$ that submitted to fragmentation gave three principal product ions at 783, 656 and 301. The fragmentation pattern agrees with data reported in literature and database [25] and thus the compound was identified as casuarictin.

Compound **25** showed a pseudo molecular ion at *m/z* 933.0632 corresponding to the molecular formula $C_{41}H_{25}O_{26}$, in comparison with database (Knapsack; mass bank) and data present in literature the compound was tentatively identified as castalagin, it was previously reported by Fujita et al. (2015) [26] in camu-camu fruit.

Compound **26** showed a pseudo molecular ion at *m/z* 937.0945 corresponding to the molecular formula $C_{41}H_{30}O_{26}$, in comparison with database (Knapsack; mass bank) and data reported in literature [24], it was tentatively identified as tellimagrandin II.

Compounds identified, pertaining to this class are reported in Fig. 4.

Compound **21** was identified as ellagic acid hexoside and compound **31** as ellagic acid, the two compounds were previously reported in *Myrtus communis* leaves [9, 27].

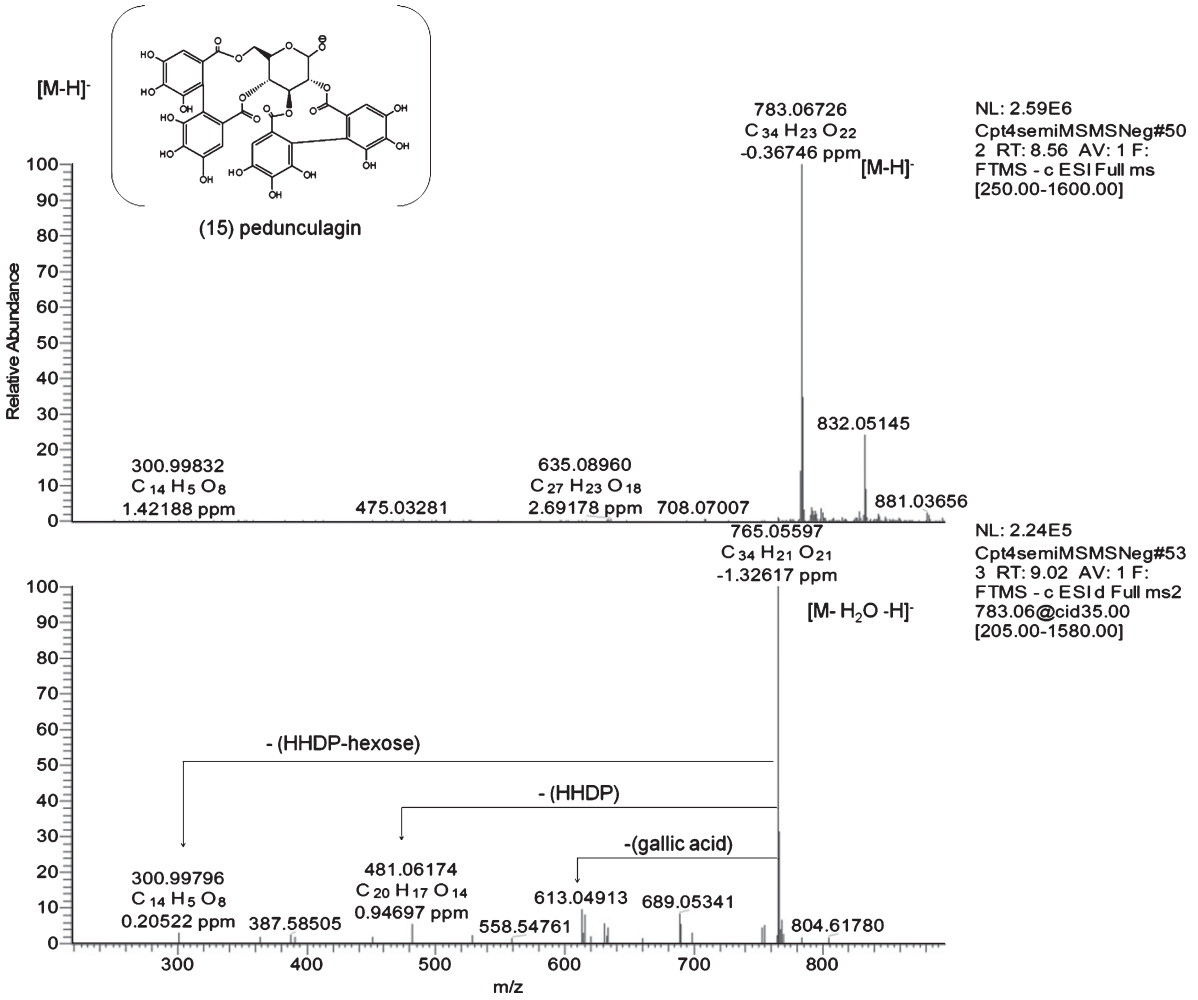

Fig. 3. MS/MS spectrum of pedunculagin (compound **15**).

3.1.2. Gallomyrtucommulones

This class of compounds was detected at 360 nm and identified through LC-MS and MS/MS experiments by operating in negative ion mode.

Compound **35** showed a pseudomolecular ion at *m/z* 567.2075 corresponding to the molecular formula $C_{27}H_{35}O_{13}$, producing in MS/MS fragment ion at 331/313/271/169. The same fragmentation pattern was previously reported by Taamalli et al., 2014 in leaves of *Myrtus communis*, thus the compound was identified as gallomyrtucommulone C.

3.1.3. Hydroxycinnamic derivatives

Compound **18** showed a pseudo molecular ion at *m/z* 341.0871 corresponding to the molecular formula $C_{15}H_{13}O_6$, producing in MS/MS two principal product ion 179 ([M-162]$^-$ loss of glucose moiety) and 161, in comparison with database and data present in literature [28], the compound was identified as caffeoylhexose;. This is the first time it was found in the myrtle berry.

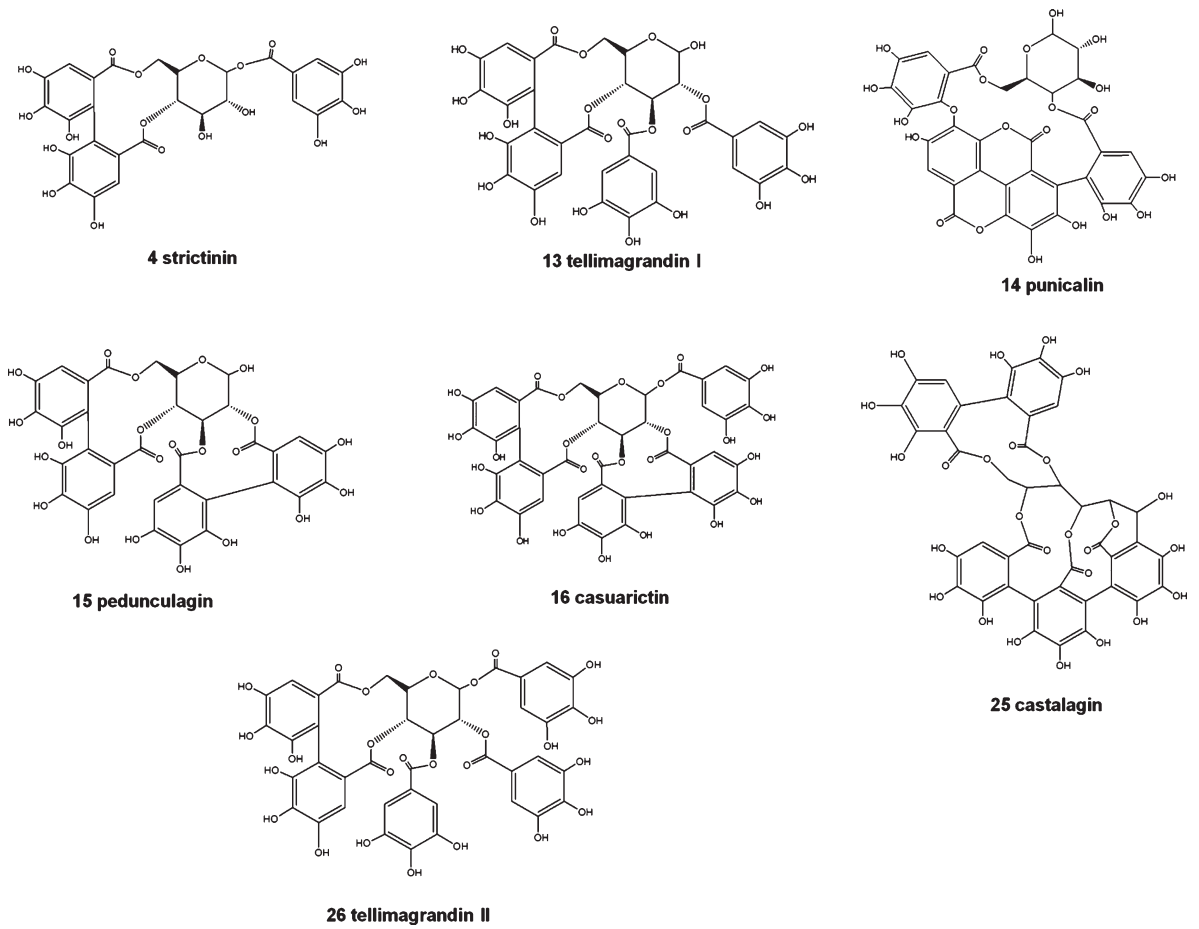

Fig. 4. Chemical structures of hydrolysable tannins tentatively identified for the first time in different parts of myrtle berries.

3.1.4. Flavanols

This class of compound was detected at 360 nm and identified through LC-MS and MS/MS experiments by operating in negative ion mode.

Compounds **7** and **22** were identified as flavanols; compound **7** was identified as epigallocatechin and compound **22** as catechin/epicatechin by comparison with standard compounds. They were previously reported in *Myrtus communis* berry [27].

3.1.5. Flavonols

This class of compound was detected at 360 nm and identified through LC-MS and MS/MS experiments by operating in negative ion mode.

Compounds **23**, **24**, **27**, **28**, **29**, **30**, **32**, **33** and **34** were identified as flavonols, most of them already reported in the myrtle berry; these compounds were present in all the parts of the fruits, but their content was higher in the peel and pulp than in the seeds (see Fig. 1 and 2).

Compounds **23**, **28**, **33** and **34** were identified for the first time in the Myrtle berry. Compound **23** showed pseudo molecular ion at *m/z* 631.0934 corresponding to molecular formula $C_{28}H_{23}O_{17}$, producing in MS/MS

two principal product ions at 479, which correspond to the loss of galloyl unit (152 Da) and 317 which could be a result of the loss of a hexosyl moiety from the molecule at *m/z* 479 and correspond to myricetin, The compound was tentatively identified as myricetin-galloyl-hexoside, which was previously reported by Taamalli et al. (2014) in the leaves of *Myrtus communis*.

Compound **28** showed a pseudo molecular ion at *m/z* 615.0985 corresponding to the molecular formula $C_{28}H_{23}O_{16}$, that in MS/MS gave two principal product ions 463, which correspond to the loss of galloyl unit (152 Da) and 301 which could be a result of the loss of a hexosyl moiety from the molecule at *m/z* 463 and corresponds to quercetin; in comparison with database and data reported in literature, the compound was tentatively identified as quercetin-galloyl-hexoside, which was previously reported by Romani et al. 2004 in leaves of *Myrtus communis* [29].

Compound **33** with pseudo molecular ion 615.0984 corresponding to the molecular formula $C_{28}H_{23}O_{16}$ showing two principal product ions 463, which correspond to the loss of galloyl unit (152 Da) and 317 which could be following a loss of a deoxyhexose moiety leading to myricetin. In comparison with literature, the compound was tentatively identified as myricetin-galloyldeoxyhexose, which was previously reported by Taamalli et al.(2014)in leaves of *Myrtus communis*.

Compound **34** showed a pseudo molecular ion at *m/z* 625.1403 corresponding to the molecular formula $C_{27}H_{29}O_{17}$, that in MS/MS gave two principal product ions 479, which correspond to the loss of deoxyhexose unit (146 Da) and 317 which correspond to the loss of deoxyhexose and hexose units. In comparison with literature the compound was tentatively identified as myricetin-deoxyhexose-hexose, which was previously reported in leaves of the *Myrtus communis* [29].

3.1.6. Anthocyanins

This class of compound was detected at 520 nm and identified through LC-MS and MS/MS experiments by operating in positive ion mode

Compound **3**, **6**, **8**, **9**, **10**, **11**, **12**, **17**,**19** and **20** were identified as anthocyanins, most of them already reported in *Myrtus communis* [30, 31]. These compounds are not present in the seeds and were identified operating in positive ion mode at the same LC condition used when negative ion mode was performed. Compounds **12** and **19** were reported for the first time in *Myrtus communis*. Compound **12** showed a pseudo molecular ion at *m/z* 419.0978 corresponding to the molecular formula $C_{20}H_{19}O_{10}^{+}$, that in MS/MS gave one principal product ion at *m/z* 287, which corresponds to the loss of a pentose moiety, thus in comparison with data reported in literature [32] the compound was tentatively identified as cyanidin pentoside. Compound **19** showed a pseudo molecular ion at 433.1134 corresponding to the molecular formula $C_{21}H_{21}O_{10}^{+}$ that submitted to fragmentation. It gave one principal product ion at *m/z* 301, which corresponds to the loss of a pentose moiety, thus in comparison with data reported in literature [32] the compound was tentatively identified as peonidin pentoside.

3.2. Multivariate data analysis

The Multivariate Data Analysis was basically performed following the experimental protocol described by D'Urso et al. (2015). LC-ESI-Orbitrap-MS chromatograms were pre-processed using MZmine (a free software) in a way to compensate for variations in retention time and *m/z* value between each chromatographic run. A peak list table was obtained from pre-processed chromatograms. In the final table, rows represented the individual samples (36 samples: 18 biological samples in technical duplicates), and columns represented integrated and normalized peak areas. Furthermore, this data was used through an approach of untargeted analysis and treated with an unsupervised Multi Variated Data Analysis method (PCA).

Principal Component Analysis (PCA) was performed by applying the peak areas of each peak observed in the LC-MS dataset (excluding the noisy), and a matrix was constructed by using these areas (variables), with the different analyzed samples as columns of the matrix. The resulted score scatter plot is reported

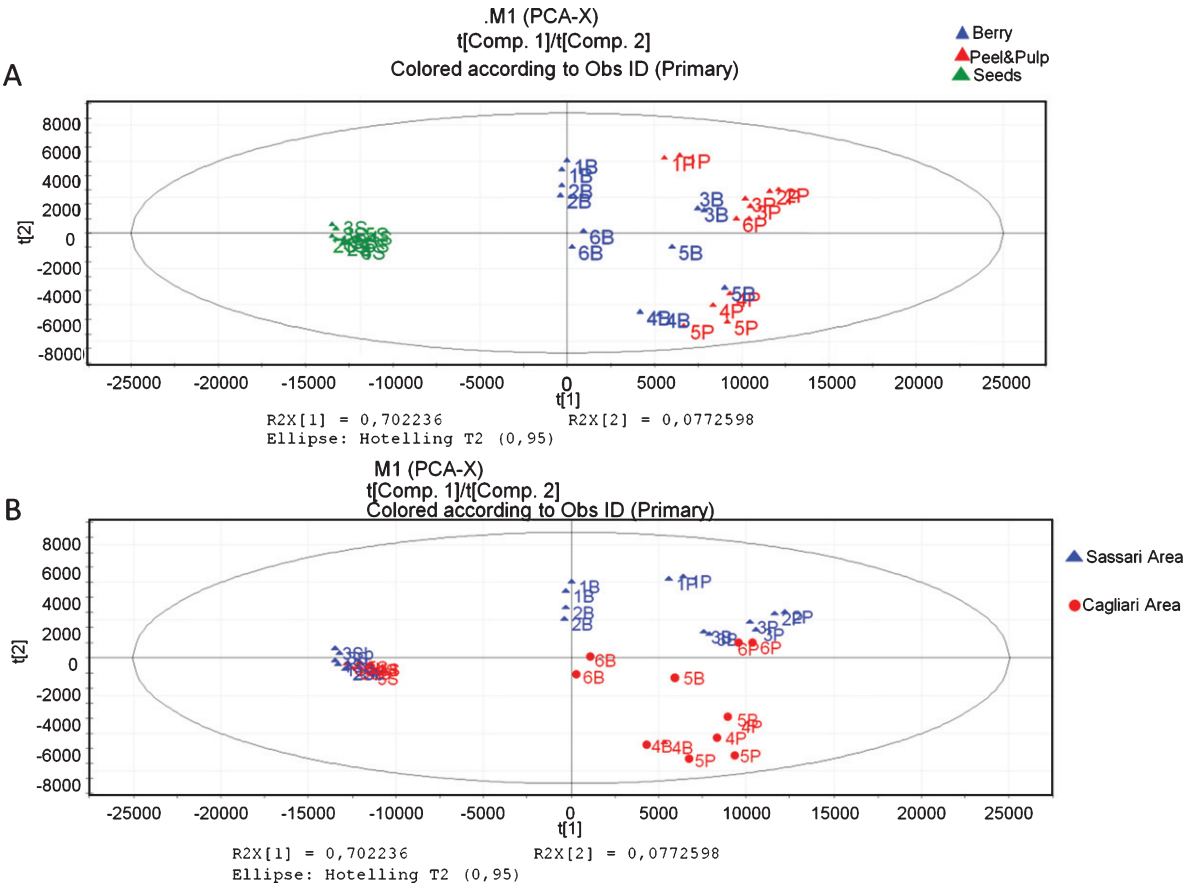

Fig. 5. Principal Component Analysis, score scatter plot, LC-ESI-Orbitrap-MS analysis. Panel A: colored according to myrtle berries parts. Panel B: colored according to geographic origin of samples.

in Fig. 5. The first component describes 70.22% of variance while the second only 7.72%. The choice of principal components was realized on the basis of the fitting (R2X) and predictive (Q2X) values for the PCA model.

The score scatter plot in Fig. 5A is colored according to the parts of fruit under analysis. There is clearly a differentiation of the samples based on this parameter: seeds are all grouped on the left part of the plot, peel & pulp and whole berry are distributed on the right part. In Fig. 5B the same scatter plot is displayed, colored according to the cultivar of samples (different cultivar is related to the different geographical origin of the seeds). By using this visualization mode, the peel and pulp appear as an important discriminator for samples of different cultivar and different geographic origin. Really, samples pertaining to the cultivar from the area of Sassari are on the right upper part of the plot while samples pertaining to the cultivar from the area of Cagliari are on the bottom right. However, this differentiation, mainly showed by peel and pulp samples is not present between the seeds, because all the samples are grouped in the same area. Thus, this is evidence that the metabolites that make the two cultivar different are the phenolic compounds, and principally anthocyanins and flavonoids, present in big amount specially in peel and pulp. In this study, these phenolic compounds appear to be marker compounds for geographic and genetic origin of myrtle berry.

4. Conclusion

An untargeted metabolomics approach together with the use of chemometric methods was developed for the discrimination of different cultivars of myrtle berry, whose seeds where collected from different geographic area of Sardinia and then grown in the experimental station of the University of Sassari located in Oristano (Sardinia Italy). For the first time methanolic extracts of different parts of myrtle berries were analyzed through liquid chromatography coupled with high resolution mass spectrometry. The study of metabolite profiling using liquid chromatography combined with high resolution mass spectrometry (LC-ESI-Orbitrap-MS) has proved to be a powerful tool for discovering changes of composition in different fruit parts of *Myrtus communis*. Through LC-ESI-Orbitrap-MS analysis and MS/MS experiments, 35 compounds were tentatively characterized on the basis of their retention time, UV/Vis absorbance, MS spectra and MS fragmentation patterns and a new class of hydrolysable tannins was identified in these berries (mainly in seeds) for the first time.

Moreover, using a large scale metabolomics approach, in this work we confirm that, in the classification based on geographic origin of the *Myrtus communis* berries, metabolites present in the peel and pulp are represented mainly by polyphenolic compounds like antocyanins and flavonoids and are more influent then metabolites present in the seeds, mainly gallotannins derivatives. Thus, flavonoids and anthocyanins, mainly found in the peel and pulp of the myrtle berry are putative marker compounds related to cultivar developed by seeds from different geographic origin.

Acknowledgments

This study was supported by a grant from the Autonomous Region of Sardinia; "Myrtle: genomics and metabolomics, an integrated approach to the industrial valorisation" CRP_60361_2012.

Conflict of interest

None to report.

References

[1] Zadernowsky R, Naczk M, Nesterowicz J. Phenolic acid profiles in some small berries. J AgricFood Chem. 2005;53:2118-24.

[2] Hancock RD, McDougall GJ, Stewart D. Berry fruit as 'superfood': Hope or hype. Biologist. 2007;54:73-9.

[3] Kool MM, Comeskey DJ, Cooney JM, McGhie TK. Structural identification of the main ellagitannins of a boysenberry (*Rubus loganbaccus baileyanus* Britt.) extract by LC–ESI-MS/MS, MALDI-TOF-MS and NMR spectroscopy. Food Chem. 2010;119:1535-154.

[4] Gasperotti M, Masuero D, Vrhovsek U, Guella G and Mattivi F. Profiling and accurate quantification of *Rubus* ellagitannins and ellagic acid conjugates using direct UPLC-Q-TOF HDMS and HPLC-DAD analysis. J Agric Food Chem. 2010;58:4602-16.

[5] Buzzini P, Arapitsas P, Goretti M, Branda E, Turchetti B, Pinelli P, Ieri F, Romani A. Antimicrobial and antiviral activity of hydrolisable tannins. Mini-Rev Med Chem. 2008;8:1179-18.

[6] Khan N, Afaq F, Kweon M, Kim K, Mukhtar H. Oral consumption of pomegranate fruit extract inhibits growth and progression of primary lung tumors in mice. Cancer Res. 2007;67:3475-82.

[7] Tuominen A, Toivonen E, Mutikainen P, Salminen J-P. Defensive strategies in *Geranium sylvaticum*. Part 1: Organ-specific distribution of water-soluble tannins, flavonoids and phenolic acids. Phytochemistry. 2013;95:394-407.

[8] Romani A, Campo M, Pinelli P. HPLC/DAD/ESI-MS analysis and anti-radical activity of hydrolyzable tannins from different vegetal species. Food Chem. 2012;130:214-21.

[9] Taamalli A, Iswaldi I, Arráez-Román D, Segura-Carretero A, Fernández-Gutiérrezc A, Zarrouka M. UPLC–QTOF/MS for a Rapid Characterisation of Phenolic Compounds from leaves of *Myrtus communis* L. Phytochem Analysis. 2014;25:89-96.

[10] Nuvole F, Spanu D. Analisi e prospettive economiche dell'utilizzazione industriale del mirto, Rivista Italiana EPPOS. 1996;12:231-6.

[11] Mulas MR, Melis AM. Essential oil composition of myrtle (*Myrtus communis*) leaves. J Herbs Spices Med Plants. 2011;17(1):21-34.

[12] Bouzabata A, Cabral C, Goncalves MJ, Cruz MT, Bighelli A, Cavaleiro C, Casanova J, Tomi F, Salgueiro L. *Myrtus communis* L. as source of a bioactive and safe essential oil. Food Chem Toxicol. 2015;75:166-72.

[13] Rahimmalek M, Mirzakhani M, Ghasemi PA. Essential oil variation among 21 wild myrtle (*Myrtus communis* L.) populations collected from different geographical regions in Iran. Ind Crops Prod. 2013;51:328-33.

[14] Piras FM, Dettori MF, Magnani A. ToF-SIMS PCA analysis of *Myrtus communis* L. Appl Surf Sci. 2009;255:7805-11.

[15] Melito S, La Bella S, Martinelli F, Cammalleri I, Tuttolomondo T, Leto C, Fadda A, Molinu MG, Mulas M. Morphological, chemical, and genetic diversity of wild myrtle (*Myrtus communis* L.) populations in Sicily. Turk J Agric For. 2016;40:249-61.

[16] Tuberoso CIG, Rosa A, Bifulco E, Melis MP, Atzeri A, Pirisi FM, Dessí MA. Chemical composition and antioxidant activities of *Myrtus communis* L. berries extracts. Food Chem. 2010;123:1242-51.

[17] Wannes WA, Marzouk B. Differences between Myrtle fruit parts (*Myrtus communis* var Italica) in phenolics and antioxidant contents. J Food Biochem. 2013;37:558-94.

[18] Mulas M, Francesconi AHD, Perinu B. Myrtle (*Myrtus communis* L.) as a new aromatic crop: Cultivar selection. J Herbs Spices Med Plants. 2002c;9(2/3):127-31.

[19] De Vos RCH, Moco S, Lommen A, Keurentjes JJB, Bino RJ, Hall RD. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat Protoc. 2007;2:778-91.

[20] D'Urso G, D'Aquino L, Pizza C, Montoro P. Integrated mass spectrometric and multivariate data analysis approaches for the discrimination of organic and conventional strawberry (*Fragaria ananassa* Duch.) crops. Food Res Int. 2015;77:264-72

[21] Mari A, Montoro P, Pizza C, Piacente S. Liquid chromatography tandem mass spectrometry determination of chemical markers and principal component analysis of *Vitex agnus-castus* L. fruits (Verbenaceae) and derived food supplements. J Pharmaceut Biomed. 2012;70:224-30.

[22] Safer S, Serhat Cicek S, Pieri V, Schwaiger S, Schneider P, Wissemann V, Stuppner H. Metabolic fingerprinting of *Leontopodium* species (Asteraceae) by means of 1H NMR and HPLC–ESI-MS. Phytochemistry. 2011;72:1379-89.

[23] Boulekbache-Makhlouf L, Meudec E, Mazauric J-P, Madania K, Cheynier V. Qualitative and semi-quantitative analysis of phenolics in *Eucalyptus globulus* leaves by High-performance Liquid Chromatography coupled with diode array detection and electrospray ionisation mass spectrometry. Phytochem Analysis. 2013;24(2):162-70.

[24] Simirgiotis MJ, Bórquez J, Schmeda-Hirschmann G. Antioxidant capacity, polyphenolic content and tandem HPLC–DAD–ESI/MS profiling of phenolic compounds from the South American berries *Luma apiculata* and *L. chequén*. Food Chem. 2013;139:289-99.

[25] Grace MH, Warlick CW, Neff SA, Lila MA. Efficient preparative isolation and identification of walnut bioactive components using high-speed counter-current chromatography and LC-ESI-IT-TOF-MS. Food Chem. 2014;158:229-38.

[26] Fujita A, Sarkar D, Wu S, Kennelly E, Shetty K, Genovese MI. Evaluation of phenolic-linked bioactives of camu-camu (*Myrciaria dubia* Mc. Vaugh) for antihyperglycemia, antihypertension, antimicrobial properties and cellular rejuvenation. Food Res Int. 2015;77:194-203.

[27] Barboni T, Venturini N, Paolini J, Desjobert J-M, Chiaramonti N, Costa J. Characterization of volatiles and polyphenols for quality assessment of alcoholic beverages prepared from Corsican *Myrtus communis* berries. Food Chem. 2010;122:1304-12.

[28] Jaiswal R, Matei MF, Glembockyte V, Patras MA, Kuhner N. Hierarchical key for the LC–MS^n identification of all ten regio- and stereoisomers of caffeoylglucose, J AgrFood Chem. 2014;62:9252-65.

[29] Romani A, Coinu R, Carta S, Pinelli P, Galardi C, Vincieri FF, Franconi F. Evaluation of antioxidant effect of different extracts of *Myrtus communis* L. Free Radical Res. 2004;38(1):97-103.

[30] Montoro P, Tuberoso CIG, Piacente S, Perrone A, De Feo V, Cabras P, Pizza C. Stability and antioxidant activity of polyphenols in extracts of *Myrtus communis* L. berries used for preparation of myrtle liqueur. J Pharmaceut Biomed. 2006;41:1614-9.

[31] Montoro P, Tuberoso CI, Perrone A, Piacente S, Cabras P, Pizza C. Characterisation by liquid chromatography-electrospray tandem mass spectrometry of anthocyanins in extracts of *Myrtus communis* L. berries used for the preparation of myrtle liqueur. J Chromatogr A. 2006;1112:232-40.

[32] Gardana C, Ciappellano S, Marinoni L, Fachechi C, Simonetti P. Bilberry adulteration: Identification and chemical profiling of anthocyanins by different analytical methods. J Agr Food Chem. 2014;62:10998-1004.