UNIVERSITÀ DEGLI STUDI DI SALERNO
Dipartimento di Farmacia

Dottorato di ricerca
in Scienze Farmaceutiche
Ciclo XIV—Anno di discussione 2016

Coordinatore: Chiar.mo Prof. *Gianluca Sbardella*

INTEGRATED METABOLOMICS APPROACHES FOR BERRY FRUIT USED IN NUTRACEUTICAL FORMULATIONS

settore scientifico disciplinare di afferenza: BIO/15

Dottorando   **Tutore**

Dott. *Gilda D'Urso*   **Chiar.ma Prof.** *Paola Montoro*

Preface

My PhD in Pharmaceutical Science at the Department of Pharmacy of Salerno University was started in January 2013 under the supervision of Prof. Paola Montoro.

Targeted and *untargeted* metabolomics approaches were applied to investigate the effect of environment, genotype, or both on variation of many metabolites in some berry fruits.

In particular the species under investigation during these three years of PhD course were: *Fragaria ananassa, Fragaria vesca*, *Morus alba*, *Morus nigra* and *Myrtus communis*. All these species are characterized by the production of small fruit, and all of them are plant species that can be used for the formulation of plant food supplements, in fact they are reported into the official list of Italian legislation (DM July 9, 2012-G.U. 21-7-2012, serie generale n. 169, and update on March 27, 2014). Some of them are recognized as Italian traditional food products, like *Fragaria vesca*, typical of Campania region and *Myrtus communis*, endemic of Sardinia region.

Furthermore, to improve my knowledge on metabolomics software, I spent a period of research activity at the Department of Bioscience of Wageningen University, Netherlands, under the supervision of Dr. Ric de Vos, where I focused my reaserch on *Morus alba* and *Morus nigra*.

During the three years course of PhD, liquid chromatography coupled to high resolution mass spectrometry and in conjunction with Multivariate Data Analysis were applied to investigate the metabolic composition of different berry species; with these approaches we could confirm that metabolomics represent a useful tool to evaluate the contribution of environmental and genetic factors to the differences in metabolite composition or content of berry fruit. Moreover, with the characterization of polar compounds we could also

assert the berry fruit as important food products with a recognized value in the production of food supplements.

Table of contents

Chapter 1: **Metabolomics and "Berry" fruit**

Chapter 2: **Metabolomics approaches used for discrimination of different cultural system**

2.1 Integrated Mass Spectrometric and Multivariate Data Analysis approach for the discrimination of organic and conventional strawberry (Fragaria ananassa Duch.) Crops

Chapter 3: **Metabolomics approaches used for typization of Berry fruit**

3.1 Characterisation of Fragaria vesca fruit from Italy following a metabolomics approach through integrated Mass Spectrometry techniques

3.2 Targeted and Untargeted Mass Spectrometric approaches in discrimination between Myrtus communis cultivars from Sardinia region

Chapter 4: **Metabolomics approaches used for identification of bioactive metabolites**

4.1 Combined use of α- glucosidase inhibitory activity and UHPLC-High Resolution mass spectrometry for the identification of active metabolites in Morus alba and Morus nigra extracts

Chapter 1
Metabolomics and "Berry" fruit

1 Introduction

The term "berry fruit" generally refers to any small fruit that can be eaten whole. Botanically, a berry is the most common type of fleshy fruit in which the entire ovary wall ripens into an edible pericarp. They may have one or more carpels with a thin covering and fleshy interiors, in which the seeds are usually embedded. Many fruits, commonly known as berries, are not actual berries by the scientific definition, but fall into one of these categories, for example blackberry, mulberry and strawberry[1]. Berries are widely recognized for their several health-promoting properties including a reduced risk of cancer, cardiovascular disease, and other chronic disease. These properties may be due to the presence of polyphenols, particularly anthocyanins, which are a class of flavonoids that produce the bright red, purple, and blue pigments seen in many berries and are among the most consumed flavonoids in the american diet[2,3].Polyphenols are a class of secondary metabolites found throughout the plant kingdom and can be categorized into a variety of families, for example phenolic acids, flavonoids, anthocyanins, stilbenes, coumarins, and tannins. In plants, phenols are essential to plant pigment, growth, reproduction, resistance to pathogens, and mediation of solar radiation and photosynthesis byproducts[3].Moreover, when phenolic compounds are introduced to biological systems, they show antiallergic, antianxiety, anticarcinogenic, anti-inflammatory, antioxidant, antiproliferative, antitumorigenic, and antiviral properties[3]. Howewer systemic bioactivities ascribed to these compounds seems to be largely mediated by their metabolites, in fact, a large proportion of potentially protective berry

polyphenols are unable to enter the circulation and influence cellular interaction[4]. Therefore most polyphenols from berries remain in the gastrointestinal tract and pass through to the large intestine where they are subject to biotransformation by colonic microbiota[5]. Those compounds may also modulate nutrient availability through the inhibition of digestive enzymes involved in lipid and starch breakdown, which could influence obesity and blood glucose control[5], and due to their high concentration and qualitative diversity in berry fruit, these fruit are increasingly often referred to as natural functional food.

Due to climatic conditions, fresh berries are generally available several months a year while some of the harvested fruits are processed to juice, fruit beverages, frozen product wine, jam and jelly. Studies report that constitutional and sensory qualities of berries are affected by many factors such as cultivars, geographic region, storage conditions, ripeness, climate and others may affect the concentration of phenolic compounds and in the antioxidant capacity[6,7].

Furthermore, there has been a growing trend in the use of berry extracts as ingredients in functional food and dietary supplements, and omics "approaches" such as genomics, proteomics and metabolomics, are increasing the possibility of carrying out better strategies to improve crops (including berry) for fruit quality enriched in bioactive compounds with nutraceutical properties[8].

1.2 Common Berry found in Italian regions

Several plant species characterized by the production of berry fruit are cultivated or spontaneous in Italy. Most of these species are present in the list

of plant species that can be used for the formulation of plant food supplements (DM July 9, 2012)[9].

They are: blackberry (*Rubus ulmifolius* L.), highbush blueberry (*Vaccinium corymbosum* L.), black currant (*Ribes nigrum* L.), white currant (*Ribes rubrum* L.), gooseberry (*Ribes grossularia* L.) black mulberry (*Morus nigra* L.), white mulberry (*Morus alba* L.), raspberry (*Rubus idaeus* L.), strawberry (*Fragaria × ananassa* Duch.), wild strawberry (*Fragaria vesca*) and myrtle berry (*Myrtus communis*).

Some of these species are endemic of Italian regions like *Fragaria vesca*, endemic and typical of Campania region in particular from "Alburni" and "Alto Sele" areas. In this area, the wild strawberry fruit are recognized as a traditional food product under the name "Fragolina degli Alburni". The fruits are used for the production of liqueur and sweets. In medicine, fruits and leaves of this specie are used in the formulation of food supplements for their antioxidant activities. Another endemic species is *Myrtus communis*, that is a pleasant annual shrub endemic of Sardinia, commonly used for the production of myrtle liqueur. Fruit and leaves of *Myrtus communis* are also used in the preparation of food supplements used for fluidity of bronchial secretions and for balsamic effect. White mulberry (*Morus alba*) and black mulberry (*Morus nigra*) are deciduous tree originating from China and Japan, spread in America and Europe for the silkworm breeding. In south Italy trees of *Morus alba* and *Morus nigra* are common spread and are used for the production of juice or jam. Fruit and leaves are used in the preparation of food supplements for their several properties like: antioxidant, antihyperglicemic and fluidity of bronchial secretions.

1.3 Foodomics

Foodomics is defined as a discipline that studies the food and nutrition domains through the application of -omics technologies to characterize and demonstrate the beneficial effects on human health of food ingredients. Foodomics includes nutritional genomics sub-reaserch areas, such as nutrigenomics, nutrigenetics, proteomics and metabolomics [10].

In the last two decades, large investments were made to develop analytical approaches to analyze the different cell products, such as those from gene expression (transcripts), proteins, and metabolites. All of these so-called 'omics approaches, including genomics, transcriptomics, proteomics, and metabolomics, are considered important tools to be applied and utilized to understand the biology of an organism and its response to environmental stimuli or genetic perturbation[10].

1.4 Metabolomics

Metabolomics as follow-on from transcriptomics and proteomics was a term coined at the end of the 1990s[11]. the concept entailed the analyses of the metabolite composition of biological material, aimed to be fully complementarily with the others potentially unbiased or non targeted "-omics" approaches[12].

Metabolome can be defined as the full set of endogenous or exogenous low molecular weight metabolic entities of approximately <1,000 Da (metabolites).

Metabolites are, in general, the final downstream products of the genome, and reflect most closely the operation of the biological system, its phenotype. One of the main challenges in metabolomics is to face the complexity of any metabolome, usually composed by a huge number of compounds of very

4

diverse chemical and physical properties (sugar, amines, aminoacids, peptides, organic acid nucleic acid or steroids). Sample preparation is especially important in metabolomics, because the procedure used for metabolite extraction has to be robust and highly reproducible. Sample preparation will depend on the sample type and the targeted metabolites of interest (fingerprint or profiling approach)[11].

Metabolomics has proven to be a valuable tool for the comprehensive profiling of plant derived samples for the study of plant systems and natural products research, especially when combined with chemometric data analysis approaches[13].

From the methodological point of view, there are basically only two different approaches used in any metabolomics study: "targeted analysis" and "metabolite profiling". Targeted analysis is restricted to quantitative analysis of a class of compounds that are related to a specific pathway or to intersecting pathways. Targeted analysis is very useful for the study of the primary effect of a genetic alteration, and the analytical procedures must include identification and absolute quantification of the selected metabolites in the sample. Metabolite profiling (or, sometimes, metabolic profiling), on the other hand, involves rapid analysis, often not quantitative, of a large number of different metabolites with the objective to identify a specific metabolite profile that characterizes a given sample[14].

1.4.1 Setup of metabolomics experiments

The analysis of data obtained from metabolomics experiments is complex. Different platforms and experimental designs produce data in various format and units, making the study of metabolomics data difficult. Providing full information regarding the experimental design a sample preparation route for

the development of an analytical methodology and an approach to data analysis that includes all such information and is simultaneously easy to handle is challenging task[15]. However, a universally applicable standard procedure has not yet been established, and single publications only include suggestions that will have to be refined in the future[16].

Sample preparation

The sample preparation is an important step in metabolomics analysis; in fact it can direct influence the results of the analysis. Basically it can be divided in the following four steps: 1) harvesting of the plant materials, 2) processing before extraction (for example drying), 3) extraction and 4) pre-analytical sample preparation. The execution of all these steps may not be necessary and can be avoided in some cases depending on the aim of research and the nature and properties of the analytes of interest[15].

Harvesting

Metabolites are influenced by environmental factor. Seasonality, developmental stage and age, circadian rhythm, temperature, water availability, UV radiation, altitude, atmospheric composition, soil nutrients, and tissue damage may induce qualitative and or quantitative variations in the metabolites composition[17]. Moreover a plant metabolome may vary depending on the time or place of harvest. The influence of environmental conditions is extremely relevant for plant material collected in the natural habitat of plants; however, it may be considered minor for plants cultivated in a greenhouse or in cell culture plants[16,18]. Accurate timing is an important factor that should be considered because the levels of metabolites (primary as well as secondary metabolites) vary throughout the day, due to the different light exposition of

6

the plant[19,20]. Moreover, the developmental stage of the plant at the time of harvest affects the metabolite profile[16,21]. Additionally, different plant organs also show significantly different metabolite profiles. For this reason the separate analysis of different plant organs with separation by tissue age is fundamental to obtaining reliable results[15].

Processing before extraction

Changes in the plant's metabolism occur from seconds up to a few minutes; for this reason, harvesting must be performed as quickly as possible and metabolism should be stopped immediately after harvesting by freezing of plant material in liquid nitrogen[16,21]. An appropriate technique would be freeze-drying, which is based on lyophilization and cryodesiccation. In this way, water is removed and the enzymatic activity is reduced. In addition to freezing or freeze-drying, the samples may also be subjected to a drying process by an oven, air or by microwave heating, which destroy enzyme activities [22]. The selection of a drying method should always consider the molecules to be targeted for analysis because some methods are incompatible with certain metabolites, such as volatile and thermo-unstable compounds, among others. Further processing before extraction can include the pulverization of the plant material into small particles to improve the extraction process[15].

Extraction

Extraction methods should be as simple and as fast as possible[23]. There is no current method for the extraction of all metabolites[18,21]. Plant tissue metabolites are highly diverse. Several critical factors must be considered during extraction, such as the ratio of solvent and plant material, solvent

characteristics, the time of extraction, the temperature and the choice of an appropriate method for the desired goals. The choice of a solvent is extremely important for the achievement of reliable results because it needs to be adequate for the metabolites targeted for extraction and for the analytical method[15]. Among solvent extraction methods, methanol and solvent mixtures that contain methanol are employed the most; furthermore acidified solution are common, and chloroform is often used for lipophilic compounds[19,23,24]. Several techniques have been applied to accelerate solvent extraction, such as ultrasonic extraction, microwave and pressurized solvent extraction [19,24]. Ultrasonic extraction is also rapid and simple and is one of the most commonly used methods for solvent extraction at present[26].

1.4.2 Metabolite analysis

Two analytical platforms are currently used for metabolomics analysis: Mass Spectrometry (MS) and Nuclear Magnetic Resonance (NMR). These techniques either stand alone or combined with separation techniques (typically, LC-NMR, GC-MS, LC-MS and CE-MS), can produce complementary analytical information to attain more extensive metabolome coverage[27]. MS and NMR-based technologies are both complementary and, therefore often used in parallel in metabolomics research. Compared to NMR, MS is a more sensitive technique; also MS coupled to GC, LC or CE allows higher resolution and sensitivity for low abundance metabolites. The use of high and ultra-high resolution analyzers (namely, TOF, FTMS, Orbitrap®) is essential to obtain accurate mass measurement for the determination of elemental compositions of metabolites and to carry out their tentative identification with database[28]. On the other hand, MS/MS or MS^n experiments, especially when product ions are analyzed at high resolution (with Q-TOF,

TOF-TOF, or LTQ-Orbitrap®) provide additional structural information for the identification of metabolites[29].

ESI source: ion generation and ion suppression

The ionization source that is generally used in metabolomics studies is electrospray ionization (ESI). Notwithstanding its widespread use, it is still a young technique and for this reason, fundamental studies about reaction in the ionization source, as well as fragmentation patterns (MS/MS or MS^n) are still being investigated[15].

In an initial step, the solution containing the analyte crosses a metallic capillary where an electric potential is applied (between 2 to 4.5 kV). This potential promotes a migration of charges to the interface of the capillary with the solution, which results in an electric double layer, and thus droplets with charged surfaces are formed. The process of solvent evaporation starts within a chamber with a slight reduction of pressure under a nebulizer gas. With the reduction of droplet sizes, an approximation of the charges occurs, and consequently, the electrostatic repulsive forces are increased, which leads to a decrease in the surface tension of the droplets until rupture, a process denoted as "Coulomb explosion". This rupture results in the liberation of ions into the gas phase by a spray of charged particles, or in other words, an electrolytic current[29]. A positive ionization mode is used in a large majority of metabolomics studies and the addition of H^+ and Na^+ ions must be balanced. In metabolomics studies using the hyphenated system LC-ESI-MS, the co-elution of substances with a large difference in the ionisation potential can lead to the complete suppression of an ion, leading to an incorrect result indicating the non-appearance of a metabolite in question[15].

9

Furthermore, it is also known that the use of hyphenated systems as well as the use of direct infusion mass spectrometry, still present several limitations, such as ion suppression, radical species formation and cross reactions[15].

MALDI source: Ion generation and matrix effects

MALDI (matrix-assisted laser desorption/ionization) was introduced to the market with great success, principally in proteomic research. MALDI can be explained in the following three basic steps: 1) formation of a solid solution in which analytes are completely isolated from each other through dilution in an homogenous matrix for subsequent desorption; 2) excitation of the matrix, in this step a laser beam is applied to the matrix surface. Photons from the laser are absorbed by the chromophore parts of the matrix molecules, which causes rapid vibrational excitation followed by disintegration. Cluster of analyte molecules surrounded by the matrix and salt ions are ejected at the solution surface, producing the plume and subsequent evaporation of the matrix molecules to expose free analyte molecules to the gas phase; 3) ionization of the analyte, in the desorbed matrix, cloud ionisation (plume) occurs via proton or cation transference from the photo-activated matrix molecules, which leads to the formation of characteristic ions $[M+X]^+$ (X= H, Na, K etc). In these steps, the process exhibits limitations for the analysis of organic micromolecules. This limitation was caused by the matrix, which is cationised or anionised after desorbition and extracted into the analyser. For this reason the mass range between 0 and 500 *m/z* may have shown numerous matrix signals, complicating any analysis of low molecular weight compounds[31,32]. MALDI is a promising techniques for metabolomics studies because of its high sensitivity, high analysis scan speed and lower contaminant influence on

ionisation[15] but generally is used for proteomic analysis in which the molecular weight of the analytes of interest is higher.

MS/MS and MS^n in metabolomics

The use of fragmentation spectra can help in the structural identification of metabolites. A series of databases has been established and with each day the information increases. An additional step is necessary for MS/MS analyses, namely, the isolation of an ion, which as an increased internal energy after receiving some type of activation that allows the beginning of the fragmentation process. Collision induced dissociation (CID) is without a doubt the most frequently used process[33,34]. Kinetic energy transference occurs through collision with the gas, and the ion is transformed by internal energy, inducing the fragmentation. MS/MS or MS^n is extremely important for metabolomics studies, but for the interpretation of data, a detailed observation of the ions that show abnormal behavior is necessary because the reason might merely be the effects of the internal energy quantity of the analysed ion[15].

Hyphenated mass spectrometry methods: GC/LC-MS

Prior to MS analysis, metabolites, are separated by a separation-based method, namely gas chromatography (GC), liquid chromatography (LC) or capillary electrophoresis (CE). In GC and LC, compounds are separated by exploiting different interaction of the substances with the stationary phase. However in CE, the separation is based on the size to charge ratio of the ionic molecules[35]. The choice of separation technique is made according to the type of molecules present in the target sample. GC is suitable for hydrophobic, low molecular weight compounds, that must be heat stable and volatile[35]. The ionization method that is most commonly used in combination with GC is electron

ionisation (EI), that is an hard ionization[36]; while in LC-MS soft ionization is most commonly used, such as electrospray ionization, where it is used to analyse a wide array of metabolites from sugar to fatty acids. The separation based MS methods are higly sensitive and the costs are relatively low compared to NMR methods[20,35]. For GC-MS new columns, such as narrow bore columns, were developed to reduce the analysis times to approximately 1 and 10 min, namely the Ultra Fast-GC and Fast-GC [37,38]. In LC-MS, monolitich columns and fused core particle (2.7 µm) packing columns have been used to reduce the time and increase the resolution. The stationary phase of C18 (reverse phase) is most frequently used in plant metabolomics, despite new phases such as HILIC[39,40,41].

Direct injection MS (DIMS)

Direct injection MS (DIMS) is an analysis performed by directly injecting a sample into the ionization souce. Soft ionization techniques are preferably applied in metabolomics studies because they cause less fragmentation than hard ionization techniques, resulting in a smaller number of signals and, consequently, less complex spectra[35]. The ions are generated in the ionization source, and they are subsequently separated in the analyser for the detection, so a mass spectrum is produced. The analysers can be of two types, namely, low or high resolution, and the latter enables high mass accuracy. Despite the high cost of the high resolution equipment, it is possible to obtain high repeatability, peak annotation and more precise data alignment[15]. Direct MS analysis of foods or food extracts has been demonstrated to be a useful approach in metabolic fingerprinting when rapid classification of food- sample types or rapid screening of food adulteration is wanted[42].

The main disadvantages of the DIMS are the impossibility of distinguishing between chemical isomers and the lower quantity of chemical information that is obtained relative to hyphenated MS methods. A possible way to circumvent the chemical isomer problem is to perform MS^n analyses. By performing MS^n analyses, previous chromatographic separations can be avoided if information obtained by the fragmentation patterns can be used to differentiate the isomers. Furthermore, chemical information can also be increased by performing analyses in negative as well as in positive ion modes[24]. The type of mass analyser and the ionization method have a considerable influence on the resulting mass accuracy, resolution and detection limit, and advantages and disadvantages depend higly upon the chosen instrument type[15].

Ambient MS

The main characteristic of ambient MS (AMS) is that it allows direct analysis of samples in open air with little or no sample preparation. For this reason MALDI and API techniques, such as ESI and APCI, are not considered to belong to this group. Since the usually still require sample preparation[42]. The development of AMS was initiated with the introduction of desorption electrospray ionization (DESI) by Cooks in 2004[43].

Almost 30 ambient sampling/ionization approaches were involved in MS analysis. Among them DESI and direct analysis in real time (DART) were the two most prevalent techniques. DESI shares the advantages of the matrix-free DIOS (laser desorption/ionization MS on porous silicon) and the advantageous production of multiply charged ions of ESI. DESI-MS has also be demonstrated to be promising tool in food safety control[44]. DESI was followed by DART in 2005[45]. DART can be considered an API (Atmospheric Pressure Chemical Ionization)-related techniques based on the thermo-desorption of

condensed phase analytes by a (distal) plasma discharge in a heated gas stream. As other ambient ionization techniques DART is undergoing rapid development and it is beginning to deliver its potential in metabolomics[45]. Other techniques used for identification of metabolites are: Laser ablation electrospray ionisation (LAESI)[47,48] and paperspray[49].

1.4.3 Preprocessing and statistical analysis of plant metabolomics data

Metabolomics experiments produce an enormous amount of data. Mathematical and statistical skills are essential for extracting as much information as possible from metabolomics data[26,50]. Data preprocessing is an important step in metabolomics data analysis, and it can influence data interpretation. Preprocessing includes noise filtering, normalization, peak detection, alignment and identification. Noise filtering is designed to separate compound signals from background signals originating from chemical or instrumental interference. Normalization is applied to correct the systematic variation and for direct comparison of different samples[15]. In different samples, slightly different *m/z* values may be obtained for the same compound. To compare different samples, alignment must be performed, and depending on whether individual metabolites are identified or not, finding out the chemical identity of a compound may also be necessary[49]. The metabolomics community is continuously growing, and new and exciting analytical strategies are steadily being developed to increase the amount of information extracted from mass spectrometric data sets. Such strategies include the development of software for the facilitation of data exchange[52]. The analysis of the large data sets generated by LC-MS requires data processing tools such as those based on multivariate data analysis. These techniques are robust to noise and missing data and enable one to deal with correlated variables;

14

different software such as the commercially available Markerlynx (Micromass Ltd., Manchester, U.K.) or freeware such as metAlign (Plant Reaserch International, Wageningen, The Netherlands) MzMine or XCMS perform the automatic extraction, alignment, and retention time correction of chromatographic peaks within individual mass to charge value using different algorithms.

These software generate a data table constituted by N observations (columns) and K variables (rows); the data table obtained must be analyzed in order to have information and to interpreted the results in a comprehensible way, for instance as a graph, and the MultiVariate Data Analysis (MVDA) is used for this purpose.[53]

1.4.4 Statistical analysis: Principal Component Analysis and Partial Least Square

Two are the multivariariate projection methods used for extracting information from a large or small table of data: Principal Component Analysis (PCA) and Partial Least Square (PLS).

Principal Component Analysis (PCA)

The PCA is an unsupervised method. The starting point for PCA is a matrix of data with N rows (*observations*) and K columns (*variables*). The observation can be analytical samples, chemical compounds or reactions, process time points of a continuous process, and so on. In order to characterize the properties of the observations, one measures variables. These variables may be of spectral origin (NIR, NMR, IR, UV, X-ray), chromatographic origin (HPLC, GC, TLC) or they may be measurements from sensors in a process (temperature, flows, pressures, curves, etc.)[53].

The most important use of PCA is to represent the multivariate data table as a low-dimensional plane, usually consisting of 2 to 5 dimensions, such that an overview of the data is obtained. This overview may reveal groups of observation, trends, and outliers. This overview also uncovers the relationships between observations and variables, and among the variables themselves[53].

Principal component analysis summarizes the variation of a data matrix X, as product of two low-dimensional matrices, T and P, which can be easily overviewed and used. The data in the multi-dimensional space, defined by the measured variables, is modeled as a plane or hyperplane, the axes of which are called the principal components. Each principal component can be displayed graphically and may often be interpreted according to chemical, technical and/or biological knowledge[53].

Partial Least Square (PLS)

PLS is a regression extension of PCA which is used to connect the information in two blocks of variables X and Y to each other. It derives its usefulness from its ability to analyze data with many noisy, collinear, and even incomplete variables in both X and Y. As in PCA, each observation can be represented graphically, however, the big difference in PLS is that each row of a data table corresponds to two points rather than one, one in the X-space and one in the Y-space.

PCA is a maximum *variance* least squares projection of X, whereas PLS is a maximum *covariance* model of the relationship between X and Y[53].

1.5 References

[1]Mikulic-Petkovsek M., Slatnar A., Stampar F., Veberic R. HPLC-MS[n] identification and quantification of flavonol glycosides in 28 wild and cultivated berry species. *Food Chemistry* **2012**, 135, 2138-2146

[2]Mattivi F., Vrhovsek U. Small berries with big nutritional benefits. *Mitteilungen Klosterneuburg* **2010**, 60, 442-448

[3]Miller M. G. and Shukitt-Hale B. Berry Fruit Enhances Beneficial Signaling in the Brain.
Journal of Agricultural and Food Chemistry **2012**, 60, 5709−5715

[4]Brown E. M., McDougall G.J., Stewart D., Pereira-Caro G., Gonzales-Barrio R., Allsopp P., Magee P., Crozier A., Rowland I., Gill C. I. R. Persistence of anticancer activity in berry extracts after simulated gastrointestinal digestion and colonic fermentation. *PLoS One* **2012**, 7 (11), e4974

[5]Boat A. S., Grussu D., Stewart D. McDougall G. J. Berry Polyphenols inhibit digestive enzymes: a source of potential health benefits? *Food Digestion* **2012**, 3, 1-7

[6]Seeram N. P. Berry Fruits: compositional elements, biochemical activities, and the impact of their intake on human health, performance, and disease. *Journal of Agricultural and Food Chemistry* **2008**, 56, 627−629

[7]Zhang Y., Seeram N. P., Lee R., Feng L., and Heber D. Isolation and identification of strawberry phenolics with antioxidant and human cancer cell antiproliferative properties. *Journal of Agricultural and Food Chemistry* **2008**, 56, 670−675

[8]Vattem D. A., Maitin V. Bioactivity, Bioavailability, and human health effects of berries' bioactive compounds. *Functional Foods, Nutraceuticals and Natural Products: Concepts and Applications* **2015**, chapter 22, 561-573

[9]Disciplina dell'impiego negli integratori alimentari di sostanze e preparati vegetali. DM 9luglio 2012 (G.U. 21-7-2012 serie generale n. 169)

[10]Puiggròss F., Solà R., Bladè C., Salvadò M-J., Arola L. Nutritional biomarkers and foodomic methodologies for qualitative and quantitative analysis of bioactive ingredients in dietary intervention studies. *Journal of Chromatography A* **2011**, 1218, , 7399-7414

[11]Oliver S. G., Winson M.K., Kell D. B. et al. systematical functional analysis of the yeast genome. *Trends in Biotechnology* **1998**, 16, 373-378

[12]Hall R. D. Plant metabolomics in a nutshell: potential and future challenges. *Annual plant review* **2011**, 43, 1-24

[13]Allwood J. W. and Goodacrea R. An introduction to liquid chromatography–mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochemical analysis* **2010**, 21, 33-47

[14]Villas-Boˆas S. G., Mas S., Akesson M., Smedsgaard J.and J. Nielsen. Mass spectrometry in metabolome analysis. *Mass Spectrometry Reviews* **2005**, 24, 613– 646

[15]Ernst M., Silva D. B., Silva R.R., Vencio R. Z. N., Lopes N. P. Mass spectrometry in plant metabolomics strategies: from analytical platforms to data acquisition and processing *Natural Product Reports* **2014**, 31,784-806

[16]Bino R. J., Hall R. D., Fiehn O., Kopka J., Saito K., Draper J., Nikolau B. J., Mendes P., Roessner-Tunali U., Beale M. H., Trethewey R. N., Lange B. M., Wurtele E. S. and Summer L. W. Potential of metabolomics as a functional genomics tool. *Trends in Plant Science* **2004**, 9, 418

[17]Villas-Boas S. G., Koulman A., Lane G. A. Metabolomics. A powerful tool in system biology, ed. J.Nielsen, M.C. Jewett, Sprinfer-Verlag, Berlin Heidelberg, **2007**, 11-52

[18]Moritz T. Johansson A. I., in Metabolomics, Metabonomics and metabolite profiling, ed. W. J. Griffiths, RSC publishing, Cambrige **2008**, 10, 254-272

[19]Kim H. K. and Verpoorte R. Sample preparation for plant metabolomic. *Phytochemical Analysis* **2010**, 21, 4-13

[20]Urbanczyk E., Baxter C., Kolbe A., Kopka J., Sweetlove L. J. and Fernie A. R. Profiling of diurnal patterns of metabolite and transcript abundance in potato (*Solanum tuberosum*) leaves. *Planta* **2005**, 221,891-903

[21]Verpoorte R., Choi Y.H., Mustafa N. R. and Kim H. K. Metabolomics: back to basic. *Phytochemistry Review* **2008**, 7,525-537

[22]Zhou H. Y. and Liu C. Z. Microwave assisted extraction of solanesol from tobacco leaves. *Journal of Chromatography A* **2006**, 1129, 135-139

[23]Morgenthal K. Wienkoop S., Wolschin F., Weckwert W., in *Methods in Molecular biology*, vol 358: Metabolomics: methods and protocols, ed W. Weckwerth, Totowa, **2007**, 4, 57-75

[24]Wolfender J. L., Rudaz S., Choi Y. H. and Kim H. Plant metabolomics: from holistic data to relevant biomarkers. *Current Medicinal Chemistry* **2013**, 20, 1056-1090

[25]Yuliana N. D., Jahangir M., Verpoorte R. and Choi Y. H. Comprehensive Extraction Method Integrated with NMR Metabolomics: A New Bioactivity Screening Method for Plants, Adenosine A1 Receptor Binding Compounds in *Orthosiphon stamineus* Benth. *Analytical Chemistry* **2011**, 83,6902-6906

[26]Vinatoru M. An overview of the ultrasonically assisted extraction of bioactive principles from herbs. *Ultrasonics Sonochemistry* **2001**,8,303-313

[27]Shulaev V. Metabolomics technology and bioinformatics. *Briefings in Bioinformatics* **2006**, 7, 128-139

[28]Brown S. C., Kruppa G. and Dasseux J.-L. Metabolomics applications of FT-ICR mass spectrometry. *Mass Spectrometry Review* **2005**, 24, 223-231

[29]Herrero M., Simò C., Garcia-Canas V., Ibànez E. and Cifuentes A. Foodomics: MS-based strategies in modern food science and nutrition. *Mass Spectrometry reviews* **2012**, 31, 49-69

[30]Kebarle P., Verkerk U. H. Electrospray and MALDI mass spectrometry ed Cole R. B.,John Wiley &sons 2nd edn, **2010**, 1,3-48

[31]Karas M., Bahr U and Giessmann. Matrix-assisted laser desorption ionization mass spectrometry. *Mass Spectrometry Review* **1991**,10, 335-57

[32]Beavis R. C. Matrix-assisted ultraviolet laser desorption: evolution and principles. *Organic Mass spectrometry* **1992**, 27,653

[33]McLuckey S. A., Principles of collisional activation in analytical mass spectrometry. *Journal of the American Society for Mass Spectrometry* **1992**, 3, 559-614

[34]Baker M. and Garbyelski W. Int. Collision induced dissociation of deprotonated glycolic acid. *Journal of Mass spectrometry* **2007**, 262,128-135

[35]Hagel J. M. and Facchini P. J. Plant metabolomics: analytical platforms and integration with functional genomics. *Phytochemical Review* **2008**, 7, 479

[36]Kopka J., Fernie A., Weckwerth W., Gibon Y. and Sitt M. Metabolite profiling in plant biology: platforms and destinations. *Genome Biology* **2004**, 5,109

[37]Allwood J. W., Erban A., Koning S., Dunn S. K. B. Luedeman, Lommen A., kay L., Loscher R., Kopka J. and Goodacre R. Inter-laboratory reproducibility of fast gas chromatography-electron impact-time of flight mass spectrometry (GC-EI-TOF/MS) based plant metabolomics. *Metabolomics* **2009**, 5,479-496

[38]Bicchi C., Brunelli C., Cordero C., Rubiolo P. Galli M. and Sironi A. Direct resistively heated column gas chromatography (Ultrafast module-GC) for high-speed analysis of essential oils of differing complexities. *Journal of Chromatography A* **2004**, 1024,195-207

[39]Kuehnbaum N. L. and Britz-McKibbin P. New advances in separation science for metabolomics: resolving chemical diversity in a post-genomic era. *Chemical Review* **2013**, 113,2437-68

[40]Wang X., Sun H. , Zhang A., Wang P., Han Y. Ultra-performance liquid chromatography coupled to mass spectrometry as a sensitive and powerful technology for metabolomic studies. *Journal Separation Science* **2011**, 34,3451

[41]Spagou K., Tsoukali H., Raikos N., Gika H., Wilson I. D. and Theodoris G. Hydrophilic interaction chromatography coupled to MS for metabonomic /metabolomic studies. *Journal Separation Science* **2010**,33, 716

[42]Ibanez C., Garcia-Canas V., Valdes A., Simò C. Novel MS-based approaches and applications in food metabolomics. *Trends in Analytical chemistry* **2013**, 52, 100-111

[43]Takats Z., Wiseman A. Gologan B., Cooks R. G. Mass spectrometry sampling under ambient condtions wit desorption electrospray ionization. *Science* **2004**, 306, 471-473

[44]Chen H., Pan Z., Talaty N., Raftery D., Cooks R. G. Combining desorption electrospray ionization mass spectrometry and nuclear magnetic resonance for differencial metabolomics without sample preparation. *Rapid Communication in Mass Spectrometry* **2006**, 20,1577-1584

[45]Cody R. B. Laramee J. A., Durst H. D. Versatile new ion source for the analysis of materials in open air under ambient conditions. *Analytical Chemistry* **2005**, 77, 2297-2302

[46]Zhou M., McDonald J. F., Fernandez F. M. Optimization of a direct analysis in real time/time of flight mass spectrometry method for rapid serum metabolomics fingerprinting. *Journal of the American Society for Mass Spectrometry* **2010**, 21,68-75

[47]Deimler R. E., Razunguzwa T. T., Reschke B. R., Walsh C. M., Powell M. J. and Jackson G. P. Direct analysis of drugs in forensic applications using laser ablation electrospray ionization-tandem mass spectrometry (LAESI-MS/MS). *Analytical Methods* **2014**, 6, 4810-4817

[48]Vaikkinen A., Shrestha B., Koivisto J., R. Kostiainen, A. Vertes, T. J. Kauppila. Laser ablation atmospheric pressure photoionization mass spectrometry imaging of phytochemicals from sage leaves. *Rapid Communication in Mass Spectrometry* **2014**, 28, 2490-2496

[49]Liu J., Wang H., Manicke N. E., Lin J.-M., Cooks R. G. and Z. Ouyang. Development, Characterization, and Application of Paper Spray Ionization. *Analytical Chemistry* **2010**, 82, 2463-2471

[50]Trygg J., Holmes E. and Lundstedt T. Chemometrics in Metabonomics. *Journal of Proteome Research* **2007**, 6, 469-479

[51]Castillo S., Gopalacharylu P., Yetukuri L. ans Oresi&ccaron M. Algorithms and tools for the preprocessing of LC-MS metabolomics data. *Chemometrics and Intelligent Laboratory Systems* **2011**,108, 23

[52]Chambers M.C., Maclen B., Burke R., Amodei D., Rudermann D. L., et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology* **2012**,30,918.

[53]Eriksson L., Byrne T., Johansson E., Trygg J. and Vikström C. Multi and Megavariate Data Analysis, Basic Principles and Apllications 3th revised edition **2013** Umetrics

Chapter 2

Metabolomics approaches used for discrimination of different cultural system

Metabolomics represent a useful tool to evaluate the contribution of environmental factors to the differences in metabolite composition or content. Modern and sensitive analytical techniques coupled with Multivariate Data Analysis were developed for the discrimination of different cultural system, in particular the study was focused on organic and conventional strawberry crops.

Strawberries are one of the most widely consumed berries and are considered to be a functional food, with multiple health benefits, over and beyond nutritional needs, as demonstrated by extensive evidence regarding their antioxidant, anti-inflammatory, antihypertensive and antiproliferative properties[1].

Organic products are becoming increasingly popular and previous research has reported that organic fruits and vegetables have higher levels of flavonoids and ascorbic acid.[2]

Moreover, a higher antiproliferative activity towards cancer cells was found in extracts from organically grown strawberries than conventionally grown[3].

During this study, metabolomic fingerprint approach was carried out on methanolic extracts of strawberries fruit to discriminate conventional and organic crops with the objective of apply these approaches for quality control, for authenticity assessment and for adulteration evaluation of food products.

2.1 Integrated Mass Spectrometric and Multivariate Data Analysis approach for the discrimination of organic and conventional strawberry (*Fragaria ananassa* Duch.) crops

2.1 Introduction

The commonly known strawberry fruit [*Fragaria ananassa* Duch. (family Rosaceae)] is constituted by two botanical organs: the receptacle, which results from the enlargement of the flower receptacle upon pollination, and the achenes, the true botanical fruit, are attached to the surface of the receptacle through vascular bundles[4]. Strawberry is one of the most commonly consumed berry fruit. Together with other soft fruit, it is an important dietary source of fiber and bioactive compounds, both micronutrients and phytochemicals. In particular, strawberry fruits are a very rich source of phenolic compounds, which are well known for their antioxidant activities.

The main phenolic classes occurring in strawberries are anthocyanins, flavonoids, hydroxycinnamic acid derivatives, proanthocyanidins, ellagitannins, and ellagic acid derivates [5]. Phenolic compounds possess an aromatic ring bearing one or more hydroxyl groups and their structures may range from that of a simple phenolic molecule to that of a complex high-

molecular weight polymer. The antioxidant activity of phenolic compounds depends on the structure, in particular the number and positions of the hydroxyl groups and the nature of substitutions on the aromatic rings[6]. In plants, phenols are constituents of plant pigments and are involved in plant growth and reproduction as well in the resistance of plants to pests and pathogens, and berry fruits are a rich source of phenolic compounds, particularly anthocyanins[7]. Due to these bioactive compounds, strawberry fruits are reported to have antioxidant, anticancer, anti-inflammatory and anti-neurodegenerative biological properties[8]. Emerging research reports that consumption of berry fruits like strawberry and blueberry, has direct effects on the brain. Specifically, the ingestion of berries may help to prevent age-related neurodegeneration and resulting changes in cognitive and motor function, in fact berry fruits mediate signaling pathways involved in inflammation and cell survival in addition to enhancing neuroplasticity, neurotransmission, and calcium buffering, all of which lead to attenuation of age- and pathology-related deficits in behavior[9].

Many methods so far used to identify strawberry fruit phenolics have been optimized for particular groups of compounds such as anthocyanins[6] or ellagitanin-based compounds[10,11]. In addition many methods previously reported to identify strawberry fruit phenolics were based on spectrophotometry[12] or high performance liquid chromatography coupled with UV detection (HPLC-UV) or diode array detection (HPLC-DAD)[13,14]. The use of HPLC coupled with mass spectrometry (HPLC-MS) detection provides useful structural information and allows the identification of tentative compounds when standard reference compounds are unavailable and when peaks have similar retention time (TR) and similar UV-absorption spectra. In addition, tandem mass spectrometric (MS^n) techniques are useful for

distinguishing compounds with identical molecular weights[10, 9, 15-17]. Recently, Pavlovic et al. (2013)[18] reported a UHPLC method coupled with a hybrid mass spectrometer, which combines a linear trap quadrupole (LTQ) and an Orbitrap mass analyzer, to investigate the main markers specific to each berry species.

Numerous methods, both chemical and physical, may be used in laboratories for food research and control, mainly to evaluate the quality of the product, authenticity/adulteration, and traceability in the production and marketing chain. Besides demand for quality parameters, a very important issue is traceability of some chemical markers of food related to its origin and nutritional quality[18]. Metabolomic profiling techniques are highly relevant since they can reveal a comprehensive view on the relative levels of hundreds to thousands of metabolites present in the plant material[19].

With regard to the metabolomic profiling aims, a full scan high resolution mass analysis appears to be most designated for such approaches since there is no need for compound tuning; therefore, the bioanalytical setup is simplified and information on compounds of interest, as well as untargeted components, is readily obtained[20-24]. Full scan mass analysis offers indeed the possibility to simultaneously analyze virtually unlimited number of compounds. Furthermore, the retrospective post-acquisition evaluation of data allows screening for analytes that were not selected a priori[25]. Metabolomic approaches have increasingly been used to gain insight into the metabolic composition of plant organs and to characterize the natural variance in metabolite content[26]. Metabolomics can provide a diagnostic tool for better understanding of a biological system and has now been successfully performed on a diverse array of plant species, including models such as Arabidopsis[27], tobacco[28] and potato[29]. In contrast to trascriptomics and proteomics, which rely to a great extent on genome information, metabolomics

is mainly metabolite dependent. Integrated metabolic profiling techniques are typically directed towards providing a wide-ranging estimation of the total metabolomes, including intermediate and end-product metabolites from multiple pathways. Wide-ranging metabolic profiling typically yields large data sets which can be mined for information using an array of statistical modeling techniques. The objective of the work was to verify if modern and sensitive analytical techniques are able to discriminate strawberry fruits grown under conventional and organic crops. In addition, with respect to the similar results obtained for the two different mass spectrometric approaches, the present work underlines as a fingerprint MS analysis could be used to classify two types of products obtained by different cultivations. By working on a selected number of samples used as models, it was possible to explore the application of Mass Spectrometry techniques coupled with Multivariate Data Analysis approaches to metabolomic studies on *F. ananassa* fruits.

2.2. Results and discussion

DI-ESI-MS untargeted analysis of strawberry crude extracts

Extracts from unripe (white) and ripe (red) strawberry fruits were analyzed by direct introduction mass spectrometry. Direct infusion electrospray ionization mass spectrometry (DI-ESI-MS) in the negative ion mode and in the positive ion mode were initially used to obtain fingerprints of the extracts of different fruit samples. Twenty extracts (samples) were used, obtained from: ripe (red) fruits (TR1, TR2, TR3, TR4, TR5) and unripe (white) fruits (TB1, TB2, TB3, TB4, TB5) from the conventional farm and ripe (red) fruits (BR1, BR2, BR3, BR4, BR5) and unripe (white) fruits (BB1, BB2, BB3, BB4, BB5) from the organic farm. The full spectra of each sample were recorded in triplicate with the aim to rapidly provide visual and statistical evaluations of similarities and

differences of secondary metabolites among them. The ESI-MS fingerprints of samples showed distinctive sets of markers for each sample, mainly flavonoids in the negative ion mode, and anthocyanins in the positive ion mode. ESI-MS fingerprint obtained in the negative ion mode $[M-H]^-$ evidenced the presence of ion peaks at *m/z* values of 289 (catechin) 463 (quercetin glucoside), 477 (quercetin glucuronide) 461 (kaempferol glucoside) and 577 and 865 corresponding to dimeric and trimeric procyanidins. Other peaks were present as major peaks in the ESI-MS positive fingerprint.

ESI-MS fingerprint obtained in the positive ion mode $[M+H]^+$ evidenced the presence of ion peaks at *m/z* values of 433 (cyanidin glucoside), 579 (pelargonidin glucoside) 461 (pelargonidin rutinoside), together with other peaks characterizing the ESI-MS fingerprint in the positive ion mode. At this point, considering the large amount of data set obtained combining the positive and negative ion ESI-MS fingerprints of extracts under investigation, a chemometric approach as PCA was performed to differentiate the different fruits and to evaluate differences in terms of metabolites. The obtained data, in the positive and negative ion modes, were organized in a data matrix, using the *m/z* values of observed peaks showing different intensities in the spectra as variables, and the samples under investigation as observations. Data matrix was processed with Multivariate Data Analysis (MVDA): Principal Component Analysis (PCA), using SIMCA+ Software, was applied to the matrix. PCA is an unsupervised method and was used to reduce the dataset in order to obtain the maximum variation between the samples. PCA has been used widely in assessing the differences between plant varieties at metabolomic level[30]. Pareto scaling was chosen for scaling data. Fig.2.1 A shows the 2D projection plot of the samples. The first two principal components, having a values greater then unity, were selected. The first

component (R1X) explains the 70% of variance and the second (R2X) explains the 20%. Principal Component's choice was done on the basis of the fitting (R2X) and predictive (Q2X) values for the PCA model; in our case the second component gave the closest value to 1 for both of them. Variance was evaluated by significance level for Hotelling's T2. Fig.2.1 shows the PCA resulting to the total data analysis. PCA results showed a clear discrimination of unripe and ripe fruits, and in fact we can observe separated regions relative to these observations. Conventional and organic crop treatment is not directly separated. To better observe this separation, unripe and ripe fruits were considered separately by realizing two different matrices to submit to PCA analysis. Fig. 2.1 B shows the results of this second differentiated statistical evaluation. It is possible to observe that unripe samples are discriminated on the basis of the cultivation technique, while for ripe fruits there is a separation only on a diagonal axis. Based on these observations, a classification approach was applied to the samples.
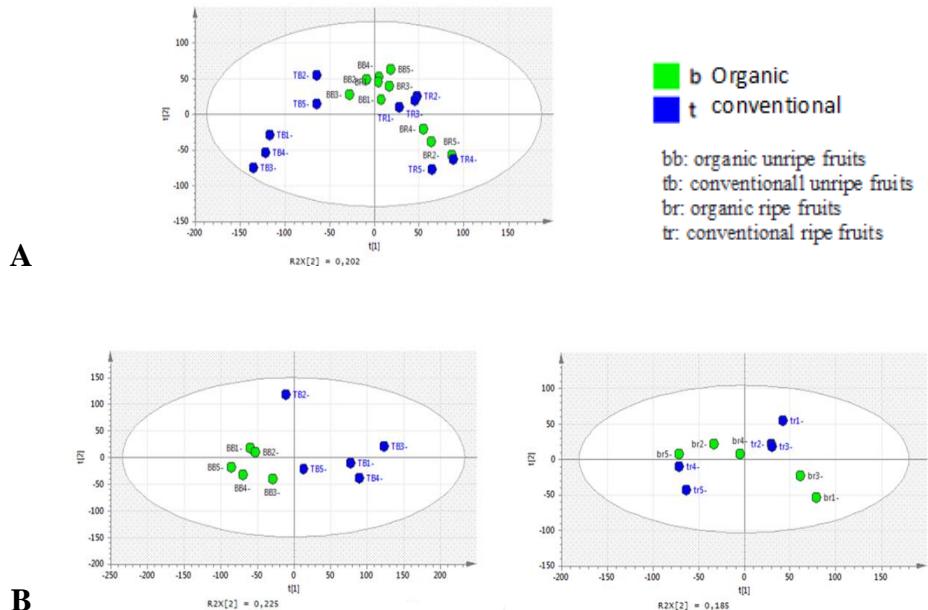
Fig. 2.1: Principal Component Analysis (PCA) score plot of unripe and ripe fruits from organic and conventional strawberry crops. **A**: PCA scatter plot total samples; **B**: PCA score scatter plot unripe and ripe fruits

The matrices were treated by a PLS-DA approach. PLS Discriminant Analysis (PLS-DA) is performed in order to sharpen the separation between groups of observations, by hopefully rotating PCA (Principal Components Analysis) components such that a maximum separation among classes is obtained, and to understand which variables carry the class separating information. PLS-DA consists in a classical PLS regression where the response variable is a categorical one (replaced by the set of dummy variables describing the categories) expressing the class membership of the statistical units.

The data sets were first analyzed by Principal Component Analysis, and specifically the data sets were obtained separately for the unripe and ripe fruits, defined homogeneous cluster of samples. These clusters were then used

31

as Y classes in Partial Least Squares Discriminant Analysis (PLS-DA). The data were thus modeled by Partial Least Squares Discriminant Analysis (PLS-DA) as a supervised approach in exploring clustering relationship. Fig. 2.2 shows the Score Scatter Plots obtained by matrix treatment with PLS-DA approach for unripe and ripe fruits. All the differences between fruits collected from the organic crops or conventional crops were underlined, both in the unripe fruit group and in the ripe fruit group. Fruit ripening is reflected by physiological changes associated with changes in gene expression, protein synthesis and metabolism; thus the ripe and unripe fruits were treated as different sets of data. However, changes in metabolism, generated from conventional organic cultivation, revealed by fingerprint MS analysis, gave classification of samples both in ripe and unripe fruits.
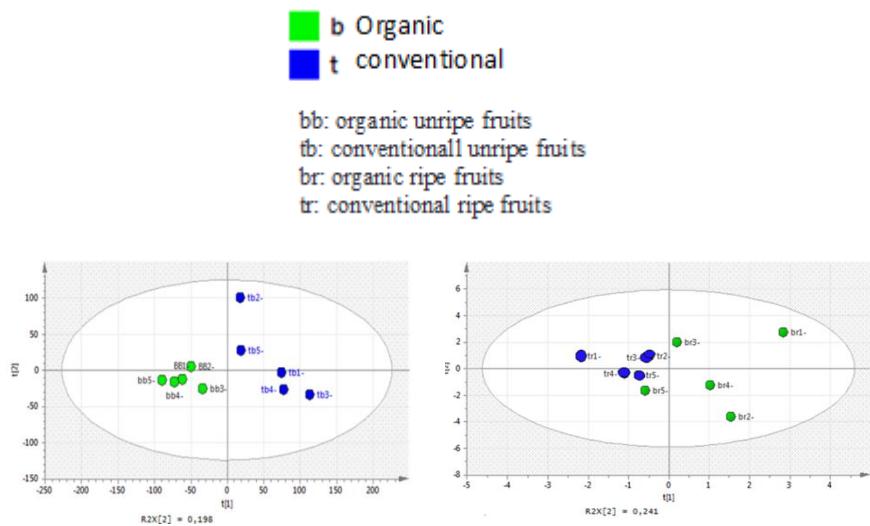
Fig. 2.2: Partial least square (PLS) score scatter plot unripe and ripe fruits.

LC–ESI-MS and LC–ESI-MS/MS analyses of crude extracts

In order to obtain a metabolite profile of the crude extract from fruits, an analytical method based on LC-ESI-Orbitrap-MS and LC-ESI-Orbitrap-MS/MS was developed, both in negative and in the positive ion modes. The negative LC-MS profile highlighted the presence of a large group of compounds corresponding to the deprotonated molecular ions of different flavonoids. The positive LC-MS profile highlighted the presence of anthocyanins (Fig.2.3).

A
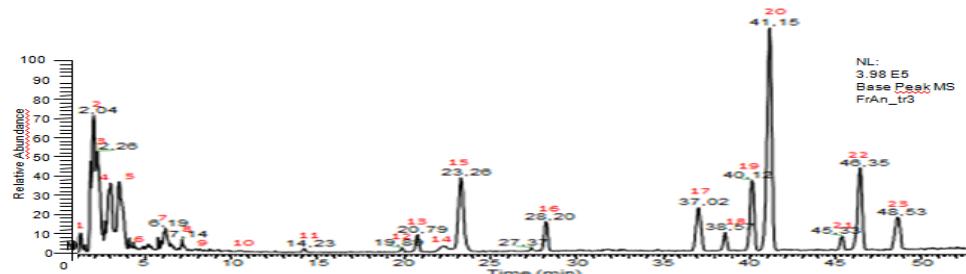
B
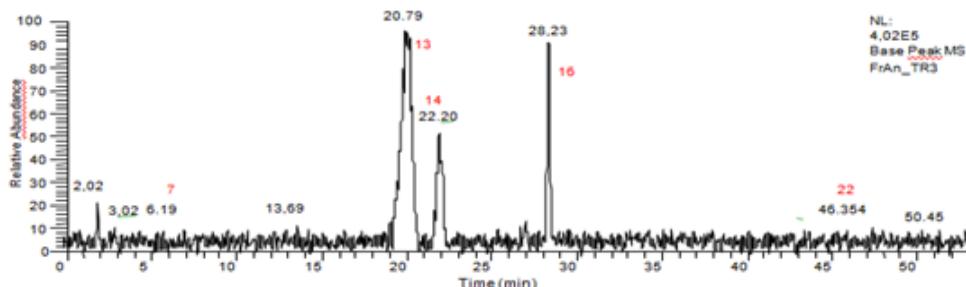
Fig. 2.3: LC-ESI-Orbitrap-MS profiles of methanolic extract of strawberry fruits (sample Tr3). Panel A. Negative ion mode, Panel B. Positive ion mode.

Individual components were identified by comparison of their *m/z* values in the Total Ion Current (TIC) profile with those of the selected compounds described in literature (Table 2.1).

Table 2.1: Retention time, high resolution pseudomolecular ions of compounds occurring in strawberry extracts by LC-ESI-Orbitrap-MS.

| Name	Retention time (min)	Ion (+/-)	[M-H]⁻ / [M+H]⁺	References
1. Catechin	1.28	−	289.0718	31,32
2. Quercetin	2.04	−	301.0349	30
3. Kaempferol	2.26	−	285.0399	18
4. Ellagic acid	2.93	−	300.9985	31
5. Unknow*	3.01	−	383.0458	−
6. Unknow*	4.86	−	411.0766	−
7. Pelargonidin-glucoside	6.19	+	433.1129	31
8. Apigenin	7.14	−	270.0529	18
9. Unknow*	8.12	−	387.1136	−
10. Unknow*	10.10	−	369.1031	−
11. Myricetin-3-O-glucoside	14.23	−	479.0826	33
12. Kaempferol-glucuronide	19.89	−	461.0720	31
13. Unknow*	20.79	−	325.0924	−
14. Cyanidin-glucoside	22.20	+	449.1078	31
15. Chlorogenic acid	23.26	−	353.0873	18
16. Pelargonidin-diglucoside	28.20	+	595.1658	31
17. Quercetin-rutinoside	37.02	−	609.1456	31
18. Kaempferol-coumaroil-glucoside	38.57	−	593.1295	31
19. Quercetin-glucuronide	40.12	−	477.0669	31
20. Proanthocyanidin B2	41.15	−	577.1346	34
21. Quercetin-glucoside	45.33	−	463.0877	31
22. Galloyl-HHDP-glucoside	46.35	+	937.6532	31
23. Unknow*	48.53	−	489.1013	−
24. Methyl-Ellagic acid-pentose	49.03	−	447.0921	31
25. Pelargonidin-rutinoside	50.4	+	579.1708	31

* Compounds not identified neither by literature or standard compounds.

Additional LC-ESI-Orbitrap-MS/MS experiments were carried out in order to select and submit these ions to fragmentation experiments using the

parameters previously chosen by ESI-MS direct infusion experiments. By matching experimental MS/MS spectra with those reported in a public repository of mass spectral data called Mass Bank [35] compounds were identified (Table 2.2).

Table 2.2: Formula, exact mass and exact masses fragmentation of major compounds in strawberry extract analysed by LC-ESI-Orbitrap-MS/MS.

	Formula	$[\text{M-H}]^-$	$[\text{M}]^+$	MS/MS	Identification
7^b	$C_{21}H_{21}O_{10}$		433.1129	271.0606 ($-C_{15}H_{11}O_5$)	Pelargonidin-glucoside
12^a	$C_{21}H_{18}O_{12}$	461.0720		415.1605 ($-C_{20}H_{27}O_{10}$) 283.8663 ($-C_{15}H_8O_6$)	Kaempferol glucuronide
14^b	$C_{21}H_{21}O_{21}$		449.1078	287.0557 ($-C_{15}H_{11}O_6$)	Cyanidin-glucoside
17^a	$C_{27}H_{30}O_{16}$	609.1496		301.0348 ($-C_{15}H_9O_7$)	Quercetin-rutinoside
18^a	$C_{30}H_{26}O_{13}$	593.1295		327.0507 ($-C_{17}H_{11}O_7$) 285.040 ($-C_{15}H_9O_6$)	Kaempferol-coumaroil glucoside
19^a	$C_{21}H_{18}O_{13}$	477.0669		301.0351 ($-C_{15}H_9O_7$)	Quercetin-glucuronide
20^a	$C_{30}H_{26}O_{12}$	577.1346		289.0459 ($-C_{15}H_{13}O_6$)	Proanthocyanidin B2
21^a	$C_{21}H_{20}O_{12}$	463.0877		301.0350 ($-C_{15}H_9O_7$)	Quercetin-glucoside
25^b	$C_{27}H_{31}O_{14}$		579.1708	433.1139 ($-C_{21}H_{21}O_{10}$) 271.0607 ($-C_{15}H_{11}O_5$)	Pelargonidin-rutinoside

a: negative ion mode
b: positive ion mode

Compounds **1**, **2**, **3**, **4**, **8**, **11**, **14**, **15**, **16**, **22** and **24** were identified by the diagnostic $[\text{M-H}]^-$ or $[\text{M+H}]^+$ ions showed in High Resolution (HR) ESI-MS

analysis, compared with standards and literature. Compounds **5**, **6**, **9**, **10**, **13** and **23** were not identified and are reported in the table as unknown.

The identities of compounds **12**, **17**, **18**, **19**, **20** and **21** were confirmed from the MS/MS data obtained by working in LC-ESI-Orbitrap-MS/MS in Product Ion Scan in the negative ion mode, and by comparing retention times and mass spectra of these compounds to those of commercial standards. Moreover, the identities of compounds **7**, **14** and **15** were confirmed from the MS/MS data obtained by working in LC-ESI-Orbitrap-MS/MS in Product Ion Scan in the positive ion mode, and by comparing retention times and mass spectra of these compounds to those of commercial standards. Full negative HR-ESI-MS profile of compound **12** was in agreement with kaempferol glucuronide, showing the diagnostic [M-H]⁻ molecular ion at 461.0721 *m/z*. The analysis of the ESI-MS/MS spectrum of compound **12** allowed one to determine the presence of a carboxihexose unit, with a prevalent fragment at 283.8663 uma. In literature the presence of kaempferol glucuronide for this species is reported, thus the compound was identified. Full negative HR-ESI-MS profile of compound **18** allowed the identification of kaempferol-coumaroyl-glucoside, showing the diagnostic [M-H]⁻ molecular ion at 593.1295 *m/z*. The analysis of the ESI-MS/MS spectrum of compound **18** allowed to determine the presence of a hexose unit derivatized with a coumaroyl esterification, with prevalent fragments at 327.0507 and 285.0401 uma. In literature the presence of kaempferol-coumaroyl-glucoside for this species is reported, thus the compound was identified. The HR-ESI-MS profile of compound **19** was in agreement with quercetin glucuronide and, showing the diagnostic [M−H]⁻ molecular ion at 477.0669 *m/z*. The analysis of the ESI-MS/MS spectrum of compound **19** allowed one to determine the presence of a carboxihexose unit, with prevalent fragment at 301.0351.

In literature the presence of quercetin glucuronide for this species is reported; thus the compound was identified. Full negative HR-ESI-MS profile of compound **20** assigned it to proanthocyanidin B2, showing the diagnostic [M-H]$^-$ molecular ion at 577.1346 *m/z*. The analysis of the ESI-MS/MS spectrum of compound **20** allowed one to identify a catechin dimer, with the major fragment ion at 289.0459. In literature the presence of proanthocyanidin B2 for this species is reported, thus the compound was identified.

Negative HR-ESI-MS profile of compound **21** was in agreement with quercetin glucoside and showing the diagnostic [M-H]$^-$ molecular ion at 463.0877 *m/z*. The analysis of the ESI-MS/MS spectrum of compound **21** allowed one to determine the presence of a hexose unit, with prevalent fragment at 301.0351. In literature the presence of quercetin glucoside for this species is reported, thus the compound was identified. The HR-ESI-MS of compound **7** suggested again an anthocyanin glucoside, showing a precursor ion at *m/z* 433.1129, corresponding to the [M]$^+$ pseudomolecular ion of pelargonidin glucoside, previously found in the fruit of this species, and a major fragment ion at *m/z* 271.0606 corresponding to the loss of an hexose unit. Full positive HR-ESI-MS of compound **14** suggested that it was a glycoside of an anthocyanin, showing the major precursor ion at *m/z* 449.1078, corresponding to the [M]$^+$ pseudomolecular ion of cyanidin glucoside and a major product ion at *m/z* 287.0557 corresponding to the loss of a hexose unit. The HR-ESI-MS of compound **25** suggested again an anthocyanin glucoside, showing the major pseudomolecular ion at *m/z* 579.1708, corresponding to the [M]$^+$ pseudomolecular ion of pelargonidin rutinoside, previously found in the fruit of this species, and two major fragment ions at 433.1139 *m/z* 271.0607 corresponding to the sequential loses of two hexose units. All these compounds could be considered chemical markers of the species [14].

A new set of experiments was executed in LC-ESI-Orbitrap-MS in the negative ion mode working on 12 samples, with the aim of comparative qualitative analysis between traditional and organic ripe and unripe fruits. These analyses were run on a reduced number of samples based on the more time-consuming protocol. Samples included in the set were: Three samples of ripe fruits TR1, TR2, and TR3 and three samples of unripe fruits TB1, TB2, and TB3 from the conventional farm and three samples of ripe fruits BR1, BR2, and BR3 and three samples of unripe fruits BB1, BB2, and BB3 from the organic farm.

The negative ion mode at this stage was selected because of the major number of compounds detected in this operative mode.

A different approach was used for untargeted multivariate data analysis. For analysis of the acquired dataset with multivariate methods, LC-MS chromatograms were pre-processed using MZmine to compensate for 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, and columns representing the integrated and normalized peak areas. Moreover these data were used to confirm the MVDA (PCA and PLS-DA), through an approach of untargeted analysis. All the differences between fruits collected from the organic crops or conventional crops were underlined. Principal Component Analysis was performed by applying the peak areas of the total peaks present in the LC-MS dataset (excluding the noisy), thus a matrix was obtained by using these areas (variables), and the columns of the matrix were different analyzed samples. The data were also modeled by Partial Least Squares Discriminant Analysis (PLS-DA) as a supervised approach to exploring clustering relationship (Fig. 2.4).

38
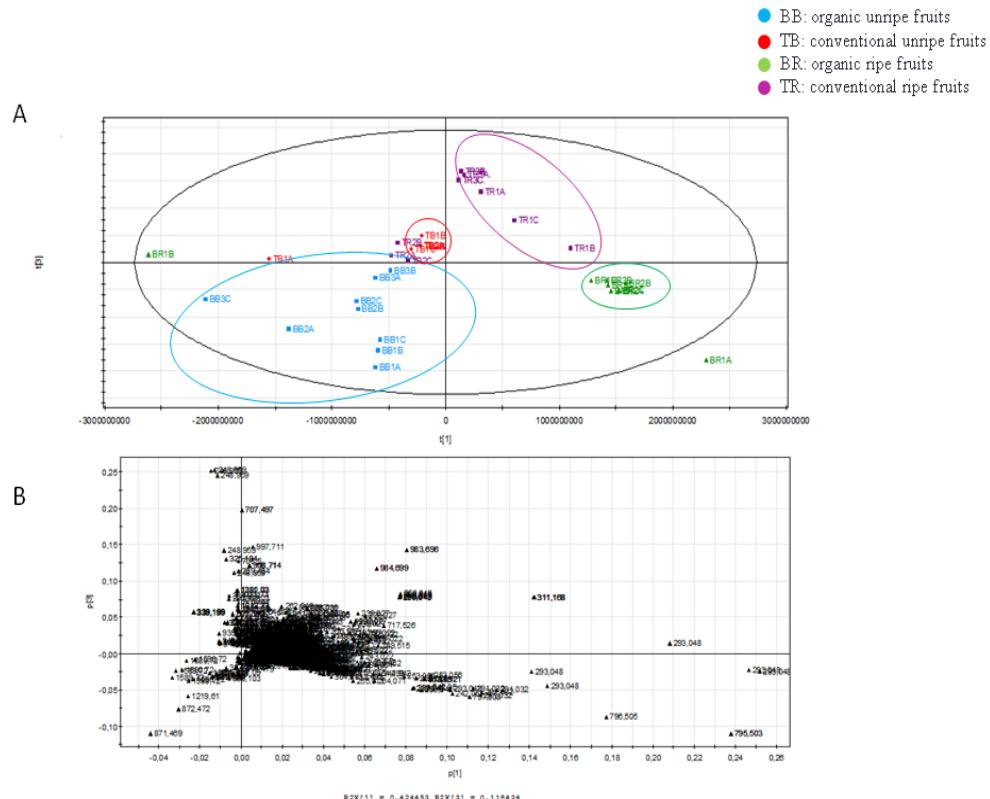
Fig.2.4: PLS-DA analysis on LC-MS data of unripe and ripe fruits of *Fragaria ananassa*, A) score plot B) loading plot

Moreover these data were used to confirm the MVDA (PCA and PLS-DA), through an approach of untargeted analysis. The resulted score scatter plot is reported in Fig. 2.4. Results show a good discrimination between the four classes of samples, with a profile very similar to what was observed in the fingerprint multivariate approach, demonstrating the two multivariate data analysis approaches led to the same results because probably the fingerprint analysis combining the positive and negative ion modes results were representative of the variability of the complex datasets obtained for the

samples under investigation. 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 showed in Fig. 2.4B: Interestingly, the loading plot shows the *m/z* values corresponding each region of the 2D space to the peak observed in the specific samples. In particular variables that contribute most to the differentiation of the samples and to their location in a specific area of the space can be highlighted. No compounds pertaining to the classes of flavonoids and anthocyanins are responsive for the classification of samples.

2.3 Conclusions

Organic farming is a cultivation approach that refuses the use of synthetic fertilizers, pesticides and herbicides in food production and aims for the maintenance of soil fertility and general environmental sustainability of the agriculture process. Following the growing concern in the public opinion associated with the massive use of chemicals in agriculture, organic farming has undergone a massive growth in the developed countries in the last years and, to date, two main approaches to the agriculture practice can be observed, one being the organic farming itself and another usually named as conventional approach, i.e., nonorganic, that makes use of synthetic fertilizers, pesticides and herbicides in addition, or instead of, organic fertilizers and biological control means. From the point of view of the consumer, discrimination between the two cultivation approaches is currently made possible through the traceability of the food product. The aims of this work were to develop and compare modern and sensitive analytical techniques coupled with multivariate data analysis and be able to discriminate strawberry fruits grown under conventional and organic crops.

For the first time high resolution LC-ESI-Orbitrap-MS and LC-ESI-Orbitrap MS/MS were applied to the evaluation of the phytochemical metabolic profile of strawberry fruits. Subsequently, metabolomic fingerprint approach was applied to the fruit's methanolic extracts to discriminate conventional and organic crops and ripe (red) and unripe fruits. The combination of DI-ESI–MS analysis with MVDA, and mainly PCA and PLS-DA led us to achieve a good discrimination among the fruits sampled from the conventional crops and the fruits sampled from the organic crops. The differences were most underlined in the unripe fruits, for which a PCA analysis was able to discriminate samples from different types of cultivation. This may suggest that fruits in an early maturation stage could be used as marker samples to discriminate conventional versus organic crops. With respect to the fact that the commercially significant products are ripe fruits, PLS-DA could be used to discriminate the crop technique by using ripe fruits. The same classification results were obtained by an untargeted High Resolution LC-ESI Orbitrap MS metabolomic approach, combining data pre-treatment by MzMine and PCA and PLS-DA chemometric analyses. In conclusion, analysis by DIESI-MS or LC–ESI-MS could be used for fast quality control screening, after a validation of the method here was proposed on a larger number of commercial products, both from organic and conventional farms. In this way, the approach here reported based on DI-ESI-MS technique confirmed by LC-ESI-MS techniques followed by PCA and PLS-DA analysis can find useful applications in the field of qualitative and quantitative analyses of berry fruits and to evaluate the environmental conditions sensu lato used for cultivation.

2.4 Materials and methods

Materials

Plant material

Ripe (fully expanded, completely red) and unripe (fully expanded, completely white) fruit of *F. ananassa* var. Candonga were collected from two different farms both located in the surroundings of Paestum [Sele River Plain, Campania Region, South Italy], distant as far as 2 km from each other here are named "T farm" and "B farm". The T farm grew strawberry plants following a conventional technique, i.e. nonorganic, using synthetic fertilizers applied through watering and following an integrated pest and pathogen management protocol, with early treatments with synthetic pesticides against powdery mildews, late treatments with sulfur against powdery mildews and mites, Orius sp. and Phytoseiulus persimilis application against thrips and mites. The B farm grew strawberry plants following an organic farming approach, using neither synthetic fertilizers nor synthetic pesticides and applying Orius sp. and Aphidius sp. against thrips, mites and aphids. Fruit samples were harvested during June 2013 at different ripening stages. Ripe and unripe fruits were collected from five distinct rows both from conventional crops (T) and organic crops (B), in order to obtain five biological replicates per each type of fruit and farm. Before analysis, the samples were stored at -18 °C, to preserve the content of phenolic compounds and radical-scavenging activity of berries[36,37].

Chemicals

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

Rutin and cyanidin-3-*O*-glucoside were purchased from Extrasynthese (Genay Cedex, France).

Methods

Sample preparation

For DI-ESI-IT-MS analyses twenty samples were used: five ripe (red) samples of fruits (ten fruits for each sample) (TR1, TR2, TR3, TR4, TR5) and five unripe (white) samples of fruits (ten fruits for each sample) (TB1, TB2, TB3, TB4, TB5) from a conventional farm; five ripe (red) samples of fruits (ten fruits for each sample) (BR1, BR2, BR3, BR4, BR5) and five unripe (white) samples of fruits (ten fruits for each sample) (BB1, BB2, BB3, BB4, BB5) from an organic farm. All the samples were extracted for 24 h with acidified methanol[31] and the filtered extract (filtered by using Filter devices 0.2 μm PTFE membrane) was used for direct introduction. For each sample the analysis were recorded for 1 min in negative and positive ion modes. For a more time-consuming LC-MS analysis, a reduced number of samples was used (12): Three samples of ripe fruits TR1, TR2, and TR3 and three samples of unripe fruits TB1, TB2, and TB3 from a conventional farm and three samples of ripe fruits BR1, BR2, and BR3 and three samples of unripe fruits BB1, BB2, and BB3 from an organic farm. Samples were extracted for 24 h with acidified methanol and then 20 μL of filtered extract (filtered by using Filter devices 0.2 μm PTFE membrane) were dried under N_2 and diluted in 1 mL of water. This solution was used in LC-MS where 10 μL was injected. For LC-MS/MS analysis only one sample was used (TR3) and prepared in the same condition described for LC-MS analysis.

DI-ESI-IT-MS

Using a metabolomic approach, i.e. no targeted ion monitoring, direct infusion MS (DIMS) was carried out on LCQ-Deca (Thermo Finningan) mass spectrometer equipped with an ESI source and an Ion Trap analyzer. Spectra were acquired in the negative and positive ion modes in the range *m/z* 200-2000. The analytical parameters were optimized by a standard solution of rutin (1 μg ml−1 in 50% methanol) for the negative ion mode, and a standard solution of cyanidin-3-*O*-glucoside (1 μg ml−1 in 50% methanol) for the positive ion mode, infused into the ion source at a flow rate of 10 μL/min. The optimized parameters were: Negative ion mode: capillary temperature: 280 °C, capillary voltage: −46 V, Spray voltage: 5 kV, Tube lens: 10 V, Auxiliary gas: 5 arbitrary units. Positive ion mode: capillary temperature: 280 °C, capillary voltage: +46 V, Spray voltage: 5 kV, Tube lens: 10 V, Auxiliary gas: 5 arbitrary units.

LC-ESI-Orbitrap-MS analysis

To investigate the main markers specific to each fruits, a HPLC method coupled with a hybrid mass spectrometer, which combines the linear trap quadrupole (LTQ) and Orbitrap mass analyzer, was developed in this research. Methanol extracts were used for this analysis. All experiments were performed using a Thermo Scientific liquid chromatography system constituted of a quaternary Accela 600 pump and an Accela autosampler, connected to a linear Trap-Orbitrap hybrid mass spectrometer (LTQ-Orbitrap XL, Thermo Fisher Scientific, Bremen, Germany) with electrospray ionization (ESI). Separation was performed on a Symmetry-Shield RP 18 column (Waters). The mobile phase consisted of A (water + 0.1% formic acid) and B (acetonitrile/water (8:2) + 0.1% 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 to 56 min, 100% (B); then 0% (B) for 5 min. The mass spectrometer was operated in the negative ion mode. ESI source parameters were as follows: capillary voltage -35 V; tube lens voltage -126.47; capillary temperature 280 °C; Shealt and Auxiliary Gas flow (N_2), 20 5 Sweep gas 0 Spray voltage 5. MS spectra were acquired by full range acquisition covering *m/z* 200-1600. For fragmentation study, a product ion scan was performed, selecting precursor ions corresponding to the most intensive peaks in LC-MS analysis. Phenolic compounds were identified according to the corresponding spectral characteristics fragmentation and characteristic retention time, by comparison with data from literature[32, 38, 39]. Xcalibur software version 2.1 was used for instrument control, data acquisition and data analysis.

Multivariate Data Analysis

All the data underwent a variety of chemometric analyses, including Principal Component Analysis (PCA) and Partial Least Square Discriminant Analysis (PLS-DA), to identify the true differences between the samples.

Statistical analysis and data mining of the DIMS data were performed using Principal Component Analysis (PCA) and Partial Least Square (PLS) approaches. Principal Component Analysis (PCA) is one method used to uncover important patterns related to physiological, genetic, and environmental factors[40]. Moreover the approach confirms that PCA and PLS-DA analyses could help with predictive information to discriminate organic and conventional crops. Prior to PCA, each DIMS spectrum was background subtracted and the peaks attributed to solvent impurities (in the optimized extract spectra) was discounted. The mass spectrum for each sample direct infusion underwent normalization to remove any run-to-run MS variability.

This data-set was then exported to MS-Excel and analyzed using the statistical package Simca + software 12.0 (Umetrics AB, Umea, Sweden). Chromatograms (LC-MS analysis) were evaluated using the free software package MZmine (http://mzmine.sourceforge. net/) and the resulting metabolomics data were processed using SIMCA-P+ software 12.0 (Umetrics AB, Umea, Sweden). Pareto scaling was applied to all the analytical methods. The entire data matrix (20 samples $\times$ 205 signals for the matrix obtained with manual normalization and 12 samples $\times$ 915 signals for the matrix obtained with MZmine) were first analyzed by Principal Component Analysis (PCA) to define homogeneous cluster of samples. This cluster was then used as Y classes in Partial Least Squares Discriminant Analysis (PLS-DA). Both PCA score and loading plots were generated for the entire dataset to visualize clustering of strawberry fruits. The data were also modeled by Partial Least Squares Discriminant Analysis (PLS-DA) as a supervised approach to exploring clustering relationship. Models were validated by cross-validation techniques and permutation tests according to standardized good practice to minimize false discoveries and to obtain robust statistical models. A small number of metabolites changing during the experiment was extracted and the behavior of each single metabolite was studied by linear mixed-effects model for longitudinal studies[41].

2.5 References

[1]Gasperotti M., Masuero D., Mattivi F., Vrhovsek U. Overall dietary polyphenol intake in a bowl ofbstrawberries: The influence of *Fragaria* spp. in nutritional studies. *Journal of Functional Foods* **2015**, 18, 1057-1070

[2]Jin P., Wang S.Y., Wang C. Y., Zheng Y. Effect of cultural system and storage temperature on antioxidant capacity and phenolic compounds in strawberries. *Food Chemistry* **2011**, 124, 262-270

[3]Olsson, M. E., Andersson, C. S., Oredsson, S., Berglund, R. H., & Gustavsson, K. Antioxidant levels and inhibition of cancer cell proliferation in vitro by extracts from organically and conventionally cultivated strawberries. *Journal of Agriculture and Food Chemistry* **2006**, 54, 1248–1255

[4]Perkins-Veazie, P. Growth and ripening of strawberry fruit. In J. Janick (Ed.), *Horticultural reviews* **1995** (pp. 267–297). John Wiley & Sons, Inc.

[5]Josuttis, M., Verral, S., Stewart, D., Kruger, E., & McDougall, G. J. Genetic and environmental effects on tannin composition in strawberry (*Fragaria ananassa*) cultivars grown in different European countries. *Journal of Agricultural and Food Chemistry* **2013**, 61, 790–800

[6]Balasundram, N., Sundram, K., & Samman, S. Phenolic compounds in plants and agri-industrial by-products: Antioxidant activity, occurrence, and potential uses. *Food Chemistry* **2006**, 99(1), 191–203

[7]Miller, M. G., & Shukitt-Hale, B. Berry fruit enhances beneficial signalling in the brain. *Journal of Agricultural and Food Chemistry* **2012**, 60, 5709–5715

[8] Hannum, S. M. Potential impact of strawberry on human health: A review of the science. *Critical Reviews in Food Science and Nutrition* **2000**, 44, 1–17

[9]Lopes-da-Silva, F., Pascual-Teresa, S. D., Rivas-Gonzalo, J., & Santos-Buelga, C. Identification of anthocyanin pigments in strawberry (cv Camarosa) by LC using DAD and ESI-MS detection. *European Food Research and Technology* **2002**, 212, 248–253

[10]Hakkinen, S. H., Karenlampi, S. O., Mykkanen, H. M., Heinonen, M., & Torronen, A. R. Ellagic acid content in berries: Influence of domestic processing and storage. *European Food Research and Technology* **2000**, 212, 75–80

[11]Williner, M. R., Pirovani, M. E., & Guemes, D. R. Ellagic acid content in strawberries of different cultivars and ripening stages. *Journal of the Science of Food and Agriculture* **2003**, 83, 842–845

[12]Meyers, K. J., Watkins, C. B., Pritts, M. P., & Liu, R. H. Antioxidant and antiproliferative activities of strawberries. *Journal of Agricultural and Food Chemistry* **2003**, 51, 6887–6892

[13]Gil, M. I., Holcroft, D. M., & Kader, A. A. Changes in strawberry anthocyanins and other polyphenols in response to carbon dioxide treatment. *Journal of Agricultural and Food Chemistry* **1997**, 45, 1662–1667

[14]Kosar, M., Kafkas, E., Paydas, S., & Baser, K. H. C. B. Phenolic composition of strawberry genotypes at different maturation stages. *Journal of Agricultural and Food Chemistry* **2004**, 52, 1586–1589

[15]Fischer, U. A., Carle, R., & Kammerer, D. R. Identification and quantification of phenolic compounds from pomegranate (*Punica granatum* L.) peel, mesocarp, aril and differently produced juices by HPLC-DAD–ESI/MSn. *Food Chemistry* **2011**, 127, 807–821

[16]Maatta-Riihinen, K. R., Kamal-Elin, A., & Torronen, A. R. High performance liquid chromatography (HPLC) analysis of phenolic compounds in berries with diode array and electrospray ionization mass spectrometric (MS) detection: *Ribes* species. *Journal of Agricultural and Food Chemistry* **2003**, 51, 6736–6744

[17]Motilva, M. -J., Serra, A., & Macià, A. Analysis of food polyphenols by ultra highperformance liquid chromatography coupled to mass spectrometry: An overview. *Journal of Chromatography A* **2013**, 1292, 66–82

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

[19]De Vos, R. C. H., Hall, R. D., & Moing, A. Metabolomics of a model fruit: Tomato. *Annual Plant Reviews* **2011**, 43, 109–112

[20]Allwood, J., & Goodacre, R. An introduction to liquid chromatography–mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochemical Analysis* **2010**, 21, 33–47

[21]Bedair, M., & Sumner, L. Current and emerging mass-spectrometry technologies for metabolomics. *Trends in Analytical Chemistry* **2008**, 27, 238–250

[22]Makarov, A., & Scigelova, M. Coupling liquid chromatography to Orbitrap mass spectrometry. *Journal of Chromatography* A **2010**, 1217, 3938–3945

[23]Perry, R., Cooks, R., & Noll, R. Orbitrap mass spectrometry: Instrumentation, ion motion and applications. *Mass Spectrometry Reviews* **2008**, 27, 661–669

[24]Zhang, N. R., Yu, S., Tiller, P., Yeh, S., Mahan, E., & Emary, W. B. Quantitation of small molecules using high-resolution accurate mass spectrometers — A different approach for analysis of biological samples. *Rapid Communications in Mass Spectrometry* **2009**, 23, 1085–1094

[25]Nielen, M., Van Engelen, M., Zuiderent, R., & Ramaker, R. Screening and confirmation criteria for hormone residue analysis using liquid chromatography accurate mass time-of-flight, Fourier transform ion cyclotron

resonance and Orbitrap mass spectrometry techniques. *Analytica Chimica Acta* **2007**, 586, 122–129

[26]Schauer, N., & Fernie, A. R. Plant metabolomics: Towards biological function and mechanism. *Trends in Plant Science* **2006**, 11, 508–516

[27]Kim, J. K., Bamba, T., Harada, K., Fukusaki, E., & Kobayashi, A. Time-course metabolic processing in *Arabidopsis thaliana* cell cultures after salt stress treatment. *Journal of Experimental Botany* **2007**, 58, 415–424

[28]Blount, J. W., Masound, S., Sumner, L. W., Huhman, D., & Dixon, R. A. Over expression of cinnamate 4-hydroxylase leads to increased accumulation of acetosyringone in elicited tobacco cell-suspension cultures. *Planta* **2002**, 214, 902–910

[29]Roessner, U., Wagner, C., Kopka, J., Trethewey, R. N., & Willmitzer, L. Simultaneous analysis of metabolites in potato tuber by gas chromatography–mass spectrometry. *The Plant Journal* **2000**, 23, 131–142

[30]Van De Velde, F., Tarola, A. M., Guemes, D., & Pirovani, M. E. Bioactive compounds and antioxidant capacity of Camarosa and Selva strawberries (*Fragaria x ananassa* Duch.). *Foods* **2013**, 2, 120–131

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

[32]Koolen, H. H. F., da Silva, M. A. F., Gozzo, F. C., de Souza, A. Q. L., & de Souza, A. D. L. Antioxidant, antimicrobial activities and characterization of phenolic compounds from buriti (*Mauritia flexuosa* L. f.) by UPLC–ESI–MS/MS. *Food Research International* **2013**, 51, 467–473

[33]Ornelas-Paz, J. d. J., Elhadi, Y. M., Ramirez-Bustamante, N., Perez-Martinez, J. D., EscalanteMinakata, M. d. P., Ibarra-Junquera, V., et al. Physical attributes and chemical composition of organic strawberry fruit

(*Fragaria x ananassa* Duch. Cv. Albion) at six stages of ripening. *Food Chemistry* **2013**, 138(1), 372–381

[34]Maeaettae-Riihinen, K. R., Kamal-Eldin, A., & Toerroenen, A. R. Identification and quantification of phenolic compounds in berries of *Fragaria* and *Rubus* Species (Family Rosaceae). *Journal of Agricultural and Food Chemistry* **2004**, 52(20), 6178–6187

[35]Horai, H., Arita, M., Kanaya, S., Nihei, Y., Ikeda, T., Suwa, K., Ojima, Y., Tanaka, K., Tanaka, S., Aoshima, K., Oda, Y., Kakazu, Y., Kusano, M., Tohge, T., Matsuda, F., Sawada, Y., Hirai, M. Y., Nakanishi, H., Ikeda, K., Akimoto, N., Maoka, T., Takahashi, H., Ara, T., Sakurai, N., Suzuki, H., Shibata, D., Neumann, S., Iida, T., Tanaka, K., Funatsu, K., Matsuura, F., Soga, T., Taguchi, R., Saito, K., & Nishioka, T. MassBank: a public repository for sharing mass spectral data for life sciences. *Journal of Mass Spectrometry* **2010**, 45, 703–714

[36]De Ancos, B., Gonzàlez, E. M., & Cano, M. P. Ellagic acid, vitamin C, and total phenolic contents and radical scavenging capacity effected by freezing and frozen storage in raspberry fruit. *Journal of Agricultural and Food Chemistry* **2000**, 48, 4565–4570

[37]Mullen, W., Stewart, A. J., Lean, M. E. J., Gardner, P., Duthie, G. G., & Crozier, A. Effect of freezing and storage on the phenolics, ellagitannins, flavonoids and antioxidant capacity of red raspberries. *Journal of Agricultural and Food Chemistry* **2002**, 50, 5197–5201

[38]March, R., & Brodbelt. Analysis of flavonoids: Tandem mass spectrometry, computational methods, and NMR. *Journal of Mass Spectrometry* **2008**, 43, 1581–1617.

[39]Vukics, V., & Guttman, A. Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. *Mass Spectrometry Reviews* **2010**, 29, 1–16

[40]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. *Journal of Pharmaceutical and Biomedical Analysis* **2012**, 70, 224–230

[41]Laird, N. M., & Ware, J. H. Random-effects models for longitudinal data. *Biometrics* **1982**, 38, 963–974

Chapter 3

Metabolomics approaches used for typization of Berry fruit

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.

With this aim, new MS-based strategies combined to Multivariated Data Analysis were applyed for food quality and geographical origin assessment of two plant species: *Fragaria vesca* (endemic of Campania region, Italy) and *Myrtus communis* (endemic of Sardinia region, Italy). These species are recognized as traditional food product of Italian regions by "Ministry of Agriculture Food and Forestry Policies".

The objective of these approaches, combining targeted and untargeted Metabolomics analysis with Chemometrics, is finally to individuate chemical markers for quality control, for authenticity assessment and for adulteration evaluation of food products.

In particular for *Myrtus communis* berries were developed two strategies:

1) metabolomics approaches were carried out for the descrimination of 42 cultivars of *Myrtus communis*, whose seeds were collected from different area of Sardinia and cultivated then in the same experimental field of Sassari University located in Oristano (SS). In this study a new approach was designed and called "pseudo-targeted": results obtained by mass spectrometry in positive ion mode were fused with selected results obtained in negative ion mode to treat a combined data matrix with chemiometric unsupervised approach, that is finally compared with an untargeted approach (performed on negative ion mode MS data). In fact the principal obstacle for a metabolic fingerprint focused on myrtle berry is that anthocyanins have a good ionization in positive ion mode, while the other phenolic compounds have better

53

ionization in negative ion mode, resulting in the necessity of switching polarity to have a metabolic profiling.

2) metabolomics approaches were applied focusing the study on two cultivars of myrtle, that in the previous study were defined as most different between the 42 cultivar analysed, and a phytochemical investigation was performed on different parts of myrtle berries extracts (Seeds, Peel and Pulp).

3.1 Characterisation of *Fragaria vesca* fruit from Italy following a metabolomics approach through integrated Mass Spectrometry techniques

3.1.1 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 ananassa*[1]. *F. vesca* is considered to be cosmopolite and is very common in the underwood in temperate areas[2]. The fruit of *F. vesca* is, therefore, a false fruit under the botanical point of view, 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, 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) and tannins (ellagitannins, procyanidins) [1,3], 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 [4,5].

Southern Italy is an important producer of *F. vesca* berries and in the Campania Region, particularly in the Alburni and the Alto Sele areas, spontaneous strawberries are officially recognized as a traditional food product under the name "Fragolina degli Alburni". In this 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 used in the preparation of traditional sweets. Unfortunately, this spontaneous native strawberry germplasm blossoms only once per year, thus supplying a fruit yield that is concentrated in the harvesting period. In order to cope with an increasing demand of wild strawberries, local farmers are, therefore, encouraged to introduce non-native strawberry germplasm 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 share local peculiarity, although chemical markers are not always available to discern among food stocks coming from different areas. Recent applications of MS-based strategies in food-omics include metabolomics studies for food quality and geographical origin assessment [6]. Moreover, it has also been reported that changes in the level of involved secondary metabolites are, to some extent, determined by environmental factors but are much more regulated at the genetic level. In the research focusing on berries, both targeted

and untargeted metabolic profiling have been used in studying phytochemical diversity and relationships between cultivars[7-9].

The aim of this study was to explore the discrimination of phytochemical content of the fruit of *F. vesca* collected in different geographic areas, under different environmental conditions, and from different germplasm using a LC-ESI-Orbitrap-MS metabolic profiling approach. The main focus was on identifying differences in the semi-polar metabolite pool, that may be useful in discriminating the geographic origin of wild strawberries.

3.1.2 Results and discussion

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

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 3.1.1, Fig. 3.1.1 and Fig. 3.1.2. 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. 3.1.2A). The positive LC-MS profile highlighted the presence of anthocyanins (Fig. 3.1.2B).
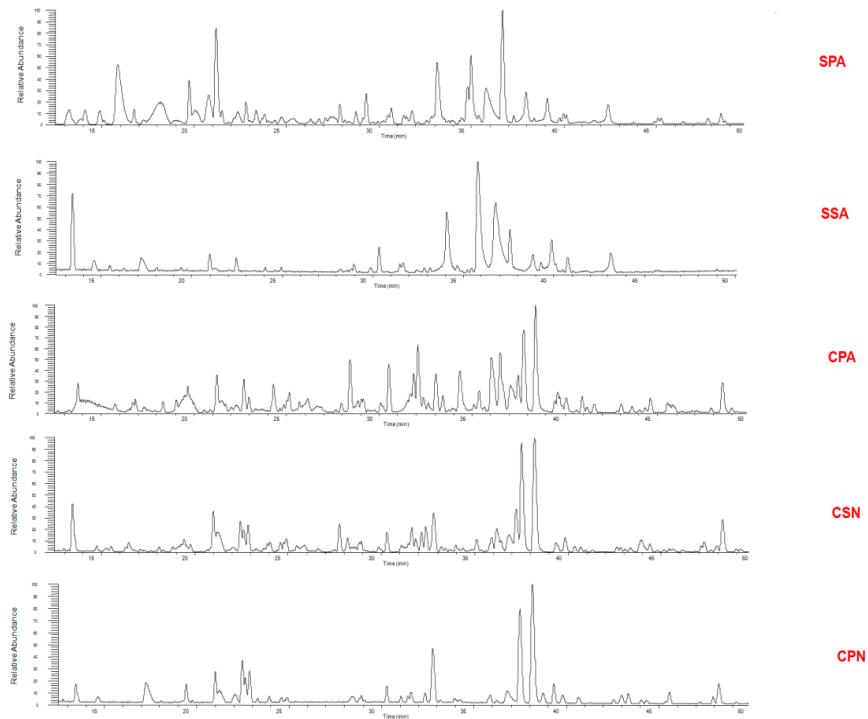
Fig.3.1.1: LC-ESI-Orbitrap-MS Full Scan (200-1600 *m/z*) profiles in negative ion mode of *Fragaria vesca* fruit's methanol extracts, for different samples: SPA Spontaneous, Petina, Autochthonous; SSA Spontaneous, Sarno, Autochthonous; CPA Cultivated, Petina, Autochthonous, CSN Cultivated, Sarno, Non Autochthonous; CPN Cultivated, Petina, Non Autochthonous.
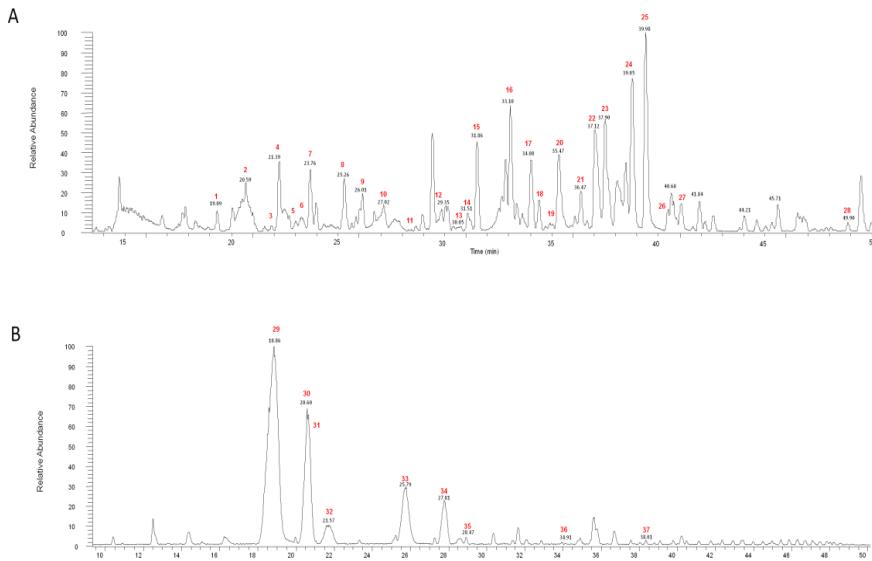
Fig.3.1.2: LC-ESI-Orbitrap-MS Full Scan (200-1600 *m/z*) in negative ion mode (A) and positive ion mode (B) of *Fragaria vesca* fruit's methanol extract obtained from sample CPA Cultivated, Petina, Autochthonous.

High Resolution MS[n] detection (positive and negative ionisation 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 [1,3,10]. 37 compounds, including anthocyanins, dihydroflavonols and flavonols, flavan-3-ols, proanthocyanidins, and ellagic acid and its derivatives were identified from the different *Fragaria vesca* samples. The molecular formula of the compounds are summarised in Table 3.1.1, where 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

59

(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 literature, these information helped to identify compounds.

Table 3.1.1: Identification of polyphenolic compounds and anthocyanins in *F. vesca* fruit using HPLC-ESI-Orbitrap-MS/MS

Compounds identified in negative ion mode

		Molecular formula	MW	[M-H]⁻	M⁺	RT (min)	MS/MS	References
1	kaempferol hexoside	$C_{21}H_{20}O_{11}$	448.1005	447.0926	19.09		285	1,3
2	B-type procyanidin (B1)	$C_{30}H_{26}O_{12}$	578.1424	577.1344	20.59		425/289	1,3
3	B-type (epi)afzelechin-(epi)catechin	$C_{30}H_{26}O_{11}$	562.1475	561.1395	21.21		543/435/289	1,3
4	Bis HHDP hexose (pedunculagin)	$C_{34}H_{24}O_{22}$	784.0759	783.0679	21.39		481/301	1,3
5a/5b	catechin/epicatechin	$C_{15}H_{14}O_6$	290.0790	289.0710	22.06		245/205/179	1,3
6	p-coumaroyl-hexoside	$C_{15}H_{18}O_8$	326.1001	325.0922	22.28		265/187/162/145	12
7	HHDP galloylhexose	$C_{27}H_{22}O_{18}$	634.0816	633.0736	23.76		301/452	1,3
8	B-type (epi)catechin trimer (procyanidin C1)	$C_{45}H_{38}O_{18}$	866.2058	865.1868	25.26		739/695/577/289	1,3
9	B-type (epi)afzelechin-(epi)catechin	$C_{30}H_{26}O_{11}$	562.1471	561.1392	26.01		435/407/289	1,3
10	B-type procyanidin (B2)	$C_{30}H_{26}O_{12}$	578.1424	577.1344	27.01		425/289	1,3
11	dihydromyricetin-hexose	$C_{21}H_{22}O_{13}$	482.1060	481.0981	27.70		463/319/301	11
12	galloylhexose-HHDP-gallic acid	$C_{34}H_{26}O_{22}$	786.0915	785.0830	29.35		615/463/301	1,3
13	procyanidin B	$C_{30}H_{26}O_{12}$	578.1424	577.1344	30.05		425/289	1,3
14	tris-galloyl-HHDP-hexose	$C_{41}H_{28}O_{27}$	952.0818	951.0725	31.51		907/783/484/301	1,3
15	casuarictin	$C_{41}H_{28}O_{26}$	936.0868	935.0789	31.86		633/451/301	1,3
16	peltatoside	$C_{26}H_{28}O_{16}$	596.1377	595.1297	33.10		463/301	1,3
17	quercetin-rhamnoside	$C_{21}H_{20}O_{11}$	448.1005	447.0926	34.08		301	1,3
18	ellagic acid rhamnoside	$C_{20}H_{16}O_{12}$	448.0641	447.0562	34.58		301/300/257	1,3
19	castalagin/vescalagin isomer	$C_{41}H_{26}O_{26}$	934.0712	933.0632	35.03		915/631/451/301	1,3
20	methyl ellagic acid glucuronide	$C_{21}H_{16}O_{14}$	492.0540	491.0460	35.47		315/300	13
21	isoquercitrin	$C_{21}H_{20}O_{12}$	464.0954	463.0875	35.74		301	1,3
22	ellagic acid	$C_{14}H_6O_8$	302.0062	300.9983	37.12		285/257/229	1,3
23	phloridzin	$C_{21}H_{24}O_{10}$	436.1369	435.1289	37.90		273	3

60

24	taxifolin 3-alpha-L-arabinofuranoside	$C_{20}H_{20}O_{11}$	436.1005	435.0926		39.05	303/285	1,3
25	quercetin glucuronide	$C_{21}H_{18}O_{13}$	478.0747	477.0667		39.98	301	1,3
26	dimethyl ellagic acid pentoside	$C_{21}H_{18}O_{12}$	462.0798	461.0718		40.05	446/315/300	1,3
27	kaempferol glucuronide	$C_{21}H_{18}O_{12}$	462.0798	461.0877		40.38	451/386/285	12
28	kaempferol coumaroyl hexoside	$C_{30}H_{26}O_{13}$	594.1373	593.1293		49.90	447/285/255	1,3

Compounds identified in positive ion mode

		Molecular formula	MW	[M-H]⁻	M⁺	RT (min)	MS/MS	References
29	cyanidin-3-*O*-glucoside	$C_{21}H_{21}O_{11}$	449.1083		449.1083	18.86	287	3
30	pelargonidin -3-*O*-glucoside	$C_{21}H_{21}O_{10}$	433.1134		433.1134	20.6	271	3
31	pelargonidin-3-*O*-rutinoside	$C_{27}H_{31}O_{14}$	579.1713		579.1713	20.66	433/271	14
32	peonidin glucoside	$C_{22}H_{23}O_{11}$	463.1240		463.1240	21.57	301	3
33	cyanidin malonyl glucoside	$C_{24}H_{23}O_{14}$	535.1087		535.1087	25.79	287	3
34	pelargonidin malonylglucoside	$C_{24}H_{23}O_{13}$	519.1138		519.1138	27.81	433/271	3
35	peonidin malonylglucoside	$C_{25}H_{25}O_{14}$	549.4576		549.4576	28.47	301/463	3
36	delphinidin -3-*O*-glucoside	$C_{21}H_{21}O_{12}$	465.1033		465.1033	34.9	303	14
37	delphinidin malonyl glucoside	$C_{24}H_{23}O_{15}$	551.1036		551.1036	38.01	303	15

Compounds identified using UHPLC system interfaced to an ABSciex (Foster city, CA, USA) API4000 QTrap in negative ion mode

		Molecular formula	MW	[M-H]⁻	M⁺	RT (min)	MS/MS	References
38	gallic acid	$C_7H_6O_5$	170,02	169.01		3.35	125	3
39	p-coumaric acid	$C_9H_8O_3$	164,04	163.03		15.15	119	3

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 *Fragaria vesca*.

Compound **11**, dihydromyricetin hexose, never detected in the genus *Fragaria*, showed a pseudolmolecular ion at *m/z* 481.0981 that submitted to fragmentation gave a principal daughter ion at *m/z* 319, corresponding to the

loss of an hexose unit. A similar fragmentation was observed and the compound was tentatively identified by High Resolution ESI-MS and High Resolution ESI-MS/MS by Abu-Reidah et al. (2014) in *Vicia faba* (Fabaceae)[11].

Compound **27**, kaempferol-glucuronide, never detected in this species, with a pseudomolecular peak at 461.0877 was identified based on the exact mass and fragmentation. MS/MS spectrum showed a most abundant peak at 285 *m/z*. The same compound was identified in *Fragaria ananassa* analysed in ESI-MS and ESI-MS/MS, by Seram et al. (2012)[12].

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**, methyl ellagic acid glucuronide, never detected in this species, with a pseudomolecular peak at 491.0460, 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 et al. (2012), in *Cochlospermum angolensis*[13].

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[3].

Compound **23**, identified as phloridzin, pertains to the class of dihydrochalcones , was described previously in *F. vesca*[3].

Compounds **29**, **30**, **31**, **32**, **33**, **34**, **35**, **36** and **37** were identified as anthocyanin derivatives, most of them previously identified in *Fragaria vesca*[3].

62

Compound **31**, pelargonidin-rutinoside, never detected in *F. vesca*, showed a pseudolmolecular ion at *m/z* 579.1713 that submitted to fragmentation gave two principal daughter ion at *m/z* 433, and *m/z* 271 corresponding to the lost of a pentose unite and a following loss of an hexose unit. This fragmentation pattern for pelargonidin-rutinoside was described by Wu et al. in 2005[14].

Compound **36**, delphinidin-glucoside, never detected in *F. vesca*, showed a pseudolmolecular ion at *m/z* 465.1033 that in MS/MS gave a principal daughter ion at *m/z* 303, corresponding to the aglycon. This compound was described by ESI-MS/MS data by Wu et al. in 2005[14].

Compound **37**, delphinidin-malonyl-glucoside, never detected in *F. vesca*, showed a pseudolmolecular ion at *m/z* 551.1036 that in MS/MS gave a principal daughter ion at *m/z* 303, corresponding to the aglycon. This compound was tentatively identified by ESI-MS/MS data by Dugo et al. in 2003 in red orange juice[15].

Multivariate Data Analysis on qualitative data

Preliminary qualitative investigation on CPA (cultivated Petina autochthonous), CSA (cultivated Sarno autochthonous), CPN (cultivated Petina non-autochthonous), SPA (spontaneous Petina autochthonous), SSS (spontaneous Sarno autochthonous), showed a different flavonoid and anthocyans distribution (Fig. 3.1.3).

For 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 take place in a different area of the plot.

Principal Component Analysis (PCA) was performed by applying the peak areas of each peak present in the LC-MS dataset (excluding the noisy), and a matrix was obtained by using these areas (variables), and the columns of the matrix were the different analysed samples. The resulted score scatter plot is reported in Fig. 3.1.3 (A and B).

The first component explains 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 model.

Results show a good discrimination 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.

Figure 3.1.3 shows the score scatter plot coloured according to spontaneous or cultivated specimens (Panel A), and it is clear a differentiation of the samples based on this parameter, while the same scatter plot in panel B is coloured 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.

64
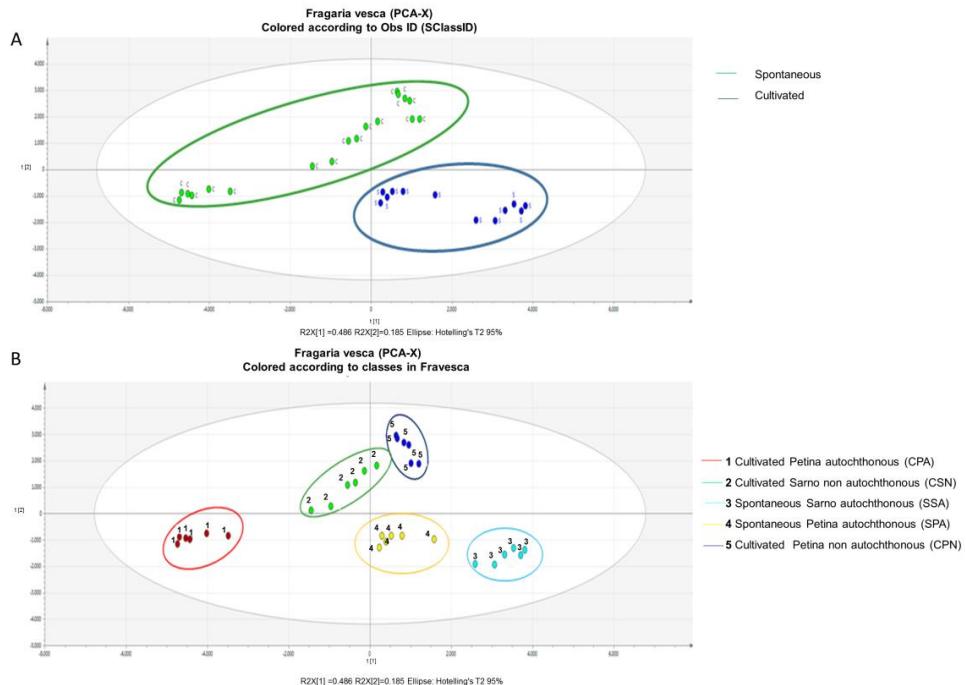
Fig 3.1.3: Principal Component Analysis, score scatter plot, untargeted LC-ESI-Orbitrap-MS analysis. Panel A: colored according spontaneous (S) versus cultivated (C) species. Panel B: colored according samples classification.

LC-ESI-QTrap-MS/MS quantitative analysis

The results obtained from the targeted analysis by LC-ESI-QTrap-MS of *Fragaria vesca* fruit by quantification of selected marker compounds (based on commercial available standards) and Principal Component Analysis are reported in Tab 3.1.2, Fig. 3.1.4. 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 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 3.1.2, that shows in addition mass spectral parameters for the standard compounds.

The table 3.1.2 in addition shows the results for quantitative analysis of phenolic compounds in the *F. vesca* samples object of this study.

Table 3.1.2: Quantitative results obtained for selected phenolic compounds in *F. vesca* fruit using an UHPLC system interfaced to an ABSciex (Foster City, CA, USA) API4000 Q-Trap instrument in ion trap mode.

	DP	EP	CE	PI	DI	CPA	SPA	SSA	CPN	CSN
Positive ion mode										
29. cyanidin-3-*O*-glucoside[a]	192	12	27.2	449	287	113.35+/-8.60	59.20+/-3.250	24.45+/-1.950	18.71+/-2.060	26.70+/-2.630
36. delphinidin-3-*O*-glucoside[a]	92.8	11	26	465	303	0.090+/-0.008	0.060+/-0.005	0.042+/-0.004	0.026+/-0.005	0.035+/-0.009
30. pelargonidin-3-*O*-glucoside[a]	76.9	10	28.8	433	271	7.730+/-0.108	6.85+/-0.131	7.15+/-0.098	6.61+/-0.028	9.20+/-0.094
31. pelargonidin-3-*O*-rutinoside[a]	98	13	40	579	271	0.025+/-0.009	0.028+/-0.003	0.033+/-0.001	0.033+/-0.003	0.054+/-0.003
Negative ion mode										
5a. catechin[a]	-93	-9	-22	289	245	0.049+/-0.004	0.617+/-0.003	0.007+/-0.001	0.030+/-0.003	0.094+/-0.005
5b. epicatechin[a]	-93	-9	-22	289	245	1.480+/-0.050	2.98+/-0.070	0.44+/-0.010	1.159+/-0.020	1.563+/-0.040
2. procyanidin B1[a]	-99	-3	-35	577	289	2.147+/-0.070	3.76+/-0.110	0.838+/-0.020	1.624+/-0.020	2.449+/-0.050
10. procyanidin B2[a]	-99	-3	-35	577	289	0.035+/-0.003	0.035+/-0.005	0.033+/-0.001	0.024+/-0.002	0.056+/-0.004
21. isoquercitrin[a]	-115	-4	-40	763	301	2.613+/-0.063	0.313+/-0.015	0.078+/-0.006	0.180+/-0.006	0.630+/-0.011
38. gallic acid[a]	-59	-6	-20	169	125	9.500+/-0.320	5.986+/-0.080	11.70+/-0.155	4.14+/-0.159	5.72+/-0.051
39. p-coumaric acid[a]	-56	-2	-20	163	119	0.625+/-0.006	0.620+/-0.004	0.671+/-0.005	0.597+/-0.003	0.600+/-0.004
23. phloridzin[a]	-89	-4	-22	435	273	0.342+/-0.023	0.235+/-0.010	0.042+/-0.010	0.165+/-0.010	0.234+/-0.017
Total						137.98+/-0.77	80.68+/-0.31	45.48+/-0.19	33.29+/0.19	47.33+/-2.90

[a]Mean in mg/100g fresh weight, DP Declustering Potential EP Entrance potential, CE Collision energy, PI Product ion, DI Daughter ion,
CPA Cultivated, Sarno, Autochthonous, SPA Spontaneous, Petina, Autochthonous, SSA Spontaneous, Sarno, Autochthonous, CPN Cultivated, Petina, Non Autochthonous, CSN Cultivated, Sarno, Non Autochthonous

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 area of "Alburni".

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 0.001-0.012 μg/ml and LOQ from 0.004-0.09 μg/ml, indicating that the developed method exhibited good sensitivity. The results for each compound are given in Table 3.1.3.

Table 3.1.3: Validation data for the quantitative UPLC-MS/MS method

		r^2	Regression equation	LOD	LOQ	Precision RSD (%)	
						Intraday	Interday
				(mg/ml)	(mg/ml)	(n=3)	(n=3)
29	cyanidin-3-*O*-glucoside	0.9940	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=2.33e6 x +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

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 3.1.3).

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%, demonstrating good recovery and precision.

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, using SIMCA-P+ Software, 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). Figure 3.1.4 A 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, 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 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 figure 3.1.4 B: 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. 3.1.4 B). Derivatives of cyanidin (i.e. cyanidin-3-*O*-glucoside) appear to be overexpressed 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.
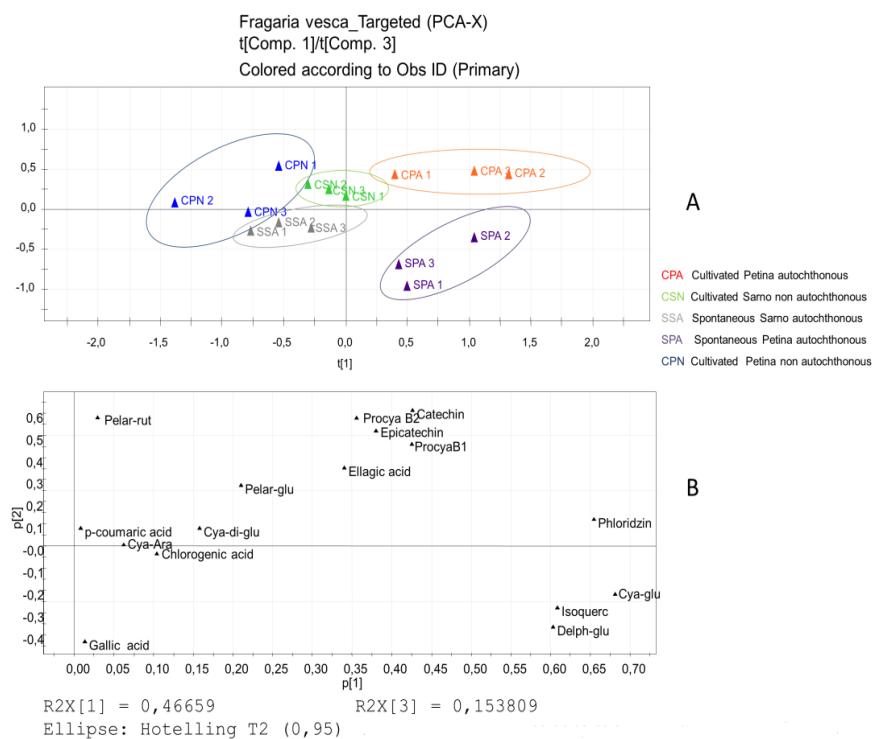

Fig. 3.1.4: Principal Component Analysis, targeted analysis data. Panel A: score scatter plot; Panel B: loading scatter plot.

70

3.1.3 Conclusions

The goal of the present study was to explore the discrimination of phytochemical content between two different geographic origins of spontaneous strawberry and two different germplasms (one autochthonous of Alburni and another non-autochthonous of Alburni) by applying untargeted and targeted liquid chromatography-mass spectrometry techniques combined with Principal Component Analysis. 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 higher amount of anthocyanins if compared to the ones from other geographical origin. The study highlights the potential efficacy of untargeted metabolite profiling in the assessment of fruits origin. This is a very important issue because the fruit value is truly 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 equipments are expensive and not routine techniques, such approaches are a useful way to identify specific origin and authenticity markers, that, when defined, can be analysed with cheaper and easier techniques in common routine labs for quality control.

3.1.4 Materials and Methods

Fruit samples

Wild strawberry fruits were collected in June 2014 in two different location in Campania Region, Petina (Alburni Mountains Area) and Sarno (Sarno River Plain Area), from populations spontaneously growing in the underwood (spontaneous) and in crop (cultivated) and from autochthonous and non-autochthonous germplasm. For each typology (5 types) three samples were collected (biological triplicates) with a result of 15 biological samples, classified in 5 groups as in the following list: 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, because *Fragaria vesca* is not a typical production of the area of Sarno, thus crops in this area are developed by imported plant material.

Reagents and solvents

Standards as procyanidin B1, procyanidin B2, isoquercitrin, epicatechin, catechin, gallic acid, p-coumaric acid, phloridzin, cyanidin-glucoside, delphinidin glucoside, pelargonidin glucoside, pelargonidin-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).

Sample preparation

Fruits were extracted immediately after collection, to prevent degradation of secondary metabolites. The extraction was done following the method proposed by Pavlovic et al. 2013 with slight modification[16]. 5g of fresh fruits were extracted for 24 h with methanol acidified with 0.1% of formic acid, then 200 μL of filtered extracts were dried under N_2 and diluted in 1 mL of water acidified with 0.1% of formic acid. Extract samples 10 μL were used for LC-MS analysis. Each sample was analysed in technical duplicate.

LC-ESI-Orbitrap-MS analysis

To investigate the main markers specific to each fruits, an LC-ESI-Orbitrap-MS method was developed in this research. Methanol extracts were used for this 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 column (2.1x150 mm i.d., 3.5 μm d.) (Waters). The mobile phase consisted of solvent A (water + 0,1% formic acid) and solvent B (acetonitrile/water (8:2) + 0,1% 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 to 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 -12V; tube lens voltage -121.47V; capillary temperature 280°C; Sheath and Auxiliary Gas flow (N_2) 30 and 5 (arbitrary units), Sweep gas 0 (arbitrary units) Spray voltage 5V. MS spectra were acquired by full range acquisition

73

covering *m/z* 200-1600. 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 49V; tube lens voltage 120V; capillary temperature 280°C; Sheath and Auxiliary Gas flow (N_2) 30 and 5(arbitrary units), Sweep gas 0 (arbitrary units) Spray voltage 5V. 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.

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.

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 µg ml^{-1} in methanol 50%) into the source at a flow rate of 10 µL/min. Data were acquired in the positive or negative ion MS and MS/MS modes depending of the standard compound.

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 (5 μl) were injected into a XSelect HSS T3 column (Waters, Milford, MA) (100 x 2.1 mm i.d., 2.5 μm d). Mobile phase A was H_2O containing 0.1% formic acid while mobile phase B was acetonitrile containing 0.1% 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 300 μl to 350μl; 4-6 min, linear gradient 0%-12% B and the flow achieves 400 μl/min; 6-12 min, linear gradient 12%-20% B and flow constant at 400 μl/min; 16-17 min, linear gradient 20%-100% B and flow retrieves to 300 μl/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 300 μl/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).

Method validation

LC-ESI-QTrap-MS/MS method was validated according to the European Medicines Agency guidelines (EMEA Quality guidelines ICH Q2)[17] 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.

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, we detected 170 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 PCA. PCA was performed by applying the peak area obtained from LC/MS analysis (Mari et al., 2009; 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 Principal Component Analysis (PCA) in order to identify similarities among our samples. Pareto scaling was applied before multivariate data analysis.

3.1.5 References

[1]Sun, J., Liu, X., Yang, T., Slovin, J., Chen, P. Profiling polyphenols of two diploid strawberry (*Fragaria vesca*) inbred lines using UHPLC-HRMSn. *Food Chemistry* **2014**, 146, 289-298

[2]Yildirim, A. B., Turker, A. U. 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* **2014**, 169, 169–178

[3]Del Bubba, M., Checchini, L., Chiuminatto, U., Doumett, S., Fibbia, D., Giordanic, E. 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* **2012**, 47, 1207-1220

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

[5]Cheel, J., Theoduloz, C., Rodriguez, J. A., Caligari, P. D. S., Schmeda-Hirschmann, G. Free radical scavenging activity and phenolic content in

achenes and thalamus from Fragaria chiloensis spp. chiloensis, *F. vesca* and *F. x ananassa* cv. Chandler. *Food chemistry* **2007**, 102, 36−44

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

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

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

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

[10]Rohloff J., J. Kopka, A. Erban, P. Winge, R. C. Wilson, A. M. Bones, J. Davik, et al.. Metabolite profiling reveals novel multi-level cold responses in the diploid model *Fragaria vesca* (woodland strawberry). *Phytochemistry* **2012**, 77: 99–109

[11]Abu-Reidah, I. M., Contreras, M. d. M., Arraez-Rom ´an´, D., Fernandez-Gutierrez A., Segura-Carretero A. UHPLC-ESI-QTOF-MS-based metabolic profiling of *Vicia faba* L. (Fabaceae) seeds as a key strategy for characterization in foodomics. *Electrophoresis* **2014**, 35, 1571–1581

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

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

[14]Wu, X., Prior, R., L. 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***2005**, 53, 3101-3113

[15]Dugo, P., Mondello, L., Morabito, L., Dugo, G. Characterization of the Anthocyanin Fraction of Sicilian Blood Orange Juice by Micro-HPLC-ESI/MS. *Journal of Agricultural and Food Chemisty* **2013**, 51, 1173−1176

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

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

[18]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. *Journal of Pharmaceutical and Biomedical Analysis* **2012**, 70, 224-230

[19]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, 1379–1389

3.2 Targeted and Untargeted Mass Spectrometric approaches in discrimination between Myrtus communis cultivars from Sardinia region

3.2.1 Introduction

Myrtus communis L. is a plant traditionally used as an antiseptic and disinfectant drug[1]. Concerning to the chemical study of the plant, several compounds have been isolated from the leaves[2-3], the essential oil[4] and the fruits[5]. Berries of *M. communis L.* are used to produce the characteristic myrtle liqueur typical of Sardinia, Italy. The polyphenolic components of berries from *M. communis* have been subject of several investigations[6-12]. These studies have reported the presence of delphinidin-, petunidin-, malvidin-, peonidin- and cyanidin- 3-mono- and 3,5-diglucosides[5], along with glycosides of myricetin and quercetin. In addition the arabinoside derivatives of delphinidin, malvidin, petunidin and cyanidin were identified in myrtle berries extracts[11-12]. Traditionally, cultivars have been distinguished by their phenotypical characteristics like shape and color fruit, precocity of maturation, the pulp/stone ratio, size of leaves, resistance to disease, etc. Notwithstanding their major role and usefulness in characterizing cultivars, with the advent of specific chemical methodologies the chemical composition of plants (berries, leaves and flowers), and principally metabolomics and metabolic profiling began to be used as possible markers of geographic origin or myrtle cultivars.

Metabolomics is the science of the small molecules. State of the art technologies allow high throughput comprehensive analysis of complexes metabolite mixtures typical of food materials[13-14].

The main advantage of metabolomics in food authentication is in its untargeted nature, which can enable the detection of emerging frauds. Thus, several works are appearing in the last years making use of metabolomics approaches for quality assessment of plant used in food industry[15]. Modern mass spectrometry techniques, because of their sensitivity and selectivity, have become methods of choice to characterize plant metabolome[16-17].

The aim of this work is to provide a better knowledge of the chemical characteristics of different cultivars of *M. communis* and to compare an untargeted approach for metabolomics analysis to a specific metabolic profiling based on phytochemical knowledge of Myrtle berry. The characterization should lead to a better valorization of myrtle products by a standardization process of the quality.

Geographical origin of a plant crop 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 final product. Recent applications of MS-based strategies in foodomics include metabolomic studies for food quality and geographical origin assessment[18].

Thus, in the study focused on *M. communis* berries, both targeted and untargeted metabolic profiling have been used in studying phytochemical diversity and relationships between cultivars[19-21] coming from seeds collected from different geographical area of Sardinia region (Italy) and then grown in the same experimental area.

Experiments were carried on myrtle cultivars grown in an experimental orchard located in Fenosu (Oristano, Sardinia, Italy), obtained by the

germination of seeds taken from berries collected in each part of the region (and classified based on the geographical provenience of seeds).

Berries were randomly harvested from 42 plants of *M. communis* during fruit ripening stage.

Discrimination of berry's phytochemical content by LC-ESI-FT-MS metabolic profiling here is explored, and then a targeted quantitative approach by LC-PDA was used to define different contents in anthocyanidins. The profiles of the anthocyanins present in fruits and vegetables can be used as fingerprints through which the authenticity of raw materials, products, juices and extracts may be assessed. So, finally a quantitative method was developed, validated and applied for the quantitative analysis of berries' extracts based on six commercial standard compounds (the markers identified).

3.2.2 Results and discussion

Qualitative   LC-ESI-FT-Orbitrap-MS   and   LC-ESI-FT-Orbitrap-MS/MS: identification of compounds.

The final goal of an untargeted metabolome strategy is to obtain a chemical fingerprint by measuring most of the metabolites present in the extract. When working in LC-MS unfortunately not all the metabolite are detectable with the same mass spectrometric parameters.

To investigate the main markers specific to each fruits, an LC method coupled with an hybrid mass spectrometer, which combines the linear trap quadrupole (LTQ) and OrbiTrap mass analyzer, was developed in this research. Methanol extracts obtained from berry fruits by enhancing extraction with ultrasounds and then submitted to centrifugation, were used for this analysis.

In order to investigate anthocyanin derivatives, for each of the 42 samples, a full positive LC-MS scan, in the form of a total ion current chromatogram

83

(TIC), was initially acquired; successively reconstructed ion chromatograms (RICs) were generated for each of the expected *m/z* values based on the molecular weights of the possible constituents. This step was very important in order to obtain a high selectivity for the analysis.

In order to investigate polyphenolic compounds different from anthocyans, for each of the 42 samples, a full negative LC-MS scan, in the form of a total ion current chromatogram (TIC), was initially acquired; successively reconstructed ion chromatograms (RICs) were generated for each of the expected *m/z* values based on the molecular weights of the possible constituents.

LC-ESI-FT-Orbitrap-MS analysis was applied to the full characterization of the different metabolites in the extracts of fruits. The chromatograms so acquired, for a sample model are displayed in Fig. 3.2.1, and they demonstrate that different metabolites were revealed with different ionization mode.
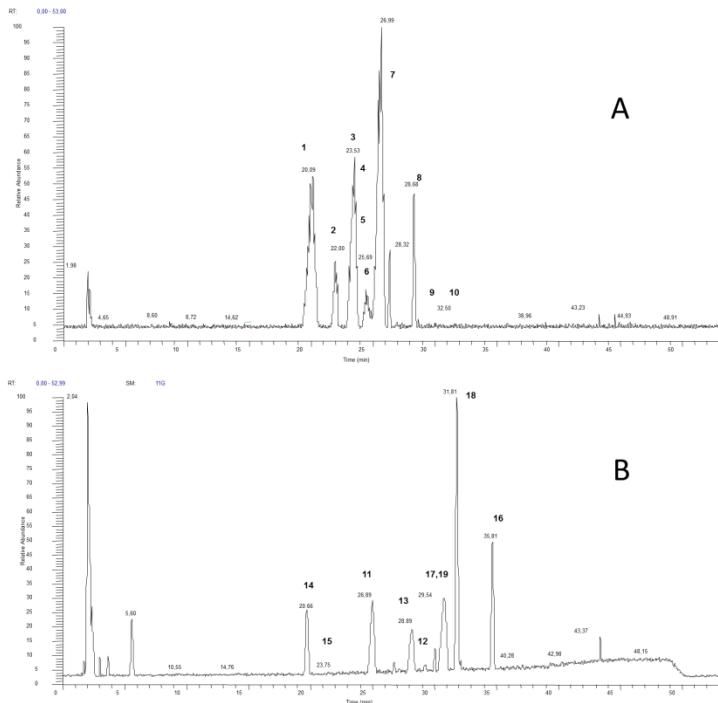
Fig 3.2.1: LC-ESI- Orbitrap (FT) MS of a samples of Myrtle berries acquired in positive (A) and negative (B) ion mode.

Berries collected from different plants showed dissimilar metabolic profiles.

In order to obtain a full identification of compounds LC-ESI-FT-Orbitrap-MS/MS experiment were run by using data dependant analysis (DDA), by selecting High Resolution experiment for precursor ions and low resolution experiments for MS/MS spectra resulting.

Only one sample for each geographical origin of the seeds, was selected for the full characterization by LC-ESI-FT-Orbitrap-MS/MS experiments.

The results obtained from the compound identification by LC-ESI-FT-Orbitrap-MS and LC-ESI-FT-Orbitrap-MS/MS are reported in Table 3.2.1.

Table 3.2.1: Identification of polyphenolic compounds and anthocyanins in *Myrtus communis* berries using HPLC-ESI-FT-(Orbitrap)-MS in positive and negative ion mode

Peak	R_T (min)	[M-H]⁻	[M]⁺	Formula	Identification
1	20.09		465.1033	$C_{21}H_{21}O_{12}$	Delph-3-*O*-Glu
2	22.00		449.1083	$C_{21}H_{21}O_{11}$	Cyan-3-*O*-Glu
3	23.53		479.1189	$C_{22}H_{23}O_{12}$	Petun-3-*O*-Glu
4	24.11		463.1240	$C_{22}H_{23}O_{11}$	Peon-3-*O*-Glu
5	25.69		493.1396	$C_{23}H_{25}O_{12}$	Malv-3-*O*-Glu
6	26.15		435.0925	$C_{20}H_{19}O_{11}$	Delph-3-*O*-Ara
7	26.88		419.0978	$C_{20}H_{19}O_{10}$	Cya-3-*O*-Ara
8	26.99		449.1083	$C_{21}H_{21}O_{11}$	Petu-3-*O*-Ara
9	31.90		433.1134	$C_{21}H_{21}O_{10}$	Peo-3-*O*-Ara
10	32.50		463.1240	$C_{22}H_{23}O_{11}$	Malv-3-*O*-Ara
11	26.89	479.0825		$C_{21}H_{20}O_{13}$	Myr-3-*O*-Gal
12	28.99	463.0876		$C_{21}H_{20}O_{13}$	Myr-3-*O*-Rha
13	28.88	449.0720		$C_{20}H_{18}O_{12}$	Myr-3-*O*-Ara
14	20.66	463.0876		$C_{21}H_{20}O_{12}$	Que-3-*O*-Glu
15	22.53	447.0927		$C_{21}H_{20}O_{12}$	Que-3-*O*-Rha
16	35.81	317.0375		$C_{15}H_{10}O_8$	Myricetin
17	29.54	301.0348		$C_{15}H_{10}O_8$	Quercetin
18	26.16	285.0399		$C_{15}H_{10}O_6$	Kaempferol
19	29.54	300.9984		$C_{14}H_5O_8$	Ellagic acid

HRMSⁿ detection in both positive and negative ionisation modes was used to get information on the structural features and the glycosylated forms of phenolic compounds. Identification of the 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 [11,22]

The basic structures of the compounds are summarised in Table 3.2.1, where the compounds are reported according to their retention time (RT). LC-ESI-FT-Orbitrap-MS/MS experiments were carried, selecting a sample for each geographical typology (see experimental section for detailed definition of the samples) in order to have a first LC-MS profile and then select and submit

specific major ions to fragmentation experiments using parameters previously chosen by explorative ESI-MS/MS direct infusion experiments realised on standard compounds. In addition to literature, these information, helped in identification of compounds.

Thus, 19 compounds were identified (Table 3.2.1): ten anthocyanin glycosides (**1–10**), five flavanol glycosides (**11–15**), and four flavonol (**16-19**) .

Qualitative LC-ESI-FT-Orbitrap-MS and Principal Component Analysis: untargeted and pseudo-targeted approach

The following step was to use the registered LC-ESI-FT-Orbitrap-MS data for multivariate data analysis. In order to carry on a comparative study of Myrtle berries extracts a multivariate approach was employed.

For untargeted Principal Component Analysis, LC-MS chromatograms were pre-processed using MZmine to compensate for 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, and columns representing the integrated and normalised peak areas.

Principal Component Analysis was performed by applying the peak areas of the total peaks present in the LC-MS dataset (excluding the noisy), thus a matrix was obtained by using these areas corresponding to *m/z* values (variables), and the column of the matrix were different analysed samples.

The resulted score scatter plot is reported in Fig.3.2.2 (LC-ESI-FT-Orbitrap-MS data obtained in negative analysis). The first component explains the 50% of variance while the second the 20%. The legend of classification is detailed as follow: **1** Plants obtained from seeds collected in Oristano Area; **2** Plants obtained from seeds collected in Sassari Area; **3** Plants obtained from seeds

collected in Olbia Area; **4** Plants obtained from seeds collected in Nuoro Area; **5** Plants obtained from seeds collected in Cagliari Area; **6** Plants obtained from seeds with a different provenience (V samples).
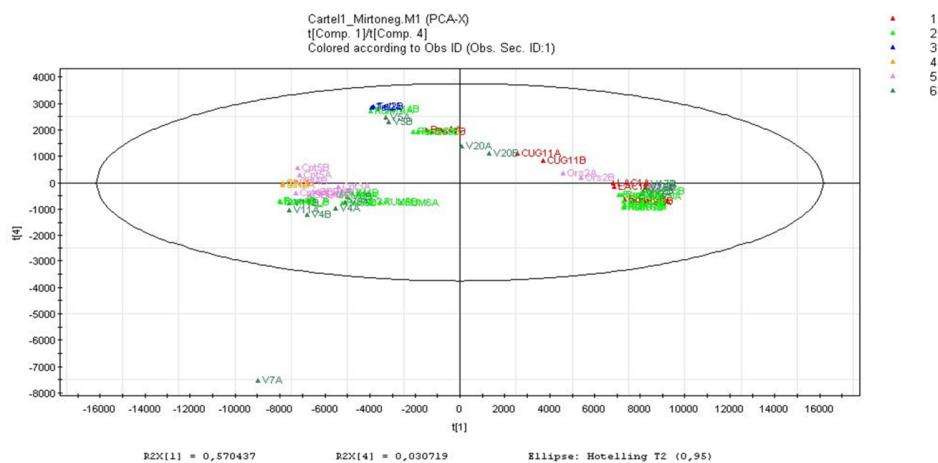

Fig. 3.2.2: Principal Component Analysis, Score Scatter Plot, untargeted LC-ESI-FT (Orbitrap)-MS

Another approach was realised for a pseudo-targeted analysis with the aim of reduce the number of variables and have more interpretable results. With this objective a data matrix was obtained by the data fusion of positive and negative ionization results, by using the peak areas of the compounds reported in Table 3.2.1 (LC-ESI-FT-Orbitrap-MS peaks corresponding to identified compounds), extracting them from the chromatograms both in negative LC-ESI-MS and in positive LC-ESI-FT-Orbitrap-MS and then organising a table with areas obtained for each sample.

The row of dataset represented the areas of each markers compounds (variables), both obtained in positive and negative mode ionization, and the

column represented the different samples (each in duplicate). The resulted score scatter plot is reported in Fig. 3.2.3 (Panel A) and the relative loading plot is in Fig. 3.2.3 (Panel B).
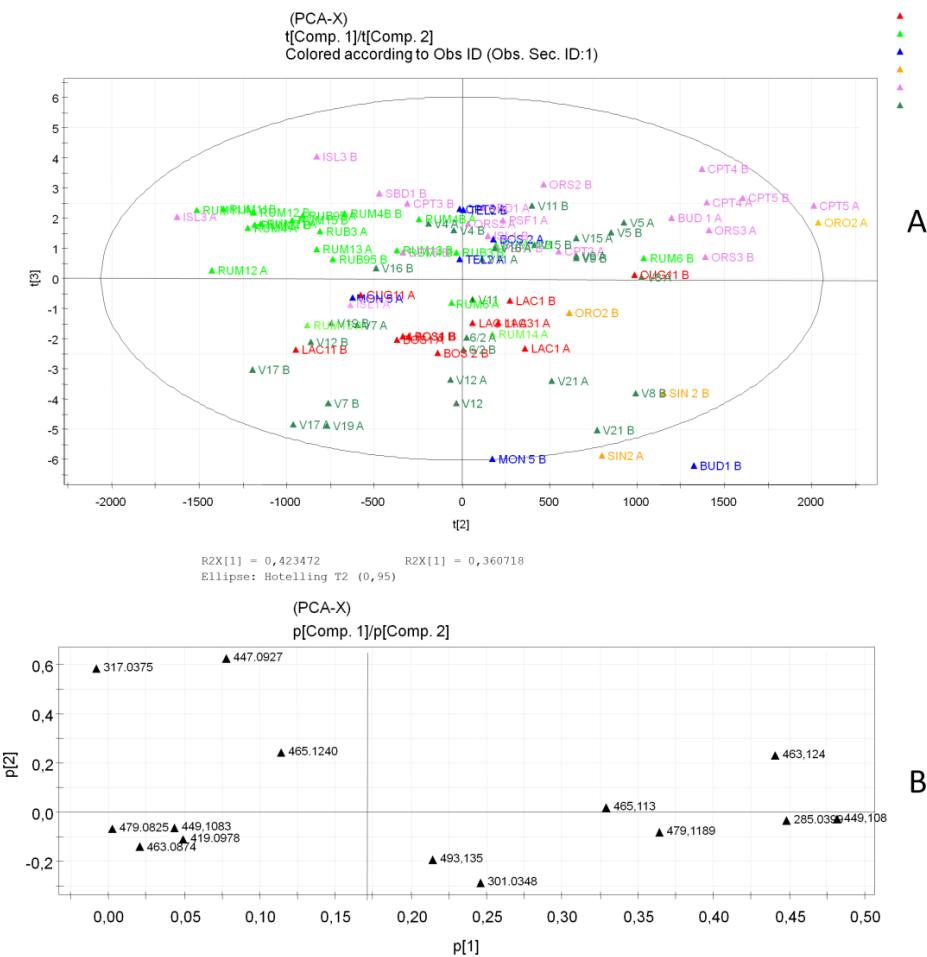

Fig. 3.2.3: Principal Component Analysis, pseudo-targeted LC-ESI-FT-(Orbitrap)-MS. score scatter plot (A), loading scatter plot (B)

Delphinidin-3-*O*-glucoside (*m/z* 447.09) and peonidin-3-*O*-glucoside (*m/z* 463.10), together with cyanidin-*O*-glucoside (*m/z* 449.10) among the

antocyans appears to be the most influent compounds in differentiating compounds relatively to the x axes.

Among flavonoids, myricetin (*m/z* 317.03) and its glycosides (*m/z* 463.08 and *m/z* 449.07) are influent in differentiating samples.

The separation of samples relatively to geographic origin of seeds is most appreciable on the y axes.

In fact **5** and **2** (from Cagliari and Sassari) samples are localised in the high area of the plot, founding their different localization based on y axis, samples **4** and **1** (Olbia and Oristano) on the lower square, characterized principally by specific markers, myricetin and its glycosides, as it is possible to confirm in the loading plot.

Other samples are heterogeneously distributed and not present a so clear classification; for them probably other factors are more influent in modify metabolism respect to the geographical area of origin of seeds.

Quantitative analysis by HPLC-DAD
Extraction Methods.

Prior to the analyses of 42 cultivars, in order to optimize the extracts' quality in association with the contents of the anthocyanins, extraction conditions were examined. The extraction efficacies were compared for three optimized and commonly deployed solvent systems (methanol, methanol/water 70:30 (v/v), and ethanol/water 70:30 (v/v)). The extraction of anthocyanins from myrtle berries was carried out at three levels of the initial solid–solvent ratio (1:10, 1:30, 1:60) and without or with ultrasound system. The mixture methanol/water 70:30 (v/v) in a solid–solvent ratio 1:60 with ultrasound assisted extraction (UAE) afforded the highest extraction yield. Only with the aqueous solution of 70% methanol, the matrix was exhausted after one cycle

of extraction. Therefore, all the analytical samples were prepared by employing this solvent extraction.

Quantitation of Anthocyanins.

All samples are collected at completely ripening in order to avoid the influence on the chemical composition of anthocyanins by maturation process. The pigment concentration of myrtle, in fact, is clearly affected by the ripening stage increasing with the ripeness of the fruits. Nine anthocyanins were identified in the *Myrtus communis* black berries extract by HPLC-DAD. Fig.3.2.4 reports an example of chromatogram of black berries extract of *Myrtus Communis* at 520 nm. All samples picked up from the "black" cultivars showed the same profile of anthocyanins with the presence of delfinidin-3-*O*-glucoside, cyanidin-3-*O*-glucoside, petunidin-3-*O*-glucoside, peonidin-3-*O*-glucoside, malvidin 3-*O*-glucoside, cyanidin 3-*O*-arabinoside, delfinidin 3-*O*-arabinoside, petunidin 3-*O*-arabinoside, and malvidin-3-*O*-arabinoside.
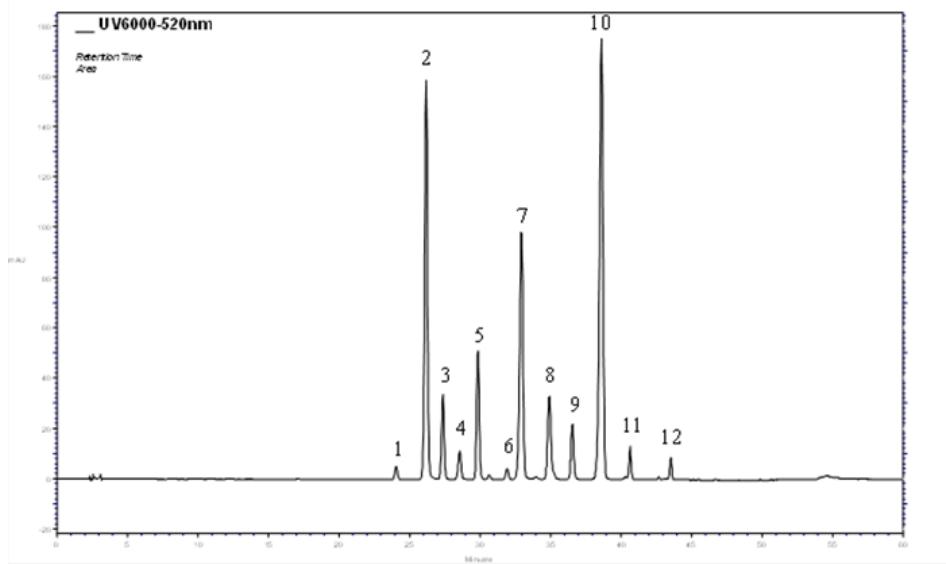
Fig. 3.2.4: HPLC-DAD chromatogram of black berries extract of *Myrtus communis* at 520 nm, where (1) unknown 1; (2) delfinidin-3-*O*-glucoside; (3) cyanidin-3-*O*-galactoside; (4) unknown 2; (5) cyanidin 3-*O*-glucoside; (6) cyanidin 3-*O*-arabinoside; (7) petunidin 3-*O*-glucoside; (8) delfinidin 3-*O*-arabinoside; (9) peonidin 3-*O*-glucoside; (10) malvidin 3-*O*-glucoside; (11) petunidin 3-*O*-arabinoside; (12) malvidin-3-*O*-arabinoside

From the analytical curve constructed by plotting area ratio between external and internal standard (cyanidin 3-*O*-galactoside) against the known concentration of each anthocyanin, the response of the UV detector at 520 nm was found to be linear in the range of 0.02–20 mg/L. The regression analysis providing the correlation coefficients between 0.9996 and 0.9999. Precision was evaluated through the determination of the repeatability and intermediary precision for peak retention times and areas of all anthocyanins quantified. In terms of repeatability the RSD values ranging between 1.17 % for cyanidin-3-*O*-glucoside and 4.27 % for malvidin-3-*O*-glucoside, while for intermediate precision the highest and the lowest coefficients of variation were 7.4 and 0.1%. The LOD, calculated at a signal-to-noise ratio of 3, and LOQ, determined with acceptable accuracy and precision at a signal-to-noise ratio of

92

10, were 0.01 and 0.02 mg/L, respectively. Each sample was analyzed three times. Table 3.2.2 shows the mean content with a standard deviation of the individual anthocyanins in the 42 cultivars expressed as milligrams of active ingredient per 100 g of dry weight (mg/100g_{dw}). As shown in Table 3.2.2, the total anthocyanin content of the black berries extracts ranged from 4658.9 to 21919.2 g/100$g_{dw.}$ The major compounds quantified in the myrtle berries extract were anthocyanins glycoside derivatives, mainly derivatives of malvidin (196.6 to 937.2 mg/100g_{dw}), delfinidin (54,3 to 486.6 mg/100g_{dw}), and petunidin (38.9 to 356.9 mg/100g_{dw}) representing more than 75% of anthocyanins fraction. Delfinidin-3-*O*-arabinoside is the most abundant anthocyanins between arabinose derivatives (12.3 to 105.7 mg/100g_{dw}), followed by petunidin-3-*O*-arabinoside (6.1 to 29.2 mg/100g_{dw}) and malvidin-3-*O*-arabinoside (3.6 to 22.2 mg/100g_{dw}).

93

Table 3.2.2: Anthocyanins concentrations in *Myrtus communis* berries extracts (mg/100g DW +SD, n=3)

	Unknown 1	Delfinidin 3-*O*-glu	Unknown 2	Cyanidin 3-*O*-glu	Unknown 3	Cyanidin 3-*O*-ara	Petunidin 3-*O*-glu	Delfinidin 3-*O*-ara	Peonidin 3-*O*-glu	Malvidin 3-*O*-glu	Petunidin 3-*O*-ara	Malvidin 3-*O*-ara	Total content
BOS1	5,9 ± 0,1	361,6 ± 14,2	7,4 ± 0,5	63,0 ± 1,0	2,9 ± 0,1	2,2 ± 0,1	274,8 ± 7,2	63,2 ± 3,8	27,0 ± 1,9	665,2 ± 36,6	20,0 ± 1,3	11,9 ± 0,5	15066,3
BOS2	11,8 ± 0,5	344,7 ± 20,4	15,45 ± 0,6	81,3 ± 40,	5,1 ± 0,4	5,2 ± 0,3	254,7 ± 13,0	67,0 ± 1,9	35,7 ± 1,1	612,8 ± 20,2	23,3 ± 01	17,3 ± 0,4	14783,0
BUD1	6,6 ± 0,2	223,5 ± 17,2	7,6 ± 0,2	38,3 ± 1,7	2,3 ± 0,0	6,8 ± 7,1	162,4 ± 10,5	45,3 ± 4,0	12,4 ± 1,2	360,6 ± 19,9	14,7 ± 1,1	8,9 ± 0,0	8892,5
CPT3	7,8 ± 2,3	330,5 ± 52,8	12,2 ± 2,7	60,3 ± 12,0	2,5 ± 0,6	3,0 ± 0,1	133,6 ± 18,2	55,4 ± 15,8	6,9 ± 3,9	234,8 ± 16,7	10,4 ± 2,1	6,4 ± 0,7	8654,3
CPT4	6,6 ± 0,6	351,5 ± 8,8	9,0 ± 0,5	38,6 ± 1,1	2,1 ± 0,0	2,1 ± 0,0	182,1 ± 13,1	43,1 ± 1,3	9,5 ± 0,2	206,5 ± 6,8	9,3 ± 0,2	4,1 ± 0,2	8644,9
CPT5	6,8 ± 0,5	272,5 ± 28,4	6,6 ± 0,1	44,8 ± 2,9	2,1 ± 0,1	2,1 ± 0,2	166,9 ± 0,9	42,6 ± 0,6	16,3 ± 1,1	320,6 ± 7,8	11,0 ± 0,3	6,3 ± 0,1	8986,9
CPT6	11,1 ± 0,2	348,5 ± 18,1	14,3± 0,9	60,0 ± 1,0	4,1 ± 0,1	4,2 ± 0,2	226,1 ± 14,0	70,5 ± 5,6	21,4 ± 2,2	497,8 ± 21,5	18,5 ± 1,1	12,2 ± 0,2	12906,1
CUG11	10,6 ± 0,1	366,5 ± 24,5	18,4 ± 1,6	143,0 ± 15,9	3,5 ± 0,3	6,9 ± 0,2	254,5 ± 18,6	89,2 ± 6,5	49,2 ± 3,5	598,5 ± 42,5	24,0 ± 1,9	13,7 ± 0,5	15807,2
ISL1	8,4 ± 0,8	323,2 ± 15,1	27,6 ± 1,1	82,7 ± 3,8	2,8 ± 0,2	5,6 ± 0,2	234,9 ± 17,4	83,8 ± 6,2	33,5 ± 2,1	490,6 ± 27,2	25,0 ± 1,0	11,1 ± 0,5	13318,9
ISL3	13,4 ± 0,3	406,7 ± 71,1	36,9 ± 5,7	87,9 ± 18,1	6,2 ± 0,2	7,1 ± 1,5	275,3 ± 49,3	105,7 ± 8,5	28,6 ± 7,0	471,5 ± 85,0	28,4 ± 4,4	9,7 ± 4,8	14809,7
LAC1	6,6 ± 0,5	272,1 ± 10,9	8,3 ± 0,8	60,2 ± 0,7	2,4 ± 0,8	2,0 ± 0,6	203,2 ± 6,2	49,3 ± 2,3	30,5 ± 1,8	576,4 ± 3,4	14,6 ± 0,5	9,9 ± 0,7	12373,2
LAC10	5,2 ± 0,1	219,7 ± 3,3	8,0 ± 0,6	59,3 ± 0,2	1,8 ± 0,1	3,3 ± 0,1	145,9 ± 2,6	42,5 ± 0,5	22,0 ± 1,7	331,2 ± 2,2	12,4 ± 0,5	6,5 ± 0,2	8598,4
LAC11	5,1 ± 0,5	229,1 ± 30,3	6,4 ± 0,4	62,2 ± 5,7	2,2 ± 0,1	2,5 ± 0,1	186,0 ± 18,7	42,5 ± 2,9	26,1 ± 3,7	465,7 ± 36,2	15,9 ± 0,7	9,8 ± 0,5	10534,8
LAC31	5,1 ± 0,1	164,8 ± 2,1	4,9 ± 0,2	41,5 ± 0,2	1,9 ± 0,0	2,9 ± 0,3	110,5 ± 5,8	37,1 ± 0,2	12,6 ± 0,8	304,8 ± 3,9	12,1 ± 0,8	8,6 ± 0,5	7073,2
MON5	9,8 ± 0,7	54,3 ± 6,4	1,9 ± 0,2	21,2 ± 1,4	4,0 ± 0,0	0,0 ± 0,0	38,9 ± 3,0	12,3 ± 1,1	23,1 ± 1,8	288,8 ± 9,8	6,1 ± 0,4	5,6 ± 0,2	4658,9
ORO2	13,0 ± 1,5	458,6 ± 14,9	9,2 ± 0,0	108,6 ± 3,4	2,3 ± 0,3	2,8 ± 0,5	311,1 ± 7,8	89,0 ± 3,0	39,6 ± 1,5	658,6 ± 7,7	22,9 ± 1,3	11,4 ± 0,4	17270,8
ORS3	5,4 ± 0,2	189,2 ± 9,9	3,5± 0,1	65,4 ± 3,4	1,7 ± 0,1	1,9 ± 0,0	120,1 ± 5,2	16,6 ± 0,0	27,5 ± 1,5	244,4 ± 9,3	7,5 ± 0,0	3,6 ± 0,0	6864,4
ORS2	5,7 ± 1,4	254,3 ± 29,6	6,3 ± 0,2	86,2 ± 2,0	2,2 ± 0,0	3,3 ± 0,1	144,8 ± 18,4	36,1 ± 5,3	30,9 ± 0,6	352,4 ± 25,6	12,6 ± 0,6	7,3 ± 0,5	9420,1
PSF1	7,8 ± 0,1	306,1 ± 12,3	11,8 ± 0,4	86,4 ± 3,1	2,8 ± 0,1	5,6 ± 0,2	194,6 ± 8,8	57,6 ± 2,8	24,5 ± 1,7	385,6 ± 19,3	14,0 ± 0,4	6,4 ± 0,2	11033,8
RUM3	6,7 ± 0,4	259,0 ± 21,7	20,9 ± 0,7	111,3 ± 2,1	2,6 ± 0,0	8,1 ± 0,3	191,5 ± 2,7	71,9 ± 2,1	70,2 ± 1,4	597,9 ± 11,4	25,5 ± 0,2	15,3 ± 0,7	13869,6
RUM4	5,5 ± 0,1	186,2 ± 11,6	15,5 ± 0,7	96,5 ± 6,6	2,2 ± 0,1	7,0 ± 0,2	152,6 ± 9,6	56,2 ± 2,9	71,6 ± 4,0	562,9 ± 32,8	19,9 ± 0,7	14,9 ± 1,0	11940,9
RUM4B	7,6 ± 0,4	257,9 ± 22,4	11,3± 0,4	88,7 ± 3,5	3,2 ± 0,2	4,9 ± 0,1	189,1 ± 13,3	56,4 ± 1,8	41,6 ± 2,5	519,7 ± 24,5	17,8 ± 0,1	11,6 ± 0,4	12116,8
RUM10	7,0 ± 0,1	441,4 ± 31,1	7,5 ± 0,1	86,7 ± 2,1	2,7 ± 0,1	2,2 ± 0,1	302,2 ± 9,2	64,0 ± 1,5	24,6 ± 1,5	430,9 ± 26,5	16,1 ± 1,1	7,1 ± 0,0	13922,7
RUM12	7,8 ± 0,4	243,8 ± 17,8	17,6 ± 1,3	77,2 ± 5,2	3,2 ± 0,3	6,7 ± 0,5	176,3 ± 13,5	58,1 ± 4,6	32,0 ± 2,8	414,1 ± 33,2	16,6 ± 2,8	10,6 ± 1,1	10679,6
RUM13	13,9 ± 1,3	486,6 ± 34,0	18,9 ± 0,0	144,6 ± 3,1	5,8 ± 0,4	8,3 ± 0,2	356,9 ± 9,0	83,6 ± 2,3	78,4 ± 5,2	937,2 ± 41,5	29,2 ± 2,4	22,2 ± 0,8	21919,2
RUM15	5,0 ± 0,2	91,5 ± 2,7	8,3 ± 0,6	48,4 ± 0,1	3,2 ± 1,1	4,0 ± 0,2	760,0 ± 5,4	28,9 ± 2,2	38,2 ± 2,3	386,0 ± 25,8	12,2 ± 1,3	13,4 ± 0,1	7169,8
RUM20	6,6 ± 0,5	328,9 ± 9,9	7,5 ± 0,2	96,9 ± 3,1	2,5 ± 0,2	3,4 ± 0,1	230,2 ± 7,0	52,6 ± 1,7	37,5 ± 1,0	511,7 ± 15,4	15,9 ± 0,6	8,5 ± 0,3	13051,5
SBD1	5,9 ± 0,7	229,8 ± 17,4	6,8 ± 0,1	40,9 ± 0,7	2,3 ± 0,1	2,1 ± 0,0	153,9 ± 0,5	44,8 ± 0,2	16,2 ± 0,6	399,2 ± 27,4	14,3 ± 0,9	8,4 ± 0,5	9245,7
V4	7,3 ± 0,2	232,8 ± 18,7	10,0 ± 0,3	80,0 ± 5,2	2,8 ± 0,0	5,5 ± 0,1	157,3 ± 9,7	56,8 ± 2,7	27,3 ± 0,5	311,7 ± 15,4	15,8 ± 1,0	9,3 ± 0,8	9185,1
V5	5,0 ± 0,3	166,0 ± 3,7	9,0 ± 0,3	39,0 ± 3,8	2,0 ± 0,1	3,5 ± 0,5	101,5 ± 0,9	45,0 ± 0,5	6,7 ± 0,6	196,5 ± 3,0	11,5 ± 0,0	6,6 ± 0,0	5936,3
V7	7,4 ± 0,2	243,6 ± 5,5	9,8 ± 0,1	56,8 ± 1,3	3,4 ± 0,1	3,6 ± 0,1	175,7 ± 4,5	48,7 ± 0,8	27,7 ± 0,5	468,3 ± 12,2	17,6 ± 0,2	12,4 ± 0,6	10785,4
V9	9,8 ± 0,8	353,1 ± 1,1	12,8 ± 0,2	84,7 ± 0,0	3,1 ± 0,1	4,8 ± 0,1	218,8 ± 2,3	73,1 ± 0,7	24,2 ± 0,6	364,3 ± 9,4	19,5 ± 0,0	9,2 ± 0,5	11804,8
V11	9,5 ± 0,1	273,3 ± 5,4	15,3 ± 0,0	50,4 ± 1,1	4,0 ± 0,1	4,5 ± 0,2	196,0 ± 5,7	71,8 ± 1,6	20,1 ± 1,4	424,2 ± 22,0	22,5 ± 0,4	13,7 ± 0,9	11071,9
V12	11,5 ± 0,1	336,4 ± 0,7	17,5 ± 0,2	79,6 ± 1,1	4,3 ± 0,1	6,9 ± 0,1	238,3 ± 0,9	88,1 ± 0,2	29,6 ± 0,4	505,1 ± 1,4	25,8 ± 0,6	12,9 ± 0,1	13559,6
V15	6,5 ± 0,5	174,1 ± 2,9	13,3 ± 0,0	35,4 ± 0,7	2,9 ± 0,1	3,5 ± 0,1	127,9 ± 1,9	40,9 ± 0,6	12,7 ± 0,7	293,4 ± 28,0	13,6 ± 1,2	9,8 ± 0,8	7376,9
V16	7,0 ± 0,3	195,1 ± 11,0	14,2 ± 0,6	35,0 ± 2,1	3,3 ± 0,2	3,3 ± 0,2	146,0 ± 8,7	44,7 ± 2,5	14,8 ± 0,2	342,4 ± 22,1	15,1 ± 0,8	10,9 ± 0,7	8337,1
V17	4,6 ± 0,6	99,2 ± 0,3	5,1 ± 0,2	21,2 ± 1,1	2,3 ± 0,1	2,0 ± 0,1	76,1 ± 3,7	33,5 ± 2,1	5,8 ± 0,5	239,1 ± 14,4	13,5 ± 1,0	9,6 ± 0,9	5118,3
V19	11,8 ± 0,6	309,8 ± 18,8	15,3 ± 0,2	43,8 ± 1,7	3,8 ± 0,1	3,8 ± 0,2	222,7 ± 14,8	87,5 ± 6,0	17,0 ± 0,5	501,0 ± 8,2	25,4 ± 1,0	15,1 ± 0,6	12570,6
V20	6,7 ± 0,3	166,8 ± 49,5	7,4 ± 0,0	52,4 ± 13,6	2,6 ± 0,1	3,3 ± 0,1	124,6 ± 32,7	55,3 ± 4,0	30,7 ± 8,4	451,4 ± 80,2	17,8 ± 2,8	9,4 ± 1,3	9304,5
6/2	9,2 ± 0,1	408,0 ± 19,0	10,6 ± 0,7	84,4 ± 2,8	3,0 ± 0,1	3,3 ± 0,2	296,4 ± 21,1	65,3 ± 3,1	30,0 ± 0,5	582,1 ± 12,1	20,2 ± 1,0	9,2 ± 0,1	15216,6

1 Plants obtained from seeds collected in Oristano Area (LAC, BOS and CUG samples); **2** Plants obtained from seeds collected in Sassari Area (RUM samples); **3** Plants obtained from seeds collected in Olbia Area (MON and BUD samples); **4** Plants obtained from seeds collected in Nuoro Area (ORO samples); **5** Plants obtained from seeds collected in Cagliari Area (CPT, PFS, ORS, ISL and SBD samples); **6** Plants obtained from seeds with a different provenience (V samples)

Although no qualitative differences were observed in the anthocyanin fraction between cultivars, quantitative differences were observed between anthocyanin compounds. Therefore, despite the differentiation of cultivars was possible only at the quantitative level, samples coming from different areas of

Sardinia don't seem to have significant differences on glucose and arabinose derivatives total content.

3.2.3 Conclusion

Metabolomics approaches revealed useful for the assessment of geographical origin of original seeds, and mainly for distinguish the areas of Sassari and Cagliari, although plants were cultivated in the same area. Untargeted metabolomics approach realized with an LC-MS platform, treating data by MZ-mine, by using all the peaks with the exception of the noise appears not valid in the discrimination of samples when observed in the PCA model. A pseudo-targeted approach realized with an LC-MS platform and data obtained by a selected peak list, was found useful to our goals.

Anthocyans were identified as principal chemical markers of the geographical origin of the plants, and they were quantified in different samples.

The results of the quantitative analysis of anthocyanins in 42 "black" cultivars of Sardinian *Myrtus communis*, indicate that with a such small number of samples would seem not reliable the differentiation of myrtle cultivars. Therefore, in order to correlate the quantitative variability of the single anthocyanins in samples analyzed with a different geographical origin and genetic factors, it would be desirable work with a more representative number of samples. It is not possible to confirm that these differences are necessarily geo-climatic, since they could also be caused by genetic and/or environmental factors.

3.2.4 Materials and methods

Reagents and solvents

Standards of cyanidin-3-*O*-glucoside, cyanidin-3-*O*-galacatoside, petunidin-3-*O*-glucoside, peonidin-3-*O*-glucoside, malvidin-3-*O*-glucoside, delphinidin-3-*O*-glucoside, myricetin-3-*O*-galactoside, myricetin-3-*O*-rhamnoside, quercetin-3-*O*-glucoside, quercetin-3-*O*-rhamnoside, myricetin, quercetin, kaempferol, ellagic acid, were purchased from Extrasynthese (Lyon, France). Water 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). Orthophosphoric acid were purchased from Carlo Erba (ACS ISO, for analysis, 85%).

Plant material

Seeds of different varieties of *Myrtus communis* were collected from different geographic areas of Sardinia (to obtain a total number of 42 different cultivars) and then grown in the Experimental Station of the University of Sassari in Fenosu (Oristano, Sardinia, Italy) in the same cultivation conditions.

Samples were classified in 6 groups:

1. Plants obtained from seeds collected in Oristano Area **2.** Plants obtained from seeds collected in Sassari Area **3.** Plants obtained from seeds collected in Olbia Area **4.** Plants obtained from seeds collected in Nuoro Area **5.** Plants obtained from seeds collected in Cagliari Area **6.** Plants obtained from seeds with a different provenience.

Black berries (≈ 2 kg) were randomly harvested during fruit ripening stage in 2014, December. Before analysis, samples collected were lyophilized and then stored at -18°C, to preserve the content of phenolic compounds of berries [23,24].

Extraction

Lyophilized fruits were extracted with a mixture of methanol in water (70:30, v/v), by using 5 mL of solvent each 1g of plant material. Extraction assisted by ultra sounds was carried out for an hour, then extracts obtained were centrifuged at 4000 rpm. Samples were filtered by using Filter devices 0.2 μmPTFE membrane, diluted 1:10 and 10 μl were injected in the LC-ESI-FT-Orbitrap-MS system.

MS^n standard analysis

MS and MS^n analysis on standard compounds were carried out on linear Trap-Orbitap hybrid mass spectrometer (LTQ-Orbitrap XL, Thermo Fisher Scientific, Bremen, Germany) equipped with electrospray ionization (ESI). Xcalibur software version 2.1 was used for data acquisition and data analysis. A syringe pump was used for the direct infusions of reference compound solutions, at a flow rate set at 5 μL/min in a HPLC flow (500 μL/min) composed of acetonitrile and milliQ (18.2 MΩ) water with 0.1% of formic acid. Standard solutions were prepared at a concentration of 0.1 mg/mL. For MS/MS experiments, CID was carried out using Helium as collision gas. The collision energy are reported as a percentage of the maximum five Vp–p normalized for the parent ion *m/z* (NCE: Normalized Collision Energy) [25]. Collision energy for CID was optimized between 25% and 46% of maximum, and the isolation width of precursor ions was 1.5 amu.

Negative ion mode: ESI source parameters were as follows: capillary voltage -35V; tube lens voltage -126.47; capillary temperature 280°C; Shealt and Auxiliary Gas flow (N_2) 20, Sweep gas 0, Spray voltage 5. MS spectra were acquired by full range acquisition covering *m/z* 200-1600. Resolution of FT-MS was 30000.

97

Positive ion mode: ESI source parameters were as follows: capillary voltage +35V; tube lens voltage +126.47; capillary temperature 280°C; Shealt and Auxiliary Gas flow (N_2) 20, Sweep gas 0, Spray voltage 5. MS spectra were acquired by full range acquisition covering *m/z* 200-1600. Resolution of FT-MS was 30000.

LC-ESI-FT-Orbitrap-MS and LC-ESI-FT-Orbitrap-MS/MS

All experiments were performed using a Thermo scientific liquid chromatography system constituted of a quaternary Accela 600 pump and an Accela autosampler, connected to a linear Trap-Orbitap hybrid mass spectrometer (LTQ-Orbitrap XL, Thermo Fisher Scientific, Bremen, Germany) with electrospray ionization (ESI). Separation were performed on a Synergy 4U Fusion RPC18 (Phenomenex).

The mobile phase consisted of A (2% formic acid in water) and B (2% formic acid in acetonitrile/water 1:1). 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, 20% (B) to 40%; then at 35 min 60% (B) and then after 5 minutes 100% (B) for 5 min. Then the column was re equilibrated for the next 7 min. MS spectra were acquired by full range acquisition covering *m/z* 150-1600. For fragmentation study, a Data Dependent scan was performed, selecting precursor ions corresponding to most intensive ions in LC-MS analysis. Collision energy was selected to 35%. Phenolic compounds were identified according to the corresponding spectral characteristics fragmentation and characteristic retention time, by comparison with data from literature. Xcalibur software version 2.1 was used for instrument control, data acquisition and data analysis. Each sample was investigated in positive and in negative ion mode.

Negative ion mode: ESI source parameters were as follows: capillary voltage - 35V; tube lens voltage -126.47; capillary temperature 280°C; Shealt and Auxiliary Gas flow (N_2), 20 Sweep gas 0 Spray voltage 5. MS spectra were acquired by full range acquisition covering m/z 200-1600. Resulution of FT-MS was 30000.

Positive ion mode: ESI source parameters were as follows: capillary voltage +35V; tube lens voltage +126.47; capillary temperature 280°C; Shealt and Auxiliary Gas flow (N_2), 20 Sweep gas 0 Spray voltage 5. MS spectra were acquired by full range acquisition covering m/z 200-1600. Resulution of FT-MS was 30000.

Each sample was analysed in duplicate for each modality in order to have technical duplicates for statistical analysis and multivariate data analysis.

Multivariate data analysis

In an untargeted approach, base peak chromatograms deriving from LC-ESI-FT-Orbitrap-MS analysis (only negative ion mode) were evaluated using a platform independent open source software package, MZmine (http://mzmine.sourceforge.net/). Using this toolbox with normalization by total raw signal, we detected 454 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 using PCA, principal component analysis. PCA was performed by applying the peak area obtained from LC/MS analysis [26-28]. Pareto scaling was applied before multivariate data analysis.

For targeted approach a data matrix was obtained by the data fusion of positive and negative ionization results, by using the peak areas of the compounds reported in table 3.2.1, extracting the chromatogram or in negative LC-ESI-

FT-Orbitrap-MS or in positive LC-ESI-FT-Orbitrap-MS and then organising a table with areas obtained by integration of peaks for each sample from the different row data. The row of the tabular dataset represented the area of each markers compounds (variables), both acquired in positive and negative mode ionization, and the column represented different samples. Further analysis of the data matrix were performed by SIMCA P+ software 12.0 (Umetrix AB, Umea Sweden) by using PCA, principal component analysis. Pareto scaling was applied before multivariate data analysis.

Quantitative analysis

Preparation of plant extracts

Dried berries were thinly pulverized in a home style coffee grinder, mixed thoroughly and split up in three replicates. Two hundred milligrams of every powdered sample were added to 12 mL of a mixture of methanol and distilled water without acidification at a ratio of 70:30 (v/v) in screw capped 15 mL Falcon tubes. The extraction was conducted in an ultrasonic bath, where the temperature was maintained constant by a circulation of water in an external jacket connected to a thermostat. The extract was centrifuged at 4000 rpm for 60 minutes at 20°C and diluted 1:10 with 0.22 M phosphoric acid in water containing cyanidin 3-*O*-galactoside as internal standard (IS) prior to its injection into the chromatographic system.

Stock Standard Solutions

A stock standard solution of each anthocyanin (1000 mg/L for all compounds) were prepared in methanol by weighing approximately 0.01 g of the analyte into a 10 mL volumetric flask and diluting to volume. An intermediary standard solution was prepared by diluting stock standard solution with the

methanol and subsequently diluted with 0.22 M phosphoric acid to obtain mixed reference solutions in the range of 0.02–20 mg/L. All standard solutions were stored in the dark at -20 °C until usage. To each reference standard solution was added an appropriate amount of IS (cyanidin 3-galactoside) to give a final concentration of 1 mg/mL.

Calibration curves

An internal standard method was utilized to construct the analytical curve for all anthocyanins measured at a wavelength of 520 nm. Calibration curves were prepared by injecting the reference standard solutions in triplicate and plotting the ratio of the peak areas of the external standard (at each concentration) to those of the IS against the corresponding standard concentration. Arabinoside anthocyanins were quantified as cyanidin-3-*O*-arabinoside and petunidin 3-*O*-glucoside was quantified using the calibration curves of delphinidin-3-*O*-glucoside. For each calibration curve the regression equation was determined. Both the precision under conditions of repeatability and intermediate precisions were determined, by performing either six injections of 0.02, 0.1, 1 and 5 mg/L standards in the same day or six injections of the same standards in different days, respectively. The limit of detection (LOD) and limit of quantification (LOQ) were calculated on the basis of the data of the regression of the analytical curves.

HPLC Analysis

A HPLC 1100 system (Agilent Technologies, Milan, Italy) was equipped with a quaternary pump, a degasser, an autosampler, a thermostated column compartment, and coupled with a DAD detector UV 6000 (Thermo Finnigan, Milan, Italy). The chromatographic separation was performed on a Kinetex

column (5u, C18, 100 A, Phenomenex), operated with mobile phase A (acetonitrile) and B (H_2O containing 0.22M phosphoric acid). The gradient program was as follows: from 0 min (A:B 5:95, v/v) to 30 min (A:B 13:87, v/v), from 30 to 35 min (A:B 15:85, v/v,) and finally from 35 to 50 min (A:B 30:70, v/v). A post-time of 15 minutes was used to allow the column to equilibrate before the next sample injections. The flow rate was 0.6 mL/min and the injection volume was 10 μL. Quantitative analysis of the individual anthocyanins was conducted at 520 nm by using an internal standard method. The contents of the anthocyanins were expressed as milligrams of active ingredient per 100 g of dry weight. All analyses were replicated three times. Data were expressed as the mean ± standard deviation (SD).

Acknowledgements

3.2.5 References

[1]Bravo L. Polyphenols: chemistry, dietary sources, metabolism, and nutritional significance. *Nutrition Reviews*, **1998** , *56*, 317–333

[2]Fournier P. In: *Catalogue des Plante Medicinales et Vénéneuses de France*, Vol III, edited by Lechevalier Paris, **1947**, pp.14-48

[3]Taamalli A., Iswaldi I., Arraez-Roman D., Segura-Carretero A., Fernandez-Gutierrez A., Zarrouk M. UPLC-QTOF/MS for a Rapid Characterisation of Phenolic Compounds from Leaves of *Myrtus communis* L. *Phytochemical Analysis* **2014**, *25(1)*, 89-96

[4]Mulas M., Melis R. A. M. Essential Oil Composition of Myrtle (*Myrtus communis*) Leaves. *Journal of Herbs, Spices & Medicinal Plants* **2011**, *17(1)*, 21-34

[5]Martin T., Villaescusa L., De Sotto M., Lucia A., Diaz A.M. Determination of anthocyanin pigments in *Myrtus communis* berries. *Fitoterapia* **1990**, *61*, 85

[6]Diaz A.M., Ageger A. Analytical method of tannins and their application to *Myrtus communis* seeds extract. *Fitoterapia* **1987**, *58*, pp. 395–407

[7]Martin T., Rubio B., Villaescusa L., Fernandez L., Diaz A.M. Polyphenolic compounds from pericarps of *Myrtus communis*. *Pharmaceutical Biology* **1999**, *37* , 28-31

[8]Romani A., Pinelli P., Mulinacci N., Vincieri F.F., Tattini M. Identification and quantification of polyphenol in leaves of *Myrtus communis*. *Chromatographia* **1999**, *49*, 17-20

[9]Romani A., Coinu R., Carta S., Pinelli P., Galardi C., Vincieri F.F., et al. Evaluation of antioxidant effect of different extract of *Myrtus communis* L. *Free Radical Research* **2004,** *38*, 97-103

[10]Rosa A., Deiana M., Casu M., Corona G., Appendino G., Bianchi F., et al. Antioxidant activity of oligomeric acylphoroglucinols from Myrtus communis L. *Free Radical Research* **2003**, *37*, 1013-1019

[11]Montoro P., Tuberoso C.I., 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. *Journal of Chromatography A* **2006**, *1112(1-2),* 232-240

[12]Montoro P., Tuberoso C. I. G., 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 the preparation of myrtle liqueur. *Journal of Pharmaceutical. Biomedical Analysis* **2006**, *41(5),* 1614-1619

[13]Fiehn O. Metabolomics - the link between genotypes and phenotypes. *Plant Molecular Biology* **2002**, *48*, 155–71

[14]Scalbert A., Brennan L., Fiehn O., Hankemeier T., Kristal B. S., van Ommen B., Pujos-Guillot E., Verheij E., Wishart D., and Wopereis S. Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics* **2009**, *5(4)*, 435–458

[15]Guijarro-Díez M., Nozal L., Marina M.L., Crego A.L. Metabolomic fingerprinting of saffron by LC/MS: novel authenticity markers. *Analytical and Bioanalytical Chemistry* **2015**, *407(23)*, 7197-213

[16] Ernst M., Silva D. B., Silva R. R., Vêncio R. Z. N., Peporine Lopes N. Mass spectrometry in plant metabolomics strategies: from analytical platforms to dataacquisition and processing. *Natural Product Reports* **2014**, 31, 784–806

[17]Turi C. E., Finley J., Shipley P. R., Murch S. J., Brown P. N. Metabolomics for Phytochemical Discovery: Development of Statistical Approaches Using a Cranberry Model System. *Journal of Natural Products* **2015**, *78 (4)*, 953–966

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

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

[20]Brown P. N., Murch S. J., Shipley P. Phytochemical diversity of cranberry (*Vaccinium macrocarpon* Aiton) cultivars by anthocyanin determination and

metabolomic profiling with chemometric analysis. *Journal of Agriculture and Food Chemistry* **2012**, *60*, 261–271

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

[22]Barboni T., Venturini N., Paolini J., Desjobert J.M., Chiaramonti N., Costa J. Characterisation of volatiles and polyphenols for quality assessment of alcoholic beverages prepared from Corsican *Myrtus communis* berries. *Food Chemistry* **2010**, *122*, 1304-1312

[23]Mullen W. , Stewart A. J., Lean M. E. J., Gardner P., Duthie G. G., Crozier A. Effect of freezing and storage on the phenolics, ellagitannins, flavonoids and antioxidant capacity of red raspberries. *Journal of Agriculture and Food Chem*istry **2002**, 50, 5197-5201

[24]De Ancos B., Gonzàlez E. M., Cano M. P.. Ellagic acid, vitamin C, and total phenolic contents and radical scavenging capacity effected by freezing and frozen storage in raspberry fruit. *Journal of Agriculture and Food Chem*istry **2000**, 48, 4565-4570

[25]Lopez L.L., Tiller P.R., Senko M.W., Schwartz J.C. Automated strategies for obtaining standardized collisionally induced dissociation spectra on a benchtop ion trap mass spectrometer *Rapid Communication in Mass Spectrometry* **1999**, 13, 663–668

[26]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. *Journal of Pharmaceutical and Biomedical analysis* **2012**, *70*, 224-230

[27]Soufi S., D'Urso G., Pizza C., Rezgui S., Bettaieb T., Montoro P. Steviol glycosides targeted analysis in leaves of *Stevia rebaudiana* (Bertoni) from plants cultivated under chilling stress conditions. *Food Chemistry* **2016**, 190, 572-580

[28]Safer S., Serhat C. 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, 1379–1389

3.3 Metabolic Profiling of different parts of myrtle's berry from Sardinia

3.3.1 Introduction

Polyphenols like anthocyanins, ellagitannins, flavonoids conjugates and hydroxycinnamic acid are the most abundant source of health promoting phytochemical compounds in berry fruit[1,2]. Ellagitannins, in particular, belong 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 relatively uncommon in fruit and vegetables in our diet, and are found only in a few berry fruit such as strawberries (*Fragaria ananassa* D.) raspberries (*Rubus idaeus* L.) and blackberries (*Rubus* sp)[4]. For this class of compounds recent studies have documented several biological properties which make them suitable not only for use in textiles/dyeing, but also for other applications in cosmetics, medicine, agronomy and phytotherapy[5,6].

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

Myrtus communis belongs to the Myrtaceae family, is a pleasant annual shrub with dark blue ripe berries, which have a long history of application in the perfumery, cosmetic, food and pharmaceutical industries[8]. In Sardinia (Italy), these berries are mostly employed for the industrial formulation of sweet liqueur[9]. The species *Myrtus communis* has been subject of several researches focused on the chemical composition of the essential oil[10,11] and the methanol extracts of myrtle berries, investigated for the presence of flavonoids and anthocyanins[12-14]. Nevertheless, no data are present on metabolite composition and distribution in the different parts of the berry.

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

Metabolite profiling using liquid chromatography combined with high resolution mass spectrometry (LC-ESI-FT-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-FT-Orbitrap-MS analysis and MS/MS experiments, 35 compounds were identified or tentatively characterized on the bases of their retention time, UV/Vis absorbance, MS spectra and MS fragmentation patterns and a new class of hydrolyzable tannins was identified for the first time in these berries.

3.3.2 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 anthocyanins. Results (data not showed) indicated that both the classes of compounds were present in the different berry parts. Therefore, LC-MS experiments were performed in positive and negative ion mode.

LC-MS metabolomic analysis

To identify the major metabolites presents in seeds, peel&pupl and whole myrtle berry, LC-MS analysis were performed on the aqueous/methanol extracts. Base-peak chromatograms of extracts in negative ion mode of the

different parts of myrtle berries are presented in Fig. 3.3.1, while base-peak chromatograms in positive ion mode are presented in figure 3.3.2.
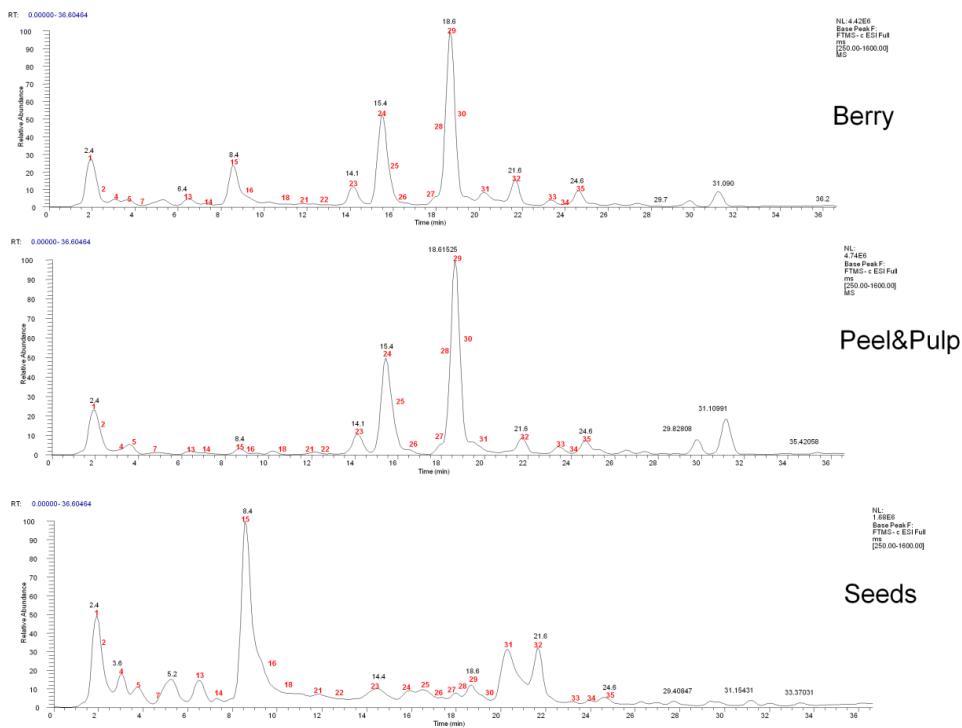

Fig. 3.3.1: Base peack chromatograms in negative ion mode of whole berry, Peel&pulp and seeds of *Myrtus communis* extracts
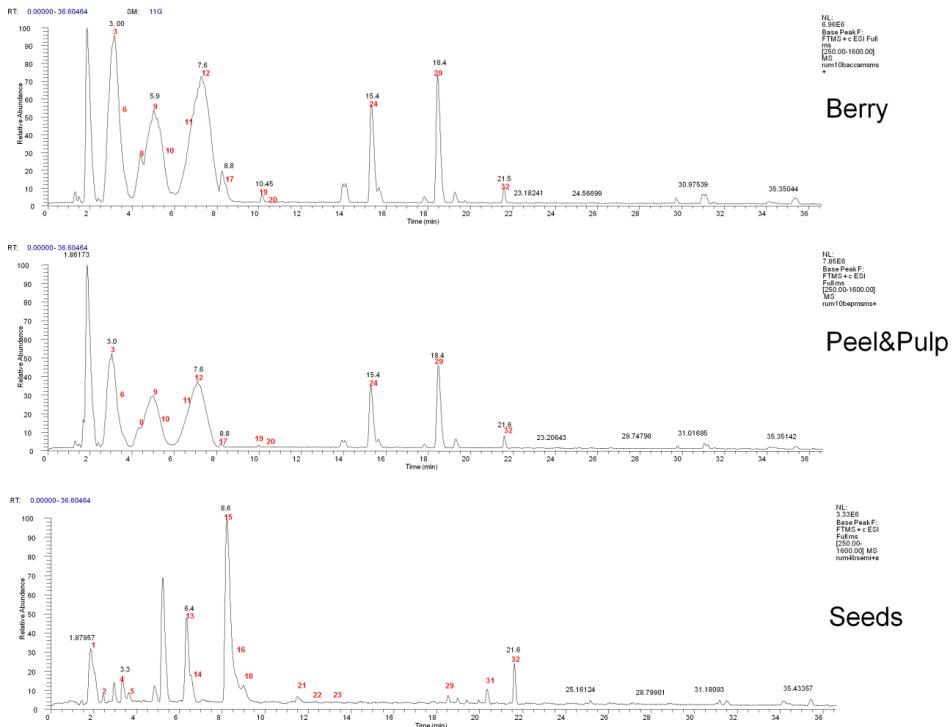
Fig. 3.3.2: Base peak chromatograms in positive ion mode of whole berry, Peel&pulp and seeds of *Myrtus communis* extracts

The compounds characterized are summarized in Table 3.3.1. In total 35 compounds were identified or tentatively identified based on retention time, accurate mass measurement, fragmentation pattern and comparison with data reported in literature. As far as we know 19 of these were identified for the first time in myrtle berries. In the present study, the compounds characterized are classified into seven groups: hydrolyzed tannins; hydroxycinnamic acid; gallomyrtucommulones; flavanols; flavonols; ellagic acid and derivatives; anthocyanins.

110

Table 3.3.1: 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	Ref.
Hydrolyzable Tannins						
1 HHDP-hexoside	$C_{20}H_{18}O_{14}$	482.0696	481.0617	301/275	2.4	16*
2 monogalloyl-hexoside	$C_{13}H_{16}O_{10}$	332.0743	331.0664	313/271/169	2.6	15*
4 strictinin (galloyl-HHDP hexoside)	$C_{34}H_{22}O_{22}$	634.0806	633.0726	301/421/615	3.6	15*
5 galloylquinic acid	$C_{14}H_{16}O_{10}$	344.0743	343.0664	191/169	3.6	15*
13 tellimagrandin I	$C_{34}H_{26}O_{22}$	786.0915	785.0836	301/483/633	6.4	17
14 punicalin	$C_{34}H_{22}O_{22}$	782.0602	781.0523	N.F.	8.1	16*
15 pedunculagin (Bis HHDP hexoside)	$C_{34}H_{24}O_{22}$	784.0759	783.0679	481/301	8.4	17
16 casuarictin	$C_{41}H_{28}O_{26}$	936.0868	935.0789	783/633/301	8.7	19
25 castalagin	$C_{41}H_{26}O_{26}$	934.0712	933.0632	N.F.	15.6	20
26 tellimagrandin II (Trigalloyl HHDP hexoside)	$C_{41}H_{30}O_{26}$	938.1025	937.0945	N.F.	17.2	17
Gallomyrtucommulones						
35 gallomyrtucommulone C	$C_{27}H_{36}O_{13}$	568.2155	567.2075	331/313/271/169	24.6	15*
Hydroxycinnamic acids						
18 caffeoylhexose	$C_{15}H_{18}O_9$	342.0950	341.0871	161/179	9.8	21
Flavanols						
7 epigallocatechin	$C_{15}H_{14}O_7$	306.0739	305.0660	174/270	4.3	22
22 catechin/epicatechin	$C_{15}H_{14}O_6$	290.0790	289.0711	245/179	13.3	22
Flavonols						
23 myricetin galloylhexoside	$C_{28}H_{24}O_{17}$	632.1013	631.0934	479/317	13.9	15*
24 myricetin hexoside	$C_{21}H_{20}O_{13}$	480.0903	479.0824	317	15.4	24
27 myricetin pentoside	$C_{20}H_{18}O_{12}$	450.0798	449.0718	317	17.9	24
28 quercetin galloylhexoside	$C_{28}H_{24}O_{16}$	616.1064	615.0985	463/301	18.2	23*
29 quercetin hexoside	$C_{21}H_{20}O_{12}$	464.0954	463.0875	301	18.6	24
30 myricetin deoxyhexoside	$C_{21}H_{20}O_{12}$	464.0954	463.0875	317	18.7	24
32 quercetin deoxyhexoside	$C_{21}H_{20}O_{11}$	448.1005	447.0926	301	21.6	24
33 myricetin galloyl deoxyhexose	$C_{28}H_{24}O_{16}$	616.1064	615.0984	317/463	23.2	15*
34 myricetin hexose deoxyhexose	$C_{27}H_{30}O_{17}$	626.1483	625.1403	479/317	23.3	23*

Ellagic acid and derivates

21 ellagic acid hexoside	$C_{20}H_{16}O_{13}$	464.0591	463.0511	301	12.3	15*
31 ellagic acid	$C_{14}H_6O_8$	302.0063	300.9983	284/256/ 185	20.2	22

Anthocyanins

			$[M]^+$			
3 delphinidin hexoside	$C_{21}H_{21}O_{12}^+$	465.1033	465.1034	303	3.0	24
6 cyanidin hexoside	$C_{21}H_{21}O_{11}^+$	449.1083	449.1083	287	4.2	24
8 petunidin hexoside	$C_{22}H_{23}O_{12}^+$	479.1189	479.1189	317	4.9	24
9 delphinidin pentoside	$C_{20}H_{19}O_{11}^+$	435.0925	435.0925	303	5.9	24
10 peonidin hexoside	$C_{22}H_{23}O_{11}^+$	463.1240	463.1240	301	6.8	24
11 malvidin hexoside	$C_{23}H_{25}O_{12}^+$	493.1346	493.1346	331	6.9	24
12 cyanidin pentoside	$C_{20}H_{19}O_{10}^+$	419.0978	419.0978	287	7.6	25
17 petunidin pentoside	$C_{21}H_{21}O_{11}^+$	449.1083	449.1083	317	8.8	24
19 peonidin pentoside	$C_{21}H_{21}O_{10}^+$	433.1134	433.1134	301	10.4	25
20 malvidin pentoside	$C_{22}H_{23}O_{11}^+$	463.1240	463.1240	331	10.7	24

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

hydrolyzable 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** were identified as hydrolyzed tannins; These compounds were present in all the part of the fruits, but their content was higher in the seeds than in the peel and pulp.

Compounds **1**, **2**, **4 5** and **14** were previously identified in leaves of *Myrtus communis*[15,16].

Compound **1** showed a pseudomolecular 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 HHDP hexoside[16].

Compound **2** showed a pseudomolecular 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-H-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[16] and thus the compound was identified as monogalloylhexoside; Compound **4** showed a pseudomolecular ion at 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 agreement with the data reported in the literature[15]. 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 ion 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*[15]. Compound **14** showed a pseudomolecular 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[16], 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 pseudolmolecular 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

113

was tentatively identified by Boulekbache-Makhlouf et al. (2013)[17] 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. 3.3.4. The pseudomolecular 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 of the resulting ion), thus it was identified as di-HHDP hexose, presumably pedunculagin;
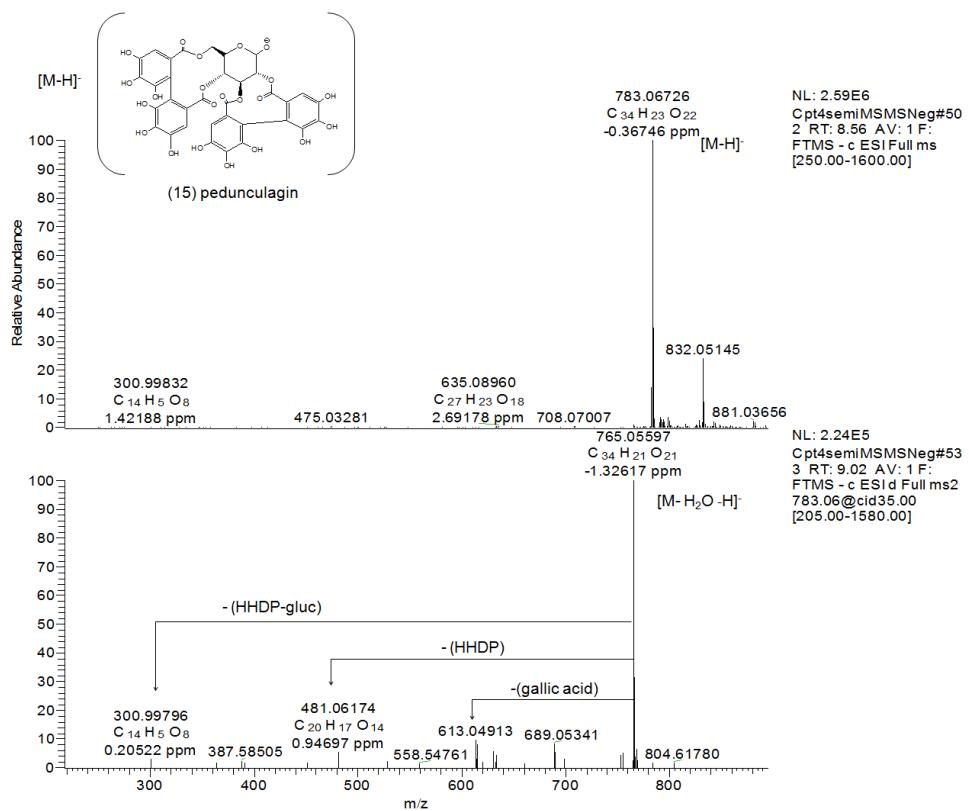
Fig. 3.3.4: MS/MS spectra of compound **15** (pedunculagin)

114

The fragmentation pattern of compound **15** is showed in fig 3.3.5 and it was previously reported by Simirgiotis et al. 2013[18] . A similar fragmentation was also observed by Boulekbache-Makhlouf et al. 2013[17] in leaves of *Eucalyptus globulus* (family Myrtaceae).
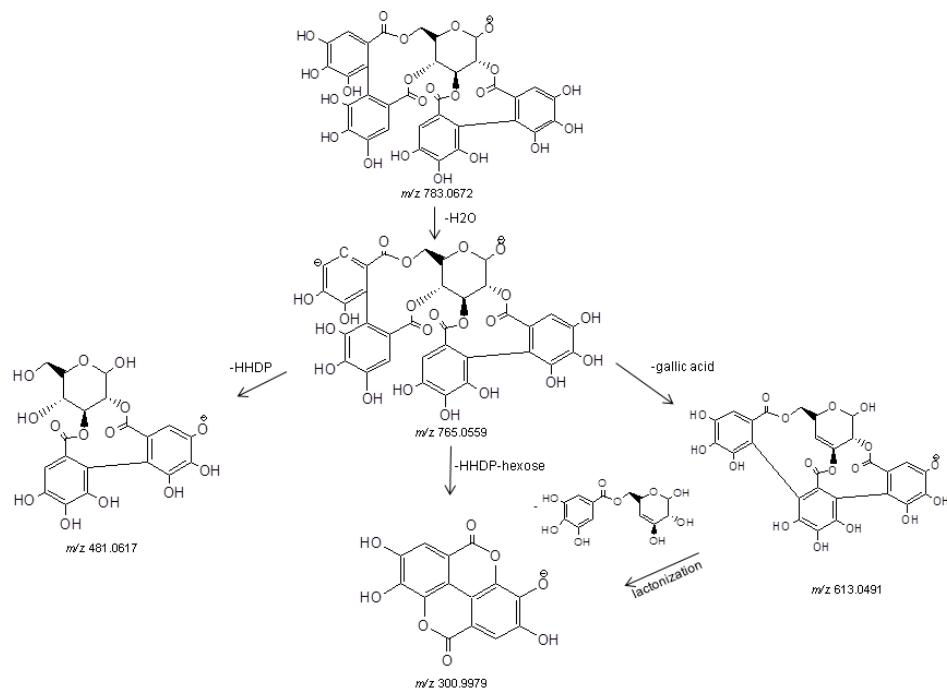

Fig. 3.3.5: Fragmentation pattern of compound **15** (pedunculagin); (HHDP:Hexahydroxydiphenic acid)

Compound **16** showed a pseudomolecular 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[19] and database, thus the compound was identified as casuarictin.

Compound **25** showed a pseudomolecular 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 in camu-camu fruit[20].

Compound **26** showed a pseudomolecular 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[17], it was tentatively identified as tellimagrandin II.

Compounds identified, pertaining to this class are reported in figure 3.3.6.
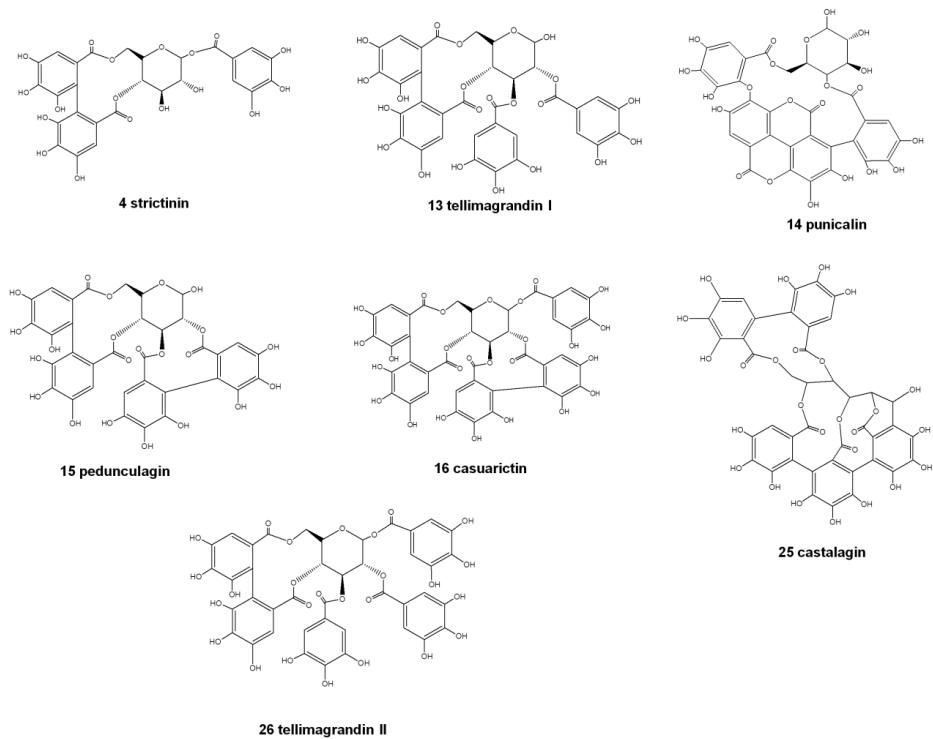
Fig. 3.3.6: Compounds identified for the first time in myrtle berry pertaining to hydrolyzed tannins

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[15].

Hydroxycinnamic acid

Compounds **18** was identified as hydroxycinnamic acid; this compound showed a pseudomolecular 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 hexose moiety) and 161, in comparison with database and data present in literature[21], the compound was identified as caffeoylhexose; it was for the first time found in myrtle berry.

Flavanols

This class of compounds 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; compounds **7** was characterised as epigallocatechin and compound **22** as catechin/epicatechin by comparison with standard compounds. They were previously reported in *Myrtus communis* berry[22].

Flavonols

This class of compounds 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 myrtle berry; these compounds were present

in all the part of the fruits, but their content was higher in the peel and pulp than in the seeds.

Compounds **23**, **28**, **33** and **34** were for the first time identified in Myrtle berry. Compound **23** showed pseudomolecular ion at *m/z* 631.0934 corresponding to molecular formula $C_{28}H_{23}O_{17}$, producing in MS/MS two principal product ion 479, which correspond to the loss of galloyl unit (152Da) and 317 which could be a result of the loss of a hexosyl moiety from the molecule at *m/z* 479 and corresponds to myricetin, The compound was tentatively identified as myricetin-galloyl-hexoside, which was previously reported in leaves of *Myrtus communis* by Taamalli et al., 2014[15].

Compound **28** showed a pseudomolecular 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 ion 463, which correspond to the loss of galloyl unit (152Da) 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 in leaves of *Myrtus communis* by Romani et al. 2004[23].

Compound **33** with pseudomolecular ion 615.0984 corresponding to the molecular formula $C_{28}H_{23}O_{16}$ showing two principal product ion 463, which correspond to the loss of galloyl unit (152 Da) and 317 which could be a following loss of a deoxyhexose moiety leading to myricetin. In comparison with literature, the compound was tentatively identified as myricetin-galloyldeoxyhexose, which was previously reported in leaves of *Myrtus communis*[15].

Compound **34** showed a pseudomolecular 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 ion 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 *Myrtus communis*[23].

Ellagic acid and derivates

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 **21** was iidentified as ellagic acid hexoside and compound **31** as ellagic acid, the two compounds were previously reported in *Myrtus communis* leaves [16,22].

Anthocyanins

This class of compounds 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*[24]. 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. Compound **12** and **19** were for the first time reported in *Myrtus communis*. Compound **12** showed a pseudomolecular 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 [25] the compound was tentatively identified as cyanidin pentoside. Compound **19** showed a pseudomolecular ion at 433.1134 corresponding to the molecular formula $C_{21}H_{21}O_{10}^{+}$ that submitted to fragmentation, 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[25] the compound was tentatively identified as peonidin pentoside.

Multivariate Data Analysis

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 (36 samples: 18 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 method (PCA).

Principal Component Analysis (PCA) was performed by applying the peak areas of each peak present in the LC-MS dataset (excluding the noisy), and a matrix was obtained by using these areas (variables), and the columns of the matrix were the different analysed samples. The resulted score scatter plot is reported in Fig. 3.3.7. The first component explains the 70.22% of variance while the second the 7.72%. 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 in figure 3.3.7A is colored according to the parts of fruit. It is clear 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. The same scatter plot in figure 3.3.7B is colored according to the cultivars of samples (different cultivar is related to 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

120

geographic origin. In fact samples pertaining to the cultivar from the area of Sassari is on the right upper part of the plot while samples pertaining to the cultivar from the area of Cagliari is on the right bottom. However, not so high differentiation is present between the seeds,   because the samples are all grouped in the same area. Thus, this is an evidence that the metabolites that make the two cultivar different are the phenolic compounds, and principally anthocyanins and flavonoids. In this study, these phenolic compounds appear to be markers compounds for geographic and genetic origin of myrtle berry.
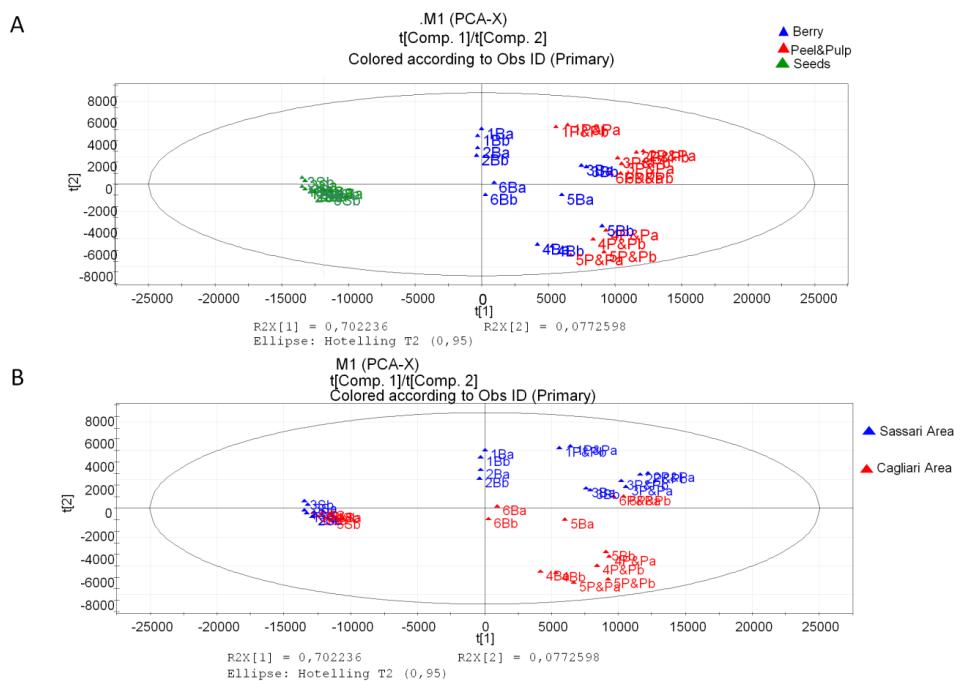

Fig. 3.3.7: Principal Component Analysis, score scatter plot, LC-ESI-Orbitrap-MS analysis. Panel A: colored according to myrtle berries parts. Panel B: colored according to geografic origin of samples

3.3.3 Conclusion

An untargeted metabolomic 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 thought liquid chromatography coupled to 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 hydrolyzable tannins was identified for the first time in these berries.

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

3.3.4 Materials and Methods

Plant materials

Seeds of two cultivars of *Myrtus communis*, were collected respectively from the geographic area of Sassari and Cagliari (Sardinia, Italy) and then grown in the Experimental Station of the University of Sassari located in Oristano (Sardinia Italy). Three samples of berries from the area of Sassari (1,2,3) and three from the area of Cagliari (4,5,6) were randomly harvested ($\approx$ 2 kg) during fruit ripening stage in 2014, December. Before analysis, samples collected were immediately frozen in liquid nitrogen and then freeze dried. 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 coming from the area of Sassari); 4S, 5S, 6S (seeds from samples coming from the area of Cagliari); 1P&P, 2P&P, 3P&P (peel&pulp from samples coming from the area of Sassari); 4P&P, 5P&P, 6P&P (peel&pulp from samples coming from the area of Cagliari); 1B,2B,3B (whole berry from the area of Sassari); 4B,5B,6B (whole berry from the area of Cagliari). Each sample was analyzed in duplicates (obtaining in total 36 samples).

Sample preparation

Seeds, Peel&pulp and whole berry were thinly pulverized and then 0.5g were extracted with 10 ml of a solution 70% of aqueous methanol acidified at 0.1% of formic acid followed by sonication for 15 minutes and centrifugation for other 15 minutes at 1750rpm. The supernatant was collected and filtered through 0.45µm filters, and then 100µl of extract were diluted in 900µl of water (ultragradient).

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

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).

HPLC-UV/VIS analysis

An Agilent (Palo Alto, CA, USA) 1260 Infinity system consisting of a G1312C binary pump, a G-1328A Rheodyne injector (20μL injection loop) a G-1379A degaser and a G1314B photodiode array detector was employed to develop the chromatographic method. Analyses were performed using a Waters XSelect CSH C18 (2.1mmx150mm 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 35min 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.

LC-ESI –Orbitrap-MS analysis

To investigate the main markers specific to each part of the berry, an HPLC method coupled with an hybrid mass spectrometer, which combines the linear trap quadrupole (LTQ) and OrbiTrap mass analyzer, was developed in this

research. All experiments were performed using a Thermo scientific liquid chromatography system constituted of a quaternary Accela 600 pump and an Accela autosampler, connected to a linear Trap-Orbitap hybrid mass spectrometer (LTQ-Orbitrap XL, Thermo Fisher Scientific, Bremen, Germany) with electrospray ionization (ESI). Separation was performed on a XSelect CSH C18 (Waters) column (2.1mmx150mm particle size 3.5 μm). The mobile phase consisted of solvent A (water + 0,1% formic acid) and solvent B (acetonitrile + 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 -12V; tube lens voltage -121.47; capillary temperature 280°C; Shealt 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. For fragmentation study, a data dependent scan was performed, selecting precursor ions corresponding to most intensive peaks in LC-MS analysis.

The experiments were also performed in positive ion mode only for a quality screening with fragmentation study. ESI source parameters were as follows: capillary voltage 49V; tube lens voltage 120; capillary temperature 280°C; Shealt and Auxiliary Gas flow (N2) 30 and 5, Sweep gas 0 Spray voltage 5. MS spectra were acquired by full range acquisition covering m/z 250-1600.

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

Multivariate Data Analysis

Raw LC-MS data of the 36 samples deriving from LC-ESI-Orbitrap-MS analysis (negative ion mode) were analyzed using a platform independent open source software package called MZmine (http://mzmine.sourceforge.net/). Using this toolbox with normalization by 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 [26,27]. Pareto scaling was applied before multivariate data analysis.

Acknowledgements

3.3.5 References

[1]Zadernowsky R., Naczk M., Nesterowicz J. Phenolic acid profiles in some small berries. *Journal of Agriculture and Food Chemistry* **2005**,53, 2118-2124

[2]Hancock R. D., McDougall G. J., & Stewart, D. Berry fruit as 'superfood': Hope or hype. *Biologist* **2007**, 54: 73–79

[3]Kool M. M., Comeskey D. J., Cooney J. M., McGhie T. K. 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 Chemistry* **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. *Journal of Agriculture and Food Chemistry* **2010**, 58, 4602–4616

[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-Reviews in Medicinal Chemistry* **2008**, 8, 1179-118

[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 Research* **2007**, 67: 3475–3482

[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]Nuvole F., and Spanu D. Analisi e prospettive economiche dell'utilizzazione industriale del mirto. *Riv. Ital. EPPOS* **1996**, 12, 231-236

[9]Mulas M.,. Melis R. A. M Essential Oil Composition of Myrtle (*Myrtus communis*) Leaves. *Journal of Herbs, Spices & Medicinal Plants* **2011** , *17(1)*, 21-34

[10]Bouzabata A., Cabral C., Goncalves M. J., Cruz M. T., Bighelli A., Cavaleiro C., Casanova J., Tomi F., Salgueiro L . *Myrtus communis* L. as source of a bioactive and safe essential oil. *Food chemical toxicology* **2015**, 75, 166-172

[11]Rahimmalek M., Mirzakhani M., Ghasemi P. A. Essential oil variation among 21 wild myrtle (*Myrtus communis* L.) populations collected from different geographical regions in Iran. *Industrial Crops and Products* **2013**, 51, 328-333

[12]Piras F. M., Dettori M. F., Magnani A. ToF-SIMS PCA analysis of *Myrtus communis* L. Applied Surface Science **2009**, 255, 7805-7811

[13]Tuberoso C. I. G., Rosa A., Bifulco E., Melis M. P., Atzeri A., Pirisi F. M., Dessì M. A. Chemical composition and antioxidant activities of *Myrtus communis* L. berries extracts. *Food Chemistry* **2010**, 123, 1242-1251

[14]Wannes W. A. and Marzouk B. Differences between Myrtle fruit parts (*Myrtus communis* var Italica) in phenolics and antioxidant contents. *Journal of Food Biochemistry* **2013**, 37, 558-594

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

[16]Romani A., Campo M., Pinelli P. HPLC/DAD/ESI-MS analyses and anti-radical activity of hydrolyzable tannins from different vegetal species. *Food Chemistry* **2012**, 130, 214-221

[17]Boulekbache-Makhlouf L., Meudec E., Mazauric J.-P., Madania K. and 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. *Phytochemical Analysis* **2013**, 24(2), 162-70

[18]Simirgiotis M. J., 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 Chemistry* **2013**, 139, 289-299

[19]Grace M. H., Warlick C. W., Neff S. A., Lila M. A. Efficient preparative isolation and identification of walnut bioactive components using high-speed

counter-current chromatography and LC-ESI-IT-TOF-MS. *Food Chemistry* **2014**, 158, 229–238

[20]Fujita A., Sarkar D., Wu S., Kennelly E., Shetty K., Genovese M. I. Evaluation of phenolic-linked bioactives of camu-camu (*Myrciaria dubia* Mc. Vaugh) for antihyperglycemia, antihypertension, antimicrobial properties and cellular rejuvenation. *Food Research International* **2015**, 77, 194–203

[21]Jaiswal R., Matei M. F., Glembockyte V., Patras M. A., and Kuhnert N. Hierarchical Key for the LC−MSn Identification of all Ten Regio- and Stereoisomers of Caffeoylglucose. *Journal of Agriculture and Food Chemistry* **2014**, 62, 9252−9265

[22]Barboni T., Venturini N., Paolini J., Desjobert J-M., Chiaramonti N., Costa J. Characterisation of volatiles and polyphenols for quality assessment of alcoholic beverages prepared from Corsican *Myrtus communis* berries. *Food Chemistry* **2010**, 122, 1304–1312.

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

[24]Montoro P., Tuberoso C. I.G., 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. *Journal of Pharmaceutical and Biomedical Analysis* **2006**, 41, 1614-1619

[25]Gardana C., Ciappellano S., Marinoni L., Fachechi C., and Simonetti P. Bilberry Adulteration: Identification and Chemical Profiling of Anthocyanins by Different Analytical Methods. *Journal of Agriculture and Food Chemistry* **2014** , 62, 10998−11004

[26]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. *Journal of Pharmaceutical and Biomedical Analysis* **2012**, 70, 224–230

[27]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, 1379–1389

Chapter 4
Metabolomics approaches used for identification of bioactive metabolites

Identify bioactive metabolites in a complex mixtures, like plant extracts, is a challenging tasks. However, new technology platforms developed in recent years have advanced the search for bioactive constituents from plant extracts. Actually hyphenated techniques like liquid chromatography coupled with high resolution mass spectrometry and nuclear magnetic resonance spectroscopy has already proven successful for fast and efficient analysis of bioactive constituents in plant extracts[1].

In this study a large scale metabolomic approach was carried out on *Morus alba* and *Morus nigra* extracts and a new protocol based on microfractionation together with microplate-based biochemical detection have been developed for the identification of α-glucosidase inhibitors direct from crude extracts of *Morus alba* and *Morus nigra* fruits in order to investigate their potential use as future functional foods.

Mass Spectrometry (MS)-based metabolomics is present in this way as a useful tool also for identification of bioactive metabolites.

The present study was developed in collaboration with Dr Ric de Vos at the Bioscience Department of Wageningen University.

132

4.1 Combined use of α- glucosidase inhibitory activity and UHPLC-High Resolution mass spectrometry for the identification of active metabolites in *Morus alba* and *Morus nigra* extracts

4.1.1 Introduction.

Mulberry belongs to the genus *Morus* of the family Moraceae. There are 24 species of *Morus* and one subspecies, with at least 100 known varieties[2]. The most commonly known species in the *Morus* genus are white mulberry (*Morus alba* L.), black mulberry (*Morus nigra* L.) and red mulberry (*Morus rubra* L.)[3]. They are deciduous tree originating from China and Japan, spread in America and Europe for the silkworm breeding.

The main use of mulberry globally is as feed for the silkworm, but depending on the location, it is also appreciated for its fruit which can be consumed both as fresh and as an ingredient in food processed products[4].

Mulberry species have been used in Traditional Chinese Medicine (TCM) for the treatment of several diseases. They were especially used for the treatment of diabetes mellitus II; in fact, they contain specific molecules that can inhibit the α-glucosidase enzyme present in the brush border of the human intestine[5,6].

133

Inhibition of α-glucosidase leads to a decrease in the rate of glucose absorption and results in a lowered postprandial blood glucose.

It has furthermore been shown that α-glucosidase inhibitors can prevent development of diabetes for people with impaired glucose tolerance and/or impaired fasting blood glucose[1].

Mulberries are also a rich source of phenolic compounds including flavonols and phenolic acids, as well as anthocyanins in the case of black and red mulberry fruits[7,8].

The present study is focused on *Morus alba* and *Morus nigra*, which are the two species of *Morus* most common in Italy.

The identification of bioactive substances in these species is key for assessing their biological and nutritional properties, for these reasons with this work we present a procedure for identification of metabolites with α-glucosidase inhibitory activity by using the combination of LC-NanoMate-FT-MS and biological assays. We firstly performed liquid chromatography with online electrospray ionization high resolution mass spectrometry LC-(ESI)-HR-MS using crude extract from both *Morus alba* and *Morus nigra* fruits. Secondly the extract was automatically fractionated in a 96-wells plate by a NanoMate LC-fraction collector/injection robot (Advion) and selected LC-fractions were subsequently analyzed thought high resolution mass spectrometry using nanospray-direct infusion; then each collected well of the plates were subsequently tested for α-glucosidase inhibitory activity. In addition, during this study, we applied an *in vitro* simulated gastrointestinal digestion model to observe if the activities of mulberries are still asserted.

134

4.1.2 Results and discussion

LC-MS analysis

To identify the major metabolites present in white and black mulberry LC-MS analysis was performed on the methanolic extract of the fruits using an Acquity UHPLC photodiode array (PDA) (Waters) coupled to an LTQ Ion trap-Orbitrap Fourier transformed Mass Spectrometer (FTMS; Thermo) hybrid system. The analysis were performed in positive ion mode. In total we could putatively identify 36 compounds in white and black fruits all reported in table 4.1.1 and numbered according to their retention time (RT).

Table 4.1.1: Metabolites identified in *Morus alba* and *Morus nigra* using UHPLC-FT-MS/MS in positive ion mode

	RT	MW	$[M+H]^+$	Molecular formula	Putative ID	MS/MS	Ref.
1	1.5	163.0844	164.0923	$C_6H_{13}NO_4$	1-deoxynojirimycin		9
2	1.7	289.2253	290.2332	$C_{15}H_{31}NO_4$	N-nonil-deoxynojirimycin	206/122	9
3	2.1	147.0895	148.0974	$C_6H_{13}NO_3$	Fagomine		9
4	2.9	181.0738	182.0817	$C_9H_{11}NO_3$	2-formyl-1H-pyrrole-1-butanoic acid	165/136	10
5	8.2	354.0951	355.1024	$C_{16}H_{18}O_9$	Caffeoylquinic acid isomer	163	11
6	9.3	449.1084	449.1084	$C_{21}H_{21}O_{11}^+$	Cyanidin hexoside	287	11
7	9.4	507.3043	508.3122	$C_{24}H_{45}NO_{10}$	Morusimic acid E	346/284	12
8	10.2	595.1662	595.1662	$C_{27}H_{30}O_{15}^+$	Cyanidin hexose deoxyhexose	449/287	11
9	10.9	433.1135	433.1135	$C_{21}H_{21}O_{10}^+$	Pelargonidin hexoside	271	11
10	11.5	354.0951	355.1024	$C_{16}H_{18}O_9$	Caffeoylquinic acid isomer	163	11
11	11.9	579.1714	579.1714	$C_{27}H_{31}O_{14}^+$	Pelargonidin hexose deoxyhexose	433/271	11
12	12.2	626.1483	627.1542	$C_{27}H_{30}O_{17}$	Quercetin hexosilhexoside	465/303	11
13	12.3	354.0951	355.1024	$C_{16}H_{18}O_9$	Caffeoylquinic acid isomer	163	11
14	12.4	772.2062	773.2135	$C_{34}H_{40}O_{31}$	Quercetin-hexose-hexose-deoxyhexose	303/465/611	13
15	13.2	466.1111	467.1190	$C_{21}H_{22}O_{12}$	Dihydroquercetin	449/305	14

					hexoside/Taxifolin hexoside		
16	14.4	712.1487	713.1544	$C_{30}H_{32}O_{20}$	Quercetin hexoside malonyl hexoside	551/463/ 303	15
17	14.7	756.2112	757.2192	$C_{33}H_{40}O_{20}$	Kaempferol hexose hexose deoxyhexose	611/449/ 287	13
18	15.3	386.1940	387.2020	$C_{19}H_{30}O_8$	Roseoside	370/208	16*
19	16.3	772.2062	773.2135	$C_{34}H_{40}O_{31}$	Quercetin-hexose-hexose-deoxyhexose	303/465/ 611	13
20	16.5	450.1162	451.1235	$C_{21}H_{22}O_{11}$	Dihydrokaempferol-hexoside	289	17
21	17.0	696.1517	697.1597	$C_{30}H_{32}O_{19}$	Kaempferol hexoside malonyl hexoside	287/449/ 535	18
22	17.2	772.2062	773.2135	$C_{34}H_{40}O_{31}$	Quercetin-hexose-hexose-deoxyhexose	303/465/61 1/ 627	13
23	18.2	756.2112	757.2192	$C_{33}H_{40}O_{20}$	Kaempferol hexose hexose deoxyhexose	611/449/28 7	13
24	18.6	756.2112	757.2192	$C_{33}H_{40}O_{20}$	Kaempferol hexose hexose deoxyhexose	611/449/28 7	13
25	19.5	491.3094	492.3173	$C_{24}H_{46}O_9N$	Morusimic acid C	330	12
26	21.3	610.1534	611.1606	$C_{27}H_{30}O_{16}$	Quercetin-hexose-deoxyhexose	303/465	19*
27	22.1	329.2566	330.2645	$C_{18}H_{35}NO_4$	Morusimic acid B	312/268/25 0	12
28	22.4	464.0954	465.1027	$C_{21}H_{20}O_{12}$	Quercetin hexose	303	11
29	24.3	594.1584	595.1664	$C_{27}H_{30}O_{15}$	Kaempferol hexose deoxyhexose	449/287	11
30	24.8	550.0958	551.1038	$C_{24}H_{22}O_{15}$	Quercetin-malonylhexoside	303	19*
31	25.1	516.1268	517.1341	$C_{25}H_{24}O_{12}$	Dicaffeoylquinic acid I	163	20*
32	25.5	448.1006	449.1078	$C_{21}H_{20}O_{11}$	Kaempferol-hexoside	287	11
33	25.6	516.1268	517.1341	$C_{25}H_{24}O_{12}$	Dicaffeoylquinic acid II	325/ 163	20*
34	27.8	516.1268	517.1341	$C_{25}H_{24}O_{12}$	Dicaffeoylquinic acid III	325/163	20*
35	28.3	534.1009	535.1088	$C_{24}H_{22}O_{14}$	Kaempferol malonyl hexoside	287	19*
36	49.1	477.3063	478.3142	$C_{24}H_{45}O_9$	Morusimic acid A	337/460	12

In the black fruits (*M. nigra*), four anthocyanins could be identified based on their exact molecular mass, their fragmentation and maxima in their light absorption spectrum at 520 nm.

Compound **6**, Cyanidin hexoside, showed a pseudomolecular ion at *m*/z 449.1084 corresponding to molecular formula $C_{21}H_{21}O_{11}^+$, that submitted to fragmentation gave one principal product ion at *m*/z 287; Compound **8**, Cyanidin hexose-deoxyhexose, showed a pseudomolecular ion at *mz* 595.1662 corresponding to molecular formula $C_{27}H_{30}O_{15}^+$, that submitted to fragmentation gave two principal product ion at 449/287. Compound **9**, Pelargonidin hexoside showed a pseudomolecular ion at *mz* 433.1135 corresponding to molecular formula $C_{21}H_{21}O_{10}^+$, that submitted to fragmentation gave one principal product ion at 271, and Compound **11**, Pelargonidin hexose-deoxyhexose showed a pseudomolecular ion at *mz* 579.1714 corresponding to molecular formula $C_{27}H_{31}O_{14}^+$, that submitted to fragmentation gave two principal product ion at 433/271. These compounds were already reported in *Morus alba* fruits from Serbia by Natic et al. in 2015[11]. The anthocyanins, responsible of black color in mulberry fruits, are also the only class of compounds find in this work, that are different from the white and black mulberry fruits.

Flavonoids, like kaempferol and quercetin with different substituent (**12**, **14**, **15**, **16**, **17**, **19**, **20**, **21**, **22**, **23**, **24**, **26**, **28**, **29**, **30**, **32**, **35**) are the other class of compounds mainly found in white and black mulberry fruits.

In particular compounds **14**,**19**,**22** never detected in this species, show different RT but same pseudomolecular ion [M+H]$^+$ at *m*/z 773.2135 and same fragmentation pattern at *m*/z 303/465/611 corresponding to the lost of two hexose and one deoxyhexose moiety, were identified as quercetin-hexose-

hexose-deoxyhexose. The data were confirmed by other data present in literature[13].

Compounds **17**,**23**,**24,** never detected in this species, show different RT but same pseudomolecular ion [M+H]$^+$ at *m*/z 757.2192 and same fragmentation pattern at *m/z* 287/449/611 corresponding to the lost of two hexose and one deoxyhexose moiety, were identified as kaempferol-hexose-hexose-deoxyhexose. The data were confirmed by other data present in literature[13].

Compound **16** never detected in this species, showed a pseudomolecular ion at *m*/z 713.1544 that submitted to fragmentation gave three principal product ion at *m*/z 303/463/551 corresponding to the lost of two hexose and one malonyl moiety; the compound was tentatively identified as quercetin hexose-malonyl-hexoside. This compound is reported in lettuce[15].

Compound **20** never detected in this species, showed a pseudomolecular ion at *m*/z 451.1235 that submitted to fragmentation gave one principal product ion at *m*/z 289 corresponding to the lost of hexose moiety; the compound was tentatively identified as dihydrokaempferol-hexoside and it is also reported in raspberry [17].

Compound **21** never detected in this species, showed a pseudomolecular ion at *m*/z 697.1597 that submitted to fragmentation gave three principal product ion at *m*/z 287/449/535 corresponding to the lost of two hexose and one malonyl moiety; the compound was tentatively identified as kaempferol hexose-malonyl-hexoside. This compound is reported in *Panax* species[18].

Three N-containing sugar like the compounds (**1**) 1-deoxynojirimicin (**2**) N-nonil deoxynojirimicin (**3**) fagomine were found in both mulberry and they were previously reported in leaves of *Morus* by Asano et al in 1994[9].

Morusimic acid A, B, C and D (**7**, **25**, **27**, **36**), were also identified based on their exact molecular mass and their fragmentation, they were previously reported by Kusano et al. in *Morus alba* fruits growing in Turkey[12].

The fruits are also rich of caffeoylquinic acid and derivates (**5**, **10**, **13**, **31**, **33**, **34**). The three isomers of caffeoylquinic acid (**5**, **10**, **13**) showed a pseudomolecular ion at 355.1024 that submitted to fragmentation gave the same daughter ion at *m/z* 163 corresponding to the lost of quinic acid moiety: these compounds were already reported in *Morus alba* fruits from Serbia by Natic et al. in 2015[11]. The three isomers of dicaffeoylquinic acid (**31**, **33**, **34**) showed a pseudomolecular ion at 517.1341 that submitted to fragmentation gave the same MS/MS base peak at 163. These compounds were previously reported in *Morus alba* leaves by Dugo et al., 2009[20].

Compound **4** identified as an alkaloid, 2-formyl-1H-pyrrole-1-butanoic acid, showed a pseudomolecular ion at *mz* 182.0817, that submitted to fragmentation gave two principal product ion at 165 and 136, the compound was previously reported in *Morus alba* fruit by Kim et al. in 2013[10].

In vitro gastrointestinal digestion

To determine the effect of *in vitro* gastointestinal digestion on metabolites of black and white mulberry, an *in vitro* gastrointestinal digestion was carried out on *Morus alba* and *Morus nigra* fruit and then LC-MS analysis were performed on the methanolic extracts of GI (gastrointestinal) and PG (post gastric digestion) of digested fruit using an Acquity UHPLC photodiode array (PDA) (Waters) coupled to an LTQ Ion trap-Orbitrap Fourier transformed Mass Spectrometer (FTMS; Thermo) hybrid system. The analysis were carried out in the same chromatographic and spectrometric conditions used for the phytochemical investigation of the original extracts (described above) and

three biological replicates for each samples were used (as described in the Materials and Methods section). Two controls were used during the analysis: the original fruit non-digested (extracts) and fruit in which the digestion protocol is followed but the enzymes were added at the end in cold, so no enzyme activity should be present (controls). A multivariate analysis was carried out on LC-MS data to visualize the changes in metabolic composition during mulberry digestion. The results of the multivarite analyses were visualized using principal components analysis plot (fig.4.1.1)
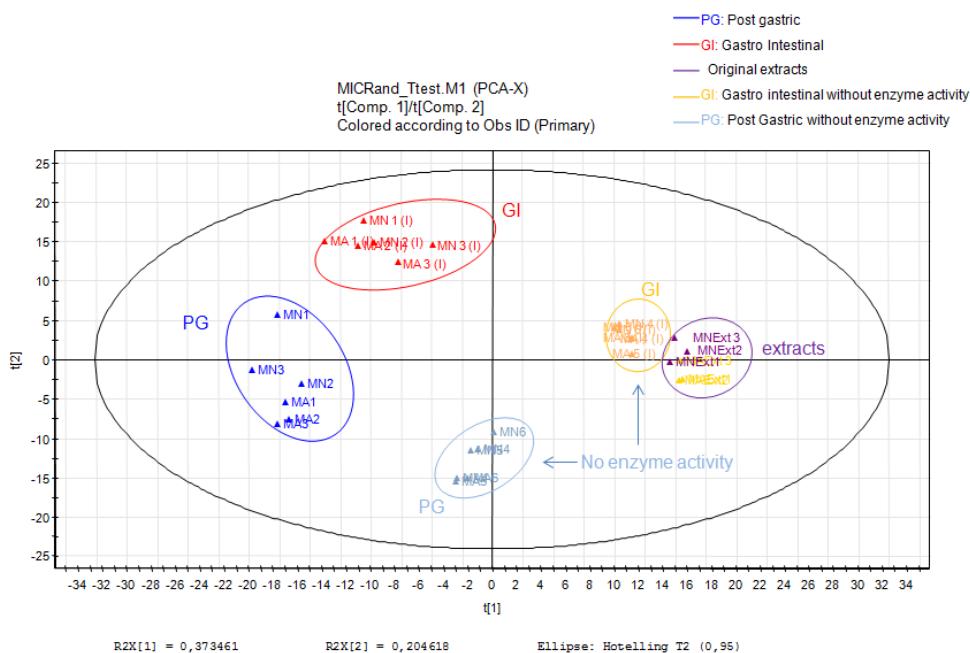
Fig. 4.1.1: Principal component analysis of untargeted LC-ESI-Orbitrap-MS data of *Morus alba* and *Morus nigra* digested samples. (MN: *Morus nigra*, MA: *Morus alba*)

PCA analysis of all analyzed samples shows that digestion has a strong effect on the metabolites composition of mulberry fruits in fact the samples after the digestion are shifted in the left part of the plot. The first principal components (X axis) explaining 37% of total variation in the data set. Digestion is known to have a profound effect on phenolic compound such as anthocyanins[4].

140

Comparing the original extracts to the digested ones, infact, a decrease of anthocyanins content was observed in *Morus nigra* samples, while no differences were present on the other phenolic compounds.

α -glucosidase inhibitory activity before and after in vitro gastrointestinal digestion

The α-glucosidase inhibitory activity was first of all evaluated on the methanol extracts of black and white mulberry fruits and was determined spectrophotometrically at 415 nm on spectrophotometer UV-Vis (TECAN SpectraFluor microplate reader) using p-nitrophenyl-α-D-glucopyranoside (p-NPG) as substrate. The released p-Nitrophenol from p-NPG was monitored in time; The results are reported in figure 4.1.2: the increasing point per minutes of the absorbance at 415nm is reported in Y axis of the histogram, while in x axis the tested sample; less is the value in Y axis higher is the activity of inhibition. The extracts are both active but the *Morus nigra* extract possess higher activity then the *alba* and in total their activity of inhibition range from 16 and 25 mM of Acarbose used as positive control (fig. 4.1.2B).
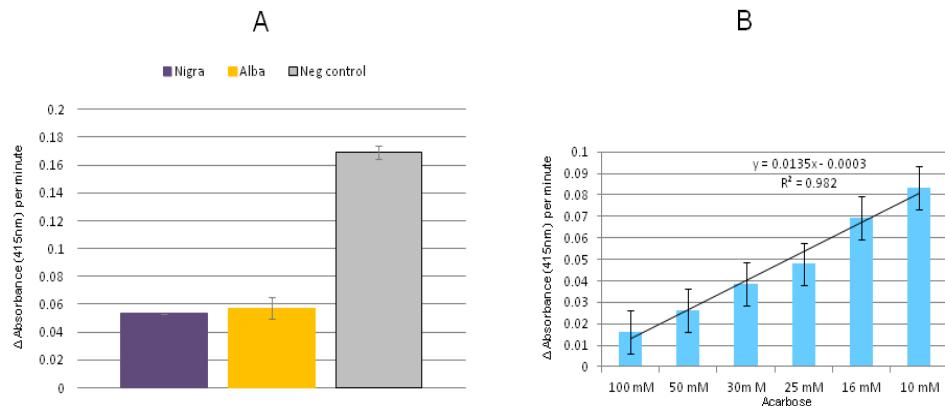
Figure 4.1.2: α-glucosidase inhibitory activity of methanol extracts
in Y axis the increasing point per minutes of the absorbance at 415 nm and in X axis the samples tested. 1A) inhibition activity from methanol extract of *Morus alba* and *Morus nigra* compared using as negative control (water) 1B) inhibition activity from Acarbose (positive control) in different concentration.

An *in vitro* simulated gastrointestinal digestion model was then applied to observe how the α-glucosidase inhibitory activity change before and after the gastrointestinal digestion of mulberry fruit.

After the *in vitro* digestion the activity is still maintained both after the stomach digestion (PG) and at the end of digestion (GI) (fig 4.1.3).
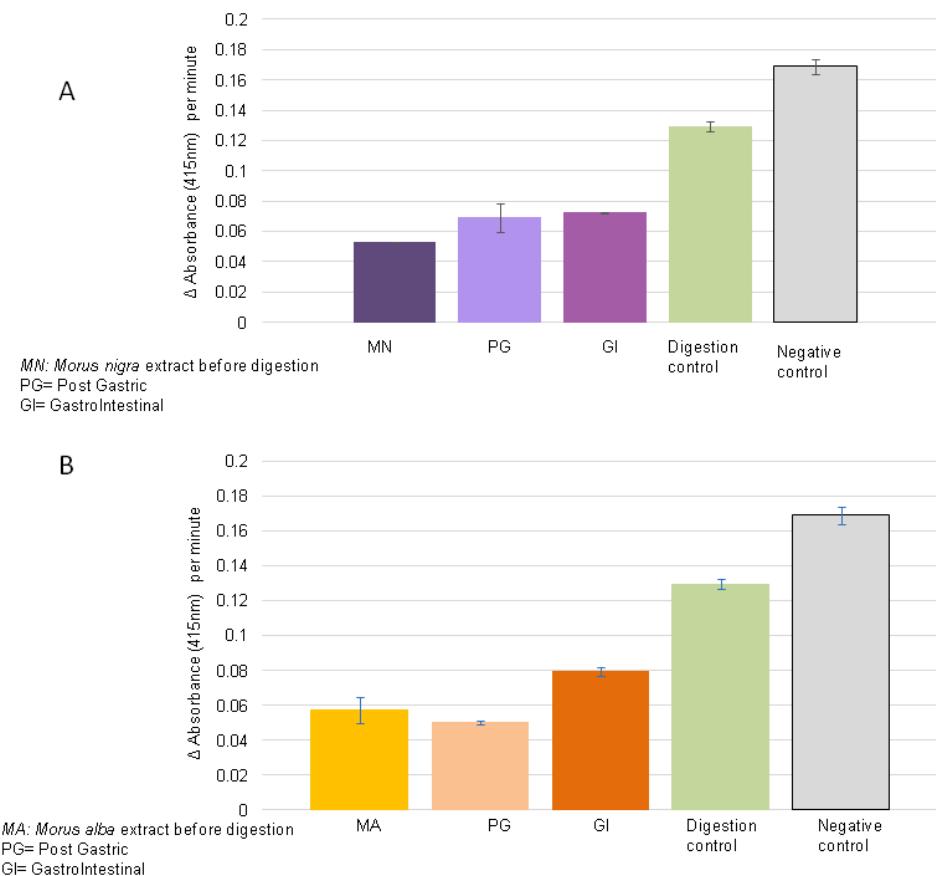
Figure 4.1.3: α-glucosidase inhibitory activity after the *in vitro* gastrointestinal digestion
2A) inhibition activity of *Morus nigra* extract, Post gastric (PG), end of gastrointestinal digestion (GI), digestion control and negative control (water)
2B) inhibition activity of *Morus alba* extract, Post gastric (PG), end of gastrointestinal digestion (GI), digestion control and negative control (water)

In particular for *Morus alba* the activity after the stomach digestion increase while for the *Morus nigra* a slight decrease is present, probably due to the reduction of anthocyanins after the stomach digestion, which are in part responsible of inhibition activity (as demonstrated below).

LC-NanoMate-FT-MS

Actually we present a procedure for identification of metabolites with α-glucosidase inhibitory activity by using the combination of Liquid Chromatography with 96 well plate fractionation/injection robot (NanoMate Triversa) and FT-MS.

We firstly performed liquid chromatography with online electrospray ionization high resolution mass spectrometry LC-(ESI)-HR-MS using crude extract from both *Morus alba* and *Morus nigra* fruits. Secondly the extract was automatically fractionated in a 96-wells plate by a NanoMate LC-fraction collector/injection robot (Advion) and selected LC-fractions were subsequently analyzed thought high resolution mass spectrometry using nanospray-direct infusion; then each collected well of the plates were subsequently tested for α-glucosidase inhibitory activity.

The analysis were performed in triplicates and we used as blank a water control; the water was injected and fractionated in the 96 well plate; that was then dried under N_2 flow at 40°C and then each well was tested for the α-glucosidase inhibitory activity. The inhibition activity was monitored in time and the results are showed in fig. 4.1.4. In X axis we have the tested wells and in Y axis the increasing point per minutes of the absorbance at 415nm, less is this value higher is the activity of inhibition. In the tested control a variation in the wells is present (fig. 4.1.4).

water injection (blanc control)
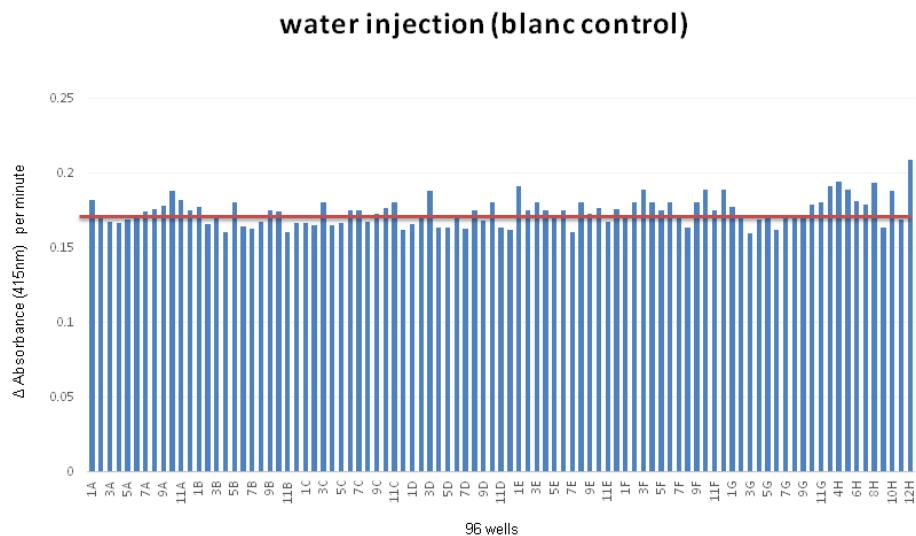
Figure 4.1.4: α-glucosidase inhibitory activity of blank (water) fractionated in 96 well plate using the combination of LC-fractionation/injection robot (nanomate) FTMS
The Y axis show the increasing point per minutes of the absorbance at 415nm in X axis the tested wells. We put a threshold value at 0.15 for the next experiments to consider the wells active.

In particular the increasing point per minutes of the absorbance at 415 nm (Y axis) range from 0.15 to 0.20 so we put a threshold value under which for the next analysis on the original extracts we can consider the wells active or not in the inhibition of the enzyme; this value was fixed at 0.15, so under these value the wells can be considered containing active metabolites. The original extract of *Morus alba* and *Morus nigra* were then fractionated in 96 well plate, the dried plate was then tested as the blank for the α-glucosidase inhibitory activity. In Fig. 4.1.5 are reported the results for *Morus alba* extract and in fig. 4.1.6 are reported the results for *Morus nigra* extract.

Retention time	Num well	Bioactive Metabolite
1.6-2.08	1A	1-2-3
4.88-5.35	8A	Not identified
12.82-13.28	1C	15
13.75-14.23	3C	16-17
19.82-20.3	9D	25
22.62-23.08	3D	28
23.55-24.02	1D	29
24.95-25.42	3E	32
31.02-31.5	9F	Not identified
34.7-35.2	1F	Not identified
35.22-35.68	1G	Not identified
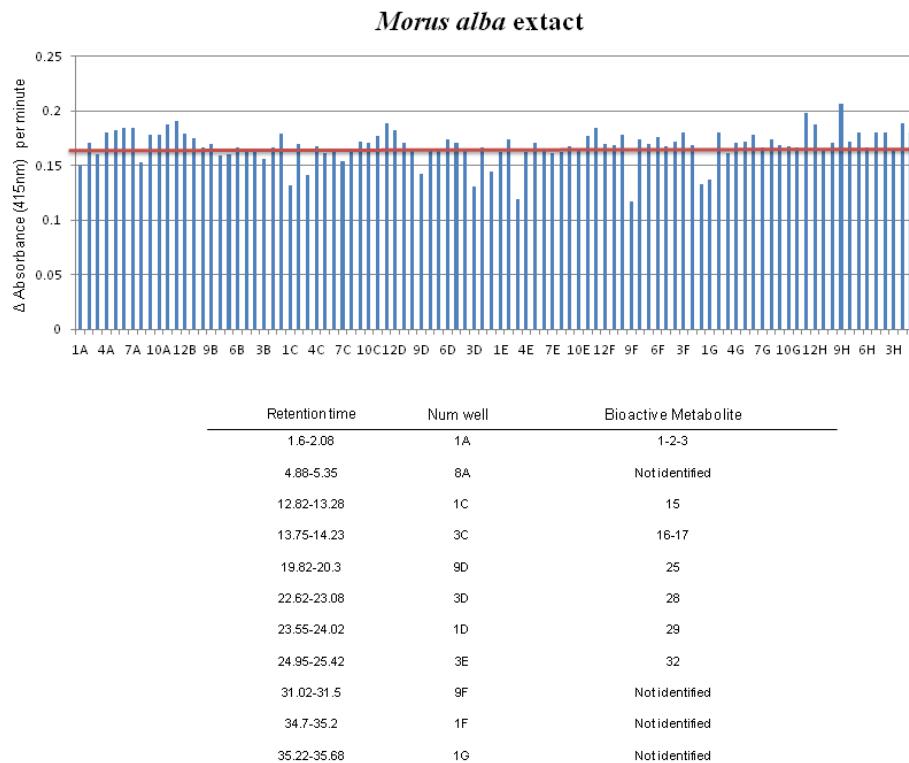
Figure 4.1.5 α-glucosidase inhibitory activity of *Morus alba* extract fractionated in 96 well plate using the combination of LC-fractionation/injection robot (nanomate) FTMS
The Y axis show the increasing point per minutes of the absorbance at 415nm in X axis the tested wells. The wells considered active are the ones under 0.15 and are all reported in the table with the corresponding number of compounds.

146

Retention time	Num well	Bioactive Metabolite
1.8-2.27	1A	1-2-3
4.13-4.6	6A	Not identified
8.33-8.8	10B	6
10.2-10.67	6B	9
12.53-13	1B	15
13.4-13.93	2C	16
14.4-14.87	4C	17
15.8-16.27	7C	19
16.27-16.73	8C	20
24.67-25.13	2E	31/32
40.07-40.53	11G	Not identified
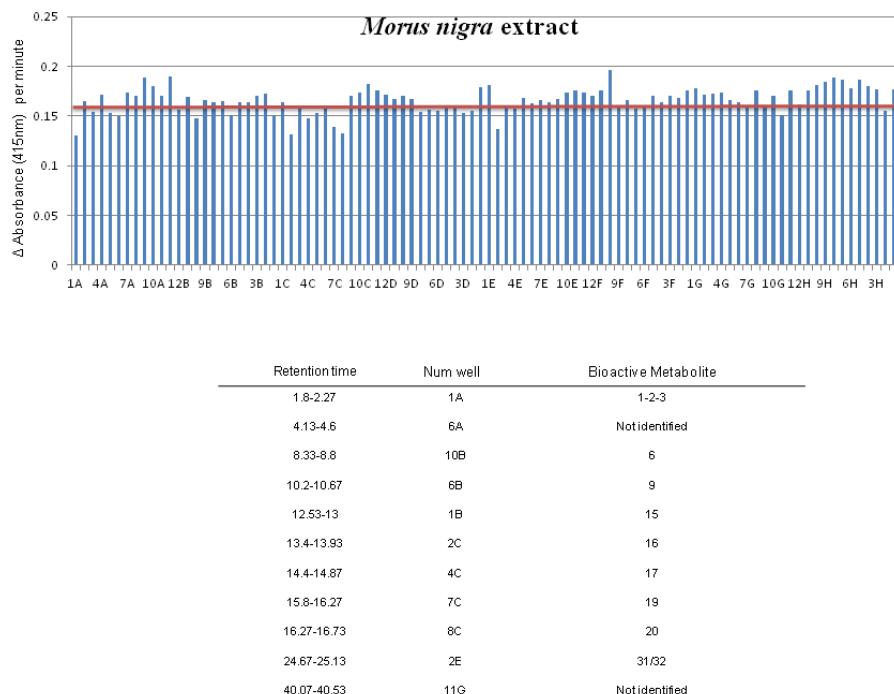
Figure 4.1.6 α-glucosidase inhibitory activity of *Morus nigra* extract fractionated in 96 well plate using the combination of LC-fractionation/injection robot (nanomate) FTMS
The Y axis show the increasing point per minutes of the absorbance at 415nm in X axis the tested wells. The wells considered active are the ones under 0.15 and are all reported in the table with the corresponding number of compounds.

The compounds active in both the fruits were: the N containing sugar **1**-**2**-**3**, the flavonoids **15**-**16**-**17**-**27**-**32** for the *nigra* and *alba* extracts and two anthocyanins **6** and **9** in *nigra*.

This new technique was useful for the identification of some metabolites with α-glucosidase inhibitory activity directly after a single LC-MS run; we can for now confirm the *α*-glucosidase inhibitory activity of anthocyanins like cyanidin hexoside and pelargonidin hexoside, that are already reported in literature as inhibitor of this enzyme [21]; and the activity of inhibition of flavonoids like kaempferol glucoside, quercetin glucoside and N-containing

147

sugar that are also reported in literature as inhibitor of α-glucosidase [6,21]; although, other wells seems to show inhibition activity, but during this study, it was not possible to identify these active metabolites on the bases of only their exact mass, however this method could be well optimized and adapted for the identification of active metabolites, in fact, according to a recent study van der Hoofr et al. confirm the importance of the so-called multistage mass spectrometry (MS^n) spectral tree approach as tool in metabolite identification in complex sample, combining the use of LC-NanoMate-FT-MS.

Targeted HPLC analysis and online antioxidant measurements

The antioxidant activity of mulberry was also compared between other fruits best known for their antioxidant activity as strawberries (*Fragaria ananassa*) and wild strawberries (*Fragaria vesca*). The activity was tested on fresh weight and performed by spectrophotometric assay and the results here reported are defined as the concentration of a standard trolox solution with the same antioxidant capacity as a 1 mg/ml of the tested extract.

The ABTS$^{\cdot+}$ assays indicated that the methanol extract of *Morus nigra* is more active than the methanol extract of *Morus alba*, in fact methanol extract of *Morus nigra* showed the same activity of methanol extract of strawberry fruits (table 4.1.2), this could be due to the presence of anthocyanins which provide the dark color of berry fruits and that are principally responsible of antioxidant activity[22].

Table 4.1.2: Antioxidant capacity of *Morus alba* and *Morus nigra* fruits compared with other fruits

	TEAC mg/ml
Morus alba (**White mulberry**)	3.58+/- 0.0187
Morus nigra (**Black mulberry**)	4.49+/-0.0005
Fragaria vesca (**Wild strawberry**)	4.60+/-0.0020
Fragaria ananassa (**Strawberry**)	4.66+/-0.0045

Based on this results an online HPLC antioxidant detection, based on a post-column reaction with ABTS$^{.+}$ cation radicals, was used to determine the relative contributions of these individual components to the total antioxidant activity[23]. A number of these components were identified by comparison of their retention time and absorption spectrum to the data of LC-MS/MS analysis. According to the results (showed in Fig 4.1.7), the most abundant antioxidant peaks in *Morus nigra* corresponded to anthocyanins, in particular compounds **6** and **9** show the highest activity. The other compounds responsible of antioxidant activity in both the methanol extract were caffeoylquinic acid like compounds **10**, **13** and flavonoids like compounds **15**,**26**, **28**, **32**.
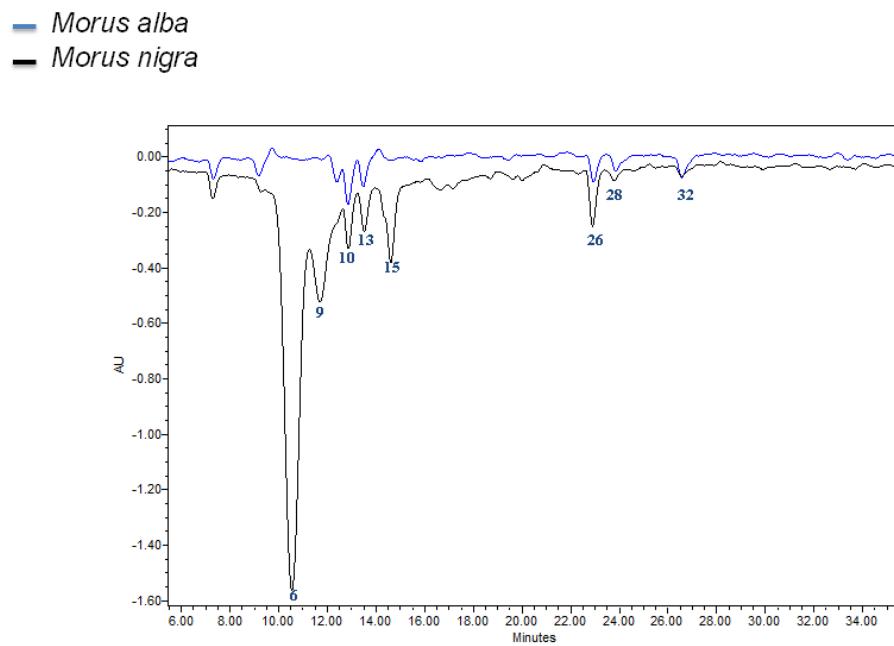
Figure 4.1.7: antioxidant activity
Representative chromatograms of antioxidant peaks in *Morus alba* (in blue) and *Morus nigra* (black) fruits determined, online, by a post column reaction with ABTS$^{\cdot+}$ cation radicals on HPLC. Number refers to the main peaks identified: **6** cyanidin hexoside, **9** pelargonidin hexoside, **10** and **13** caffeoylquinic acid isomers, **15** dihydroquercetin hexoside, **26** Quercetin hexose deoxyhexose, **28** quercetin hexose and **32** kaempferol hexoside.

4.1.3 Conclusion

In conclusion, with this new technique was possible to confirm that 1-deoxynojirimycin (**1**), N-nonil-deoxynojirimycin (**2**), fagomine (**3**), cyanidin hexoside (**6**), pelargonidin hexoside (**9**), dihydroquercetin hexoside, (**15**) quercetin hexoside malonyl hexoside (**16**), kaempferol hexose hexose deoxyhexose (**17**), quercetin hexose (**28**), kaempferol hexoside (**32**), are responsible for the major α-glucosidase inhibitory activity of methanol extracts of mulberry fruits.

This shows that *Morus alba* and *Morus nigra* hold potential as future functional foods, but further studies for investigation of optimal preparation

procedures, as well as safety of optimized preparations and NMR study, are needed.

4.1.4 Materials and Methods

Mulberry materials

Fruits of *Morus alba* and *Morus nigra* were manually picked at ripening stage in May 2014 in different geographic area of Campania region (Italy). All samples were freeze dried in Italy before being transported to The Netherlands. They were then ground to a fine powder and stored at -80^{0}C until analysis.

Extract preparation

The sample extracts which were used for the LC-MS analysis were essentially prepared in this way: exactly 30 mg of freeze dried sample were extracted with 1200 µl of methanol 75% in MQ water containing 0.1% of formic acid. Extracts were sonicated for 15 min, centrifuged at 12500 xg for 10 min, filtered through 0.45 um filters (Minisart SRP4, Biotech GmbH, Germany) and then 5µl were inject in the UHPLC-LTQ-Orbitrap.

The sample extracts, which were instead used in NanoMate LC-fractionation-FT-MS analysis and in HPLC with online antioxidant assay analysis, were prepared as described above but in this case, after centrifugation, 1ml of supernatant was collected and concentrated in SpeedVac, then, taken up in 250µl of water, filtered through 0.45um filters (Minisart SRP4, Biotech GmbH, Germany), in this case the extracts were prepared in triplicates.

Metabolomic analysis

Metabolite analysis was performed using an Acquity UHPLC photodiode array (PDA), (Waters) coupled to an LTQ Ion trap-Orbitrap Fourier transformed Mass spectrometer (FTMS; Thermo) hybrid system. A Luna 3μ C18 150x2mm column (Phenomenex, USA) was used to separate the extracted metabolites, with MQ water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B) as solvents. A linear gradient from 5 to 35% B at a flow rate of 0.19 ml/min was used. Each samples was investigated in positive ion mode. ESI source parameters were as follows: capillary voltage 43V; tube lens voltage 120V; capillary temperature 295°C; Sheath and Auxiliary Gas flow 40 and 3 (arbitrary units), Sweep gas 0(arbitrary units) Spray voltage 5V. MS spectra were acquired by full range acquisition covering *m/z* 104-1350. For fragmentation study, a data dependent scan was performed, selecting precursor ions corresponding to most intense peaks in LC-MS analysis.

In vitro simulated gastrointestinal digestion

In *vitro* digestion was carried out on dried powdered forms of both fruit samples (*Morus alba* and *Morus nigra*), following the protocol described by McDougall et al. (2005)[24] with slight modification and consisted of two sequential steps; an initial pepsin/HCl digestion for 2h at 37°C to simulate gastric conditions (GI:Gastro Intestinal) followed by a digestion with bile salts/pancreatin for 2h at 37°C to simulate small intestine conditions (PG: Post gastric). The process was performed in triplicates and three controls were used: 1) Plant material without digestion (MA ext: *Morus alba*; MN ext: *Morus nigra*), 2) the solution with all the ingredients for the digestion but without plant material, 3) plant material with all the ingredient for the digestion but without enzyme activity (The enzyme where added at the end of

152

the process in cold). Release of phytochemicals from fruit was analyzed on UHPLC-ESI-Orbitrap-MS at different stages of digestion. These stages represented the aliquots from gastric digest (post gastric, PG) and from GI (Gastrointestinal digest, GI). PG and GI samples were stored at -80^{0}C until further analysis. The samples used for LC-MS and multivariate data analysis are described as follow: Original extract of *Morus alba* (MA ext1, MA ext2, MA ext3) Original extract of *Morus nigra* (MN ext1, MN ext2, MN ext3); PG for *Morus alba* (MA1(I), MA2(I), MA3(I)); PG for *Morus nigra* (MN1(I), MN2(I), MN3(I)); GI for *Morus alba* (MA1, MA2, MA3; GI for *Morus nigra* (MN1, MN2, MN3); PG control for *Morus alba* in which the pepsine were added in cold (MA4(I), MA5(I), MA6(I)) and for *Morus nigra* (MN4(I), MN5(I), MN6(I)); GI control for *Morus alba* in which the pepsine were added in cold (MA4, MA5, MA6) and for *Morus nigra* (MN4, MN5, MN6).

α-glucosidase inhibitory activity and antioxidant activity were investigated for each of these PG and GI samples using the methods described below.

Processing of LC-MS data

Acquisition and visualization of the LC-ESI-Orbitrap-MS data were performed using Xcalibur software version 2.1. The MetAlign software package (www.metAlign.nl) was used for baseline correction, noise extimation and spectra alignment. The aligned and filtered peak table was subjected to redundancy removal using MSClust software[25] by which mass peaks originating from the same metabolites (including isotopes, adducts, and fragments) were clustered. This analysis resulted in a total of 250 reconstructed metabolites (centrotypes) of which for each cluster the highest signal per metabolite was chosen as a representative for the respective cluster and was used for further (statistical) analysis. Comparison and visualization of

the main features of the LC–MS data were performed by loading the data matrix into SIMCA P+ software 12.0 (Umetrix AB, Umea Sweden). PCA was performed and Pareto scaling was applied before multivariate data analysis.

α-glucosidase assay

The assay uses the p-nitrophenyl-α-D-glucopyranoside (pNPG) as the substrate, which is hydrolyzed by α-glucosidase to realease p-nitrophenol (pNP), a color agent that can be monitored at 415nm. Briefly 10 µl of extract were combined with 40µl of 100mM phosphate buffer (pH 6.8), and 20µl of α-glucosidase (0.6 u/ml). After shaking for intensive mixing and incubation for 5 min at 37^0C, 20µl of a 20mM pNPG solution was added to start the reaction. The reaction was monitored in time at 415 nm by TECAN SpectraFluor microplate reader. Acarbose was used as positive control, while water was used as negative control. The activity was monitored in time, the results here reported are represented as an histogram in which the Y axis shows the increasing point per minutes of the absorbance at 415nm (Δ absorbance per minute) the X axis the sample treated.

NanoMate LC-fractionation of extracts

The UHPLC–PDA–ESI–MS^n system was adapted with a chip-based nano-electrospray ionization source/fractionation robot (NanoMate Triversa, Advion BioSciences) coupled between the PDA and the inlet of the Ion Trap/Orbitrap hybrid instrument. Sample injection volume was 5µl. The gradient and flow conditions were the same as described above, with an additional 30 µl/min 100% isopropanol added into the LC flow via a T-junction between the PDA and the NanoMate. The eluents flow was split by the NanoMate, at 219.5 µl/min to the fraction collector and 0.5 µl/min to the nano-electrospray source.

LC-fractions were collected every 28.2s (i.e., 100 µl) into a 96 wells plate (Twin tec, Eppendorf). After collection, the 96 wells plate was dried at 40^0C and under N_2 flow and then tested to α-glucosidase inhibitory activity as described above.

Targeted HPLC analysis and online antioxidant measurements

The total antioxidant capacity of mulberry extracts was compared with other fruits best known for their antioxidant activity like *Fragaria vesca* (wild strawberry) and *Fragaria ananassa* (strawberry) using ABTS$^{·+}$ method as according to that used by Capanoglu et al. in 2008[26] with slight modifications. The fruits were collected in June 2014 and the antioxidant activity was tested on fresh weight. 0.5g of samples were extracted in 1.5ml of Methanol containing 0.05% of formic acid, sonicated for 15 minutes and centrifuged at 12500 xg for 15 min, filtered through 0.45 um filters (Minisart SRP4, Biotech GmbH, Germany) and then 10µl of extract were used to test the antioxidant activity and Trolox was used as reference.

10µl of sample extracts or standard was mixed with 90µl of ABTS-working solution (pH 7.4) and after 40s the remaining ABTS$^{·+}$ radical were measured at 415nm using 96-well microplates (Nunc, Roskilde, Denmark) and Infinite$^{®}$ M200 micro plate reader (Tecan, Gröding, Austria) The experiments were done in triplicates and the trolox equivalent antioxidant capacity (TEAC) value, here reported, is defined as the concentration of a standard trolox solution with the same antioxidant capacity as a 1 mg/ml of the tested extract.

The contribution of each identified antioxidant compound to the total antioxidant capacity of the extracts was determined using an HPLC-PDA system coupled to postcolumn on-line antioxidant detection [23,26]. Sample extracts of *Morus alba* and *Morus nigra* fruits, prepared following the

155

procedure described below, were analyzed using a W600 Waters HPLC system coupled to a Waters 996 photodiode array (PDA) detector, as described by Bino et al. (2005)[27]. Anthocyanins were detected at 512 nm[23], flavonols at 360 nm [28]. Eluted compounds reacted online for 30 s at 40 °C with a buffered solution of ABTS$^{\cdot+}$ cation radicals (pH 7.4). Then, the absorption of the remaining ABTS$^{\cdot+}$ radicals was monitored at 412 nm by a second detector (Waters 2487, dual-wavelength UV−vis detector). Peak identification was done by comparing absorbance spectra and retention times of eluting peaks with data taken from the literature and confirmed by UHPLC-FTMS and MS/MS analysis as described above.

Acknowledgement

4.1.5 References

[1]Kongstad, K. T., Özdemir, C., Barzak, A., Wubshet, S. G., Staerk D. Combined use of high-resolution α-glucosidase inhibition profiling and High-Performance Liquid Chromatography−High-Resolution Mass Spectrometry−Solid-Phase Extraction−Nuclear Magnetic Resonance

Spectroscopy for investigation of antidiabetic principles in crude plant extracts. *Journal of Agricultural and Food Chemistry* **2015**, 63, 2257-2263.

[2]Ercisli, S., & Orhan, E. Chemical composition of white (*Morus alba*), red (*Morus rubra*) and black (*Morus nigra*) mulberry fruits. *Food Chemistry* **2007**, 103, 1380-1384.

[3]Gundogdu, M., Muradoglu, F., Gazioglu-Sensoy, R. I., & Ylmaz, H. Determination of fruits chemical properties of *Morus nigra* L., *Morus alba* L. and *Morus rubra* L. by HPLC. *Scientia Horticulturae* **2011**, 132, 37-41.

[4]Liang, L., Wu, X., Zhao, T., Zhao, J., Li, F., Zou, Y., Mao, G., Yang, L. In vitro bioaccessibility and antioxidant activity of anthocyanins from mulberry (Morus atropurpurea Roxb) following simulated gastro-intestinal digestion. *Food Research International* **2012**, 46, 76-82.

[5]He, H. and Lu, Y-H. Comparison of inhibitory activities and mechanism of five mulberry plant bioactive components against α-glucosidase. *Journal of Agricultural and Food Chemistry* **2013**, 61,8110-8119.

[6]Banu, S., Jabir, N. R., Manjiunath, N. C., Khan, M. S., Asharf, G. M., Kamal, M. A., Tabrez. S. *Saudi Journal of Biological Sciences* **2015**, 22, 32-36.

[7]Sánchez-Salcedo, E. M., Mena P., García-Viguera C., Hernández, F., Martínez, J. J.. (Poly)phenolic compounds and antioxidant activity of white (Morus alba) and black (Morus nigra) mulberry leaves: Their potential for new products rich in phytochemicals, *Journal of Functional Foods* **2015a** doi: 10.1016/j.jff.2015.03.053

[8]Sánchez-Salcedo, E. M., Mena, P., García-Viguera, C., Martínez, J. J., Hernández, F. Phytochemical evaluation of white (*Morus alba* L.) and black (*Morus nigra* L.) mulberry fruits, a starting point for the assessment of their beneficial properties. *Journal of functional foods* **2015b**, 12, 399–408.

[9]Asano, N., Oseki, K., Tomioka, E., Kizu, H., Matsui, K. N-containing sugars fromMorus alba and their glycosidase inhibitory activities. *Carbohydrate Research* **1994**, 259,243-255.

[10]Kim S. B., Chang B. Y., Jo Y. H., Lee S. H., Han S.-B., Hwang B. Y., Kim S. Y., Lee Mi K. Macrophage activating activity of pyrrole alkaloids from Morus alba fruits. *Journal of Ethnopharmacology* **2013**, 145, 393-396.

[11]Natić,M. M., Dabić, D. Č. Papetti, P., Fotirić Akšić, M. M., Ognjanov, V., Ljubojević,

M., Tešić, Ž. Lj. Analysis and characterisation of phytochemicals in mulberry (*Morus alba* L.) fruits grown in Vojvodina, North Serbia. *Food Chemistry* **2015**, 171, 128-136.

[12]Kusano G., Orihara S., Tsukamoto D., Shibano M., Coskun M., Guvenc A., and C. S. Erdurak. Five New Nortropane Alkaloids and Six New Amino Acids from the Fruit of *Morus alba* LINNE Growing in Turkey. *Chemical and Pharmaceutical Bulletin* **2002**, 50, 185-192

[13]van der Hooft, J. J. J., Vervoort, J., Bino, R. J., de Vos R. C. H. Spectral trees as a robust annotation tool in LC–MS based metabolomics. *Metabolomics* **2012**, 8, 691-703.

[14]Zhang, W., Han, F., He, J., Duan, C. HPLC-DAD-ESI-MS/MS analysis and antioxidant activities of Nonanthocyanin phenolics in mulberry (*Morus alba* L.). *Journal of Food Science* **2008**, 73,512-518.

[15]Llorach, R., Martinez-Sanchez A., Tomas-Barberan F. A., Gil M. I., Ferreres F. Characterization of polyphenols and antioxidant properties of five lettuce varieties and escarole. *Food Chemistry* **2008**, 108, 1028-1038.

[16]Doi K., Kojima T., Makino M., Kimura Y., Fujimoto Y. Studies on the constituents of the leaves of *Morus alba* L. *Chemical and Pharmaceutical Bulletin* **2001**, 49, 151-153.

[17]Paudel, L., Wyzgoski, F. J., Scheerens, J. C., Chanon, A. M., Reese, R. N., Smiljanic, D., Wesdemiotis, C., Blakeslee, J. J., Riedl, K. M., and Rinaldi P. L. Nonanthocyanin secondary metabolites of black raspberry (*Rubus occidentalis* L.) fruits: identification by HPLC-DAD, NMR, HPLC-ESI-MS, and ESI-MS/MS analyses. *Journal of Agricultural and Food Chemistry* **2013**, 61, 12032-12043.

[18]Mao, Q., Bai, M., Xu, J.-D., Kong, M., Zhu, L.-Y., Wang, Q., li, S.-L. Descrimination of leaves of *Panax Ginseng* and *P. quinquefolium* by ultra high performance liquid chromatography quadrupole/time-of-flight mass spectrometry based metabolomics approach. *Journal of Pharmaceutical and Biomedical Analysis* **2014**, 97, 129-140.

[19]Thabtì, I., W. Elfalleh, H. Hannachi, A. Ferchichi and M. D. G. Campos. Identification and quantification of phenolic acids and flavonol glycosides in Tunisian *Morus* species by HPLC-DAD and HPLC-MS. *Journal of functional foods* **2012**, 4, 367-374.

[20]Dugo P., Donato P., Cacciola F., Germano M. P., Rapisarda A., Mondello L. Characterization of the polyphenolic fraction of *Morus alba* leaves extracts by HPLC coupled to a hybrid IT-TOF MS system. *Journal of separation science* **2009**, 32,3627-3634

[21]He, H. and Lu, Y-H. Comparison of inhibitory activities and mechanism of five mulberry plant bioactive components against α-glucosidase. *Journal of Agricultural and Food Chemistry* **2013**, 61,8110-8119.

[22]Petra, K., Ruslan, M., Adriana, E., Dagmar, M., Tomas, C., Jan, K., Dana; Juraj, P., L'udmila, T. In vitro study of biological activities of anthocyanin-rich berry extracts on porcine intestinal epithelial cells. *Journal of the Science of Food and Agriculture* **2015**. doi: 10.1002/jsfa.7181.

[23]Beekwilder, J., Jonker, H., Meesters, P., Hall, R. D., Van Der Meer, I., De Vos, R. C. H., Antioxidant in raspberry: On-line analysis links antioxidant activity to a diversity of individual metabolites. *Journal of Agricultural and Food Chemistry* **2005**, 53, 3313-3320.

[24]McDougall, G. J., Dobson, P., Smith, P., Blake, A., & Stewart, D. Assessing potential bioavailability of raspberry anthocyanins using an in vitro digestion system. *Journal of Agricultural and Food Chemistry* **2005**, 53(15), 5896–5904.

[25]Tikunov, Y. M., Laptonek, S., Hall, R. D., Bovy, A., & de Vos, R. C. MSClust: a tool for unsupervised mass spectra extraction of chromatography–mass spectrometry ion-wise aligned data. *Metabolomics* **2012**, 8, 714–718.

[26]Capanoglu, E., Beekwilder, J., Boyacioglu, D., Hall, R., De Vos, R. Changes in antioxidant and metabolite profiles during production of tomato paste. *Journal of Agricultural and Food Chemisty* **2008**, 56, 964-973.

[27]Bino, R. J., De Vos, R. C. H., Lieberman, M., Hall, R. D., Bovy, A., Jonker, H. H., Tikunov, Y., Lommen, A., Moco, S., & Levin, I. The light-hyperresponsive high pigment-2dg mutation of tomato alterations in the fruit metabolome. *New Phytologist* **2005**, 2, 427-438.

[28]Fang, F., Li, J.-M, Pan, Q.-H., & Huang, W.-D. Determination of red wine flavonoids by HPLC and effect of aging. *Food Chemistry* **2007**, 101, 428-433.

Publication List

1. Soufi S., **D'Urso G.**, Pizza C., Rezgui S., Battaieb T., Montoro P. Steviol glycosides targeted analysis in leaves of *Stevia rebaudiana* (Bertoni) from plants cultivated under chilling stress condition. *Food Chemistry* **2016**, 190, 572-580.

2. **D'Urso G.**, d'Aquino L., Pizza C., Montoro P. Integreted mass spectrometric and multivariate data analysis approach for the discrimination of organic and conventional strawberry (*Fragaria ananassa* Duch.). *Food Research International* **2015** 77, 264-272.

3. Mari A., Montoro P., **D'Urso G.**, Macchia M., Pizza C., Piacente S. Metabolic profiling of *Vitex agnus castus* leaves, fruits and sprouts: analysis by LC/ESI/(QqQ)MS and LC/ESI/(orbitrap)/MSn. *Journal of Pharmaceutical and Biomedical Analysis* **2015,** 102, 215-21.

4. **D'Urso G.**, Maldini M., Pintore G., d'Aquino L., Montoro P., Pizza C. Characterisation of *Fragaria vesca* fruit from Italy following a metabolomics approach through integrated Mass Spectrometry techniques. *LWT-Food Science and Technol*ogy submitted

5. Sarais G.*,* **D'Urso G.**, Lai C., Pirisi F. M., Pizza C., Montoro P. Targeted and untargeted mass spectrometric approaches in discrimination between *Myrtus communis* cultivars from Sardinia region. *Journal of Mass Spectrometry* submitted

6. **D'Urso G.**, Sarais G., Lai C., Pirisi F. M., Pizza C., Montoro P. Metabolic Profiling of different parts of myrtle's berry coming from Sardinia. *Analytical and Bioanalytical Chemistry* submitte