Spectroscopic and multivariate data-based method to assessthe metabolomic fingerprint of Mediterranean plants

Laura Grauso
Maurizio Zotti
Wen Sun
Bruna de Falco
Virginia Lanzotti
Giuliano Bonanomi

Spectroscopic and multivariate data based method to assess the metabolomic fingerprint of Mediterranean plants

Laura Grauso[a], Maurizio Zotti[a], Wen Sun[a], Bruna de Falco[b], Virginia Lanzotti[a] and Giuliano Bonanomi[a]*

[a]: Dipartimento di Agraria, Università di Napoli Federico II, Via Università 100, I-80055 Portici, Naples, Italy

[b]: School of Science, Engineering & Technology, Division of Food & Drink, University of Abertay, Bell Street, DD1 1HG Dundee, Scotland, UK

*To whom correspondence should be addressed. Tel.: +39 0812539015; fax: +39 081

E-mail address: giuliano.bonanomi@unina.it (G. Bonanomi).

Running Title: Metabolomic Fingerprint of Mediterranean Plants

Keywords: Metabolite profile, NMR, GC-MS, PCA, Metabolites, *Dittrihia viscosa, Hedera helix.*

Short Abstract for Table of Content

This study was aimed to a rapid and versatile assessment of the metabolomic fingerprint of Mediterranean plants through an integrated approach of spectroscopic techniques, NMR and GC-MS, and Multivariate Data Analysis. The chemical profile of eight Mediterranean species, largely used as folk remedy, was analyzed. In addition, the metabolic compositions of leaves and roots were compared for each species.

ABSTRACT

Introduction – Most secondary metabolites from plants have a prominent defensive role and repellency against predators and microbial pathogens. These properties largely varies among plant species and found potential applications as biologically active compounds in medicine as well in agriculture.

Objectives – We propose a new procedure that combine different spectroscopic techniques and multivariate data analysis to determine the chemical composition and the relative amounts of each metabolites and/or each class of organic compounds. The approach was used for a rapid identification of secondary metabolites from leaf and root of eight Mediterranean plants species.

Methodology – The polar and the apolar extracts of two leaves and roots of each plant were analyzed by ^{1}H-NMR and GC-MS, respectively. Multivariate Data Analysis was used for a faster interpretation of data.

Results – The metabolic fingerprint of the Mediterranean plants, *Acanthus mollis*, *Dittrichia viscosa*, *Festuca drymeja*, *Fraxinus ornus*, *Fagus sylvatica*, *Hedera helix*, *Quercus ilex*, and *Typha latifolia*, showed a complex chemical composition, being specific for each species and plant tissue. Two alditols, mannitol and quercitol, were found in manna ash (*F. ornus*) and holm oak (*Q. ilex)* polar leaf extracts, respectively. The highest levels of aromatic compounds were found in *D. viscosa* and *T. latifolia*. Fatty acids were the predominant class of compounds in all apolar extracts under investigation. Triterpene were almost exclusively found in roots, except for holm oak, where they constitute 58% of total extract. Steroids were widespread in leaf extracts.

Conclusion – The major advantages of the proposed approach are versatility and rapidity, thus making it suitable for a fast comparison among species and plant tissue types.

Introduction

Higher plants produce a great diversity of chemical compounds, which have often a defensive role and repellence against predators and microbial pathogens.[1-3] Moreover, environmental abiotic factors like extreme temperature, light intensity and UV, shortage in water and minerals supply, as well osmotic stress can cause the accumulation of specific metabolites in plants tissues.[4]

Exploitation of plant phytochemical diversity with antimicrobial activity in traditional medicinal, as well for biological control in agriculture, has been the focus of an increasing number of studies (Fabricant and Farnsworth, 2001, Koul and Dhaliwal, 2003). [5,6] The study of medicinal plants provides a scientific basis for the popular use against infectious diseases in the modern era. [7] Although thousands of plant species have been tested for antimicrobial properties, only a small fraction of the estimated plant species has been investigated in their phytochemical composition,[8] and, so the majority of them has not been adequately evaluated. Moreover, the fractions of plant submitted to biological or pharmacological screening are even smaller.[9],[10]

The phytochemical diversity of higher plants has been previously reviewed by examining their involvement in constitutive[11] and inducible chemical defenses,[12] mechanisms of plant resistance to biotic,[13] as well to abiotic stresses,[4] and fitness cost.[14] The potential exploitation of such molecules plant antimicrobial compounds has also been evaluated and thousands of diverse natural products, involved in plant defense, have been identified including terpenoids, saponins, phenolics, phenylpropanoids, pterocarpans, stilbenes, alkaloids, glucosinolates, tiosulfinates and indoles.[1]

Indeed, higher plants had enormous potential as sources for antimicrobial drugs with reference to antibacterial and antifungal agents. However, the majority of previous studies have focused on a single or few species, studying separately different plant parts (e.g. leaves, roots, flowers, seeds etc.) However, no attempts targeted simultaneously several species and plant tissues to provide a comprehensive description of plants metabolomics. Here, to overcome the limitations of previous studies, we selected eight plant species from the Mediterranean biome (i.e. *Acanthus mollis* L., *Dittrichia viscosa* (L.) Greuter syn. *Inula viscosa* (L.) Aiton. *Fagus sylvatica* L., *Festuca drymeja*

Mert. et W. D. J. Koch, *Fraxinus ornus* L., *Hedera helix* L., *Quercus ilex* L. and *Typha latifolia* L.) having different plant traits and being widely used as folk plants.[15] We analyzed both leaf and root organs using functionally complementary and powerful organic chemistry methods like Gas Chromatography (GC-MS) and Nuclear Magnetic Resonance (NMR) Spectroscopy. Moreover, we used a multivariate data analysis approach to manage the large amount of data produced by GC-MS and NMR.

The main objectives of our study were:

(i) to describe the phytochemical composition and diversity of the eight plants;

(ii) to compare the leaf and root chemistry of selected Mediterranean plants.

Experimental

Plant materials

Eight Mediterranean plants were chosen for a metabolomic analysis through NMR and GC-MS approaches. Leaves and roots of *A. mollis*, *D. viscosa*, *F. drymeja*, *F. ornus*, *F. sylvatica*, *H. helix*, *Q. ilex*, and *T. latifolia* were collected in Cicerale (40°19′ N, 15°07′ E), an Italian municipality belonging to the "Parco Nazionale del Cilento, Vallo di Diano e Alburni", at an altitude of 250 m a.s.l.. The study site has a typical Mediterranean climate with a mean annual temperature of 16.9 °C and 1,328 mm of annual rainfall well distributed in winter, spring and fall, but with a pronounced dry summer.

Solvent and chemicals

n-Hexane and methanol were obtained from Delchimica Scientific Laboratories (Naples, Italy). Deuterium oxide (99,8 atom %D) was acquired from ARMAR Chemicals (Switzerland) and chloroform-d (99,8 atom %D) and HCl- methanol solution 1.25 M were obtained from Sigma-Aldrich

(Steinheim, Germany). Dimethyl-4-silapentane sodium sulfonate (DSS) (Merck, Darmstadt, Germany).

Extraction procedure

Leaves and roots of eight selected Mediterranean plants, collected in triplicates, were dried under controlled temperature and powdered finely with a pestle and mortar. Four grams of each sample were extracted with 20 mL of *n*-hexane under stirring for 1 h. Then, the apolar extract was filtered, evaporated, and stored at 4 °C until analysis.

The plant material was further extracted with 20 mL MeOH/H_2O (6:4) solution under stirring for 1 h, followed by centrifugation at 3000 rpm for five minutes at 25 °C. After separation, the polar extract was collected, dried by a rotary evaporator, and stored in a refrigerator at 4 °C until analysis. The following amounts of apolar extracts were obtained: *A. mollis* (33.7mg for leaves, 1.4 mg for roots), *D. viscosa* (39.6 mg for leaves, 5.9 mg for roots), *F.drymeja* (9.4 mg for leaves, 1.8 mg for roots), *F. ornus* (32.4 mg for leaves, 2.8 mg for roots), *F.sylvatica* (6.2 mg for leaves, 1.5 mg for roots), *H. helix* (5.2 mg for leaves, 19.8 mg for roots), *Q. ilex* (6.7mg for leaves, 18.3 mg for roots), *T. latyfolia* (5.4 mg for leaves, 1.7 mg for roots). The following amounts of polar extracts were obtained: *A. mollis* (255.0 mg for leaves, 58.5 mg for roots), *D. viscosa* (468.0 mg for leaves, 285.7 mg for roots), *F.drymeja* (41.6 mg for leaves, 38.4 mg for roots), *F. ornus* (348.7mg for leaves, 95.4 mg for roots), *F.sylvatica* (20.1 mg for leaves, 60.0 mg for roots), *H. helix* (278.5 mg for leaves, 151.7 mg for roots), *Q. ilex* (165.4 mg for leaves, 340.0 mg for roots), *T. latyfolia* (150.0 mg for leaves, 54.5 mg for roots).

All samples were analyzed in triplicate to ensure their reproducibility. Apolar extracts was analyzed by GC-MS, while polar extracts was analyzed by ^{1}H-NMR.

NMR Experiment

An aliquot (10 mg) of each dried polar sample was solubilized in 600 µl of deuterium oxide (99.9% D_2O) and transferred into a 5 mm NMR tube. Dimethyl-4-silapentane sodium sulfonate (DSS)

(Merck, Darmstadt, Germany), added at a concentration of 0.2 mg/mL, was used as an internal standard. The NMR spectra were recorded at 298 K on a Varian Unity Inova spectrometer operating at 500 MHz. The ^{1}H-NMR experiments were performed with 128 transients and 16K complex data point. The recycle time was set to 5 s, and a 45° pulse angle was used. Chemical shifts were referred to DSS signal (δ 0.00 ppm). All spectra were processed using iNMR program (www.inmr.net), phased and baseline corrected. In total, 48 spectra (16 plant population × 3 replicates) were acquired. Quantification was performed by signal integration relative to the internal standard, DSS. The region of the solvent peaks was excluded from the analysis. Spectral peak assignments of the detected compounds were obtained based on pure standards purchased by Sigma-Aldrich, and on combined comparison with data reported in the literature and in Human Metabolome Database (HMDB).[16,17] All spectra were manually phased, and baseline corrected.

Gas chromatography–mass spectrometry

Apolar extracts were derivatized as methyl esters before analysis by GC-MS. For this purpose, an aliquot of each apolar extract (0.5 mg) was transferred into a vial and dissolved in 1 ml of a solution of MeOH: HCl 1.25 N. The vials were vortexed and left at 50 °C overnight, then they were neutralized with NaOH 1N, dried under nitrogen, solubilized in n-hexane and analyzed by GC-MS. Chromatographic conditions are those described by de Falco *et al*.[18,19] One µl of derivatized samples were injected in a pulsed splitless mode into an Agilent-7820A GC system with 5977E MSD operating in EI mode at 70 eV. The system was equipped with a 30 m × 0.25 mm id fused-silica capillary column with 0.25 µm HP-5MS stationary phase (Agilent Technologies, UK). The injection temperature was set at 270 °C. Helium was used as carrier gas at a constant flow rate of 1 ml/min. Separation of the nonpolar extract was achieved using a temperature program of 80 °C for 1 min, then ramped at 10 °C/min to 320 °C and held for 1 min. Both chromatograms and mass spectra were evaluated using the MassHunter Qualitative Analysis B.07.00 (Agilent Technologies, CA, USA). Mass spectra of all detected compounds were compared with standard compounds and with spectra

in National Institute of Standard and Technologies library NIST MS search 2.2. Data was processed with the AMDIS (Agilent Technologies, CA, USA) software to deconvolute co-eluting peaks. Compounds were selected from the NIST list on the basis of the probability of identification (> 90%) and by comparison with literature data. The relative amounts of separated metabolites were calculated from Total Ion Chromatography (TIC) by the computerized integrator.

Multivariate data analysis

Resulting dataset from ^{1}H-NMR and GC-MS analysis was examined through Multivariate approach in order to obtain information of its underlying structure and the effect of multiple variable on the chemical differentiations between the plant species and plant tissues object of the study. Previous to apply multivariate approach each dataset was normalized. Particularly for NMR analysis, data were normalized to total area to minimize small differences and subsequently mean-centered. In detail, for GC-MS analytical method data list of compounds was aligned according by time of retention time between samples. Aligned data was utilized to build a double entry data matrix, with plant species and portion as cases in column and identified GC-MS observed metabolite as variable in row. When metabolite was not detected for a plant species or portion, 0 value was assigned.

For ^{1}H-NMR, the description of statistical analyses refers to range scaled data, in order to preserve experimental biological information. Total dataset was plotted according to PCA, in order to explain main chemical species producing differentiations among plant and root extracts. Given the high number of resonance regions and the unbalanced presence of chemical classes that are constitutively more produced with respect to other, we perform three additional PCA on different resonance regions. Resonance regions were clustered according to common chemical classes as described following: i) Aromatic/ phenolic compounds regions (from δ 10.5 to 5.5); ii) Carbohydrates regions (from δ 5.5 to 3.0) iii) Aliphatic regions (from δ 3.0 to 0.5).

Data ordination and normalization was performed by means of Excel software, while Multivariate analysis and plotting was performed in Statistica 10 software (StatSoft, Inc., Tulsa, OK).

Results and discussion

Metabolite profiling of polar extracts

An integrate spectroscopic approach combined with multivariate data analysis was applied on eight Mediterranean plants. The metabolic profile of leaves and roots was obtained to comprehensively evaluate the metabolome of each species and how its chemical composition was distributed in two compartments of each plant species. On the basis of our previous experience,[16, 17] the polar extracts were analyzed by NMR analysis, while the apolar extracts were investigated through GC-MS, because of the strong overlapping of the methylene signals in the ^{1}H-NMR spectra.

Each polar extract showed a very intricate profile, with free aliphatic and aromatic amino acids, carbohydrates, organic acids and aromatic compounds; the qualitative and quantitative metabolite profile was peculiar of each analyzed species (Figure 1).

For more convenient data interpretation, the ^{1}H-NMR spectra were divided in three regions: the aliphatic region between 0.5-3.10 ppm, the sugar region between 3.10-5.50 ppm and the aromatic region ranging from 5.50 to 8.5 ppm (Figure 2). The aliphatic region contained signal related to amino acids and organic acids. Diagnostic methyl doublets typical of isoleucine (Ile) and valine (Val), resonated at 0.91 ppm and 1.01 ppm, respectively, and the methyl triplet of leucine (Leu) at 0.95 ppm allowed their qualitative and a quantitative assignment. Moreover, doublets at δ 1.46 (J 7.0 Hz) and δ 1.32 were associated to alanine (Ala) and threonine (Thr), respectively (Table S1). The typical region of methylene groups closes to a carbonyl group in ^{1}H-NMR spectra showed a triplet at 2.98 ppm attributed to the γ-methylene protons of γ-amino butyric acid (GABA), as well as two double doublets at 2.84 and 2.94 ppm, corresponding to the diastereotopic hydrogens of asparagine (Asn). A mention is due to proline (Pro), whose recognized has been obtained by three multiplets at δ 1.99, 2.06 and 2.34, and to glutamic acid (Glu) with the typical multiplet signals at δ 2.05, 2.10 and 2.36. Pro and Glu were not always present in the studied species, but when they occurred in the plant were

present at reasonable amounts, although it was not easy to distinguish between them (Table S1). All monosaccharides and alditols were quantified by integrating the signals indicated in Table S1. Finally, the aromatic region was selected from 5.51 to 8.50, excluding three multiplet signals at 7.32, 7.36 and 7.40 ppm corresponding to phenylalanine (Phe), and two doublets at 6.80 and 7.12 ppm, corresponding to tyrosine (Tyr). Some aromatic signals were determined, as chlorogenic acid (CA) (Table S1).

The results showed that all analyzed samples have carbohydrates as major metabolites. In detail, the analysis of the leaves indicated *F. ornus* and in *Q. ilex* to contain a rather high content of monosaccharides, due to the presence of additional alditols. In the ^{1}H-NMR spectra of *F. ornus* leaves, the signals of mannitol were easily recognized by the presence of two coupled double doublets at δ 3.66 and 3.85, a double triplet at δ 3.75 and a doublet at δ 3.79. Mannitol is the major component of manna, which is produced from *Fraxinus* sp. especially under stress conditions.[20] In our study mannitol alone represented 45.1% in weight of the total metabolome of *F. ornus*. The holm oak (*Q. ilex)* contained two metabolites deriving from the shikimic acid pathway,[21] quercitol and quinic acid (QA), whose signals resonated mostly in the sugar region. This is probably the reason for the high sugar content found for this species. It has been reported that QA and quercitol are the most abundant metabolites in *Q. ilex* and in other species of *Quercus*;[22,23] their production is a reaction to biotic and osmotic stress.[24,25] Quantitative determination of QA and quercitol was not easy due to their nearness in the ^{1}H-NMR spectra; to avoid any kind of overlapping, we choose to integrate the signal at δ 1.81 for quercitol and the signal at δ 1.87 for QA. In this way, we were able to quantitate quercitol and QA which represented 18.9% and 13.9% of the all holm oak leaves polar extract, respectively. On the contrary, *D. viscosa* had the lowest amount of carbohydrates (10.4%) (Figure 3, Table S1 and S2).

The organic acid total content was almost the same in all analyzed leaves (~10%) with the exception of *T. latifolia*, *H. helix*, and *Q. ilex*. The former contained the lowest content of organic acid (2.6%) among the analyzed leaves. The latter present the highest content, reaching respectively 23.2% and 21.0%, due to the presence of QA, absent in the other analyzed plants (Figure 3, Table

S2). Moreover, the leaves of *A. mollis* showed a high amount of betaine, recognizable from a singlet at 3.25 ppm (Table S1).

The leaves of *D. viscosa*, *F. sylvatica*, and *F. drymeja* showed a high percentage of amino acids, which was partially due to the high values of glutamic acid (31.9%, 24.8% and 20.9%, respectively) (Figure 3). Aromatic compounds were particularly abundant in *D. viscosa* and *T. latifolia* leaves, followed by *F. sylvatica*, *F. ornus* and *A. mollis*. The lowest content of aromatic compounds was found in *Q. ilex* leaves (Figure 3).

Concerning root tissue, the carbohydrate content of the analyzed species was generally around 50% of the comprehensive metabolite content of polar extract, with some notable exceptions. *A. mollis* extract had 80.6% of sugar content, due to the presence of several sugar residues, from which raffinose (Raff) was predominant with 36.5% of total extract. *F. ornus* contained 62.7% of sugars, the most abundant being sucrose (Sucr). Moreover, particularly low is the content of amino acids and organic acids in these species. On the contrary, *D. viscosa* and *T. latifolia* had the lowest percentage of carbohydrate content and the highest percentage of aromatic compounds (Figure 3).

Metabolite profiling of apolar extracts

NMR spectra of the apolar extracts of each plant showed a chemical composition almost exclusively amenable to fatty acids. The overlapping of some signals, such as the methylene groups close to a carbonyl group, did not allow to distinguish and quantify all fatty acids of apolar extracts. For this reason, we used gas chromatography coupled to mass spectrometry (GC-MS) with increasing separation capability, allowing to determine the quali-quantitative profile of the studied Mediterranean plants. In this way, it was possible to identify single fatty acids on the basis of their molecular weight.

The GC-MS data of the species under investigation allowed to characterize 60 metabolites, belonging to several classes of organic compounds. The count of metabolites extracted was higher in the root extracts, reaching the maximum value in *F. ornus*, which also contained the highest

variability of organic compounds (Figure 4). Fatty acids were ubiquitous metabolites in all analyzed samples at variable percentage, ranging from 24.5% in *D. viscosa* to 100% in *T. latifolia.* Moreover, the fatty acid profile with relative amount depend on the species analyzed with palmitic and oleic acids always present (Table S3). In particular, palmitic acid showed an elevate concentration range from a maximum value of 52.7% in *A. mollis* roots to a minimum value of 1.0% in *T. latyfolia* leaves. Similarly, cerotic acid (C26:0) has been determined in most analyzed plants at very different concentration (44.6% in *A. mollis* leaves, 30.1% in *T. latifolia* leaves and 1.9% in roots) while in others (e.g. *H. helix*) it was absent. Unsaturated fatty acids were determined in the roots of *F. ornus* and *H. helix* and identified as C16:1, C18:2, C18:1 and C20:1, the last found only in manna ash (Table S3).

The apolar extracts of leaves contained sesquiterpenes and alkanes in addition to fatty acids. The detected heptacosane, nonacosane and henicosane alkanes were determined in *F. ornus* and *F. sylvatica*, representing 48.9% and 43.9% of total extract, and in *D. viscosa*, where they were less abundant (Table S3). Sesquiterpenes were almost exclusively found in *D. viscosa*, where they represented 59.4% of the total extract and in the roots of *A. mollis* and *F. ornus* roots, present at smaller amounts (Table S3). Triterpenes were widespread in all studied roots extracts, although holm oak leaves contained a high level of these organic compounds (57.9% of total extract). Triterpenes in roots were present at variable concentrations and showed the highest content in *A. mollis*, and *F. drymeja*, *D. viscosa*, *Q. ilex*, and *T. latifolia*, the latter species containing 38.3% and 28.4% of lupan-3-one and friedelan 3-one, respectively (Table S3).

Finally, steroids were found in *A. mollis* (16.7%), *F. drymeja* (41.2%) and *H. helix* (27.3%) leaves, and *F. sylvatica* (14.1%) and *Q. ilex* roots (41.0%), where the most representative steroid was 3b,5a,6b-Cholestanetriol (Table S3).

Multivariate Data Analysis

Coupled with metabolic profiling, the multivariate approach ordinated plant species according to their respective metabolic characteristics. Through PCA we obtained a general view of the underling structure of the data. The principal components were displayed as a set of scores (PC), which highlights clustering or outliers, and a set of loadings (p), which emphasizes the influence of input variables on PC. The multivariate methodology was applied both for polar and apolar extracts and, particularly for data originated from NMR, was tested for the ordination of plant according to chemical characteristics present in three different regions putatively assigned to aromatic/ phenolic compounds (from δ 10.5 to 5.5), carbohydrates regions (from δ 5.5 to 3.0), and aliphatic compounds (from δ 3.0 to 0.5).

In PCA performed for the totality of the regions from ^{1}H-NMR spectra, the first 2 components explains the 71.4% of the variance among the samples (PC1 65.6 and PC2 5.8%). Results are showed in figure 4A and 4B for loadings and score plots, respectively. In a general view, we observed a marked ordination of loadings values according to the respective plant species. Inversely, metabolic profile of the plants does not discriminate among plant portion from which metabolite was extracted. The general variation among plant species was triggered by carbohydrates, while aliphatic and aromatic/phenolics regions has a decreased discriminant power. This is likely explained by the normal attitude to accumulate carbohydrates as nutrient source from photosynthetic pathways. [26]Given this, carbohydrates mediate the unidirectional disposition of the samples in its correspondent area showing a generalized positive association of all the samples with carbohydrate signals. However, peculiar number of specific metabolites generate distinctive disposition among plant species. For instance, *D. viscosa* leaves extracts is characterized by the presence of betaine, that also appears to be responsible of the separation of *A. mollis* leaves extracts from the other plant extracts. So far, mannitol signals are majorly associated to the well-known manna producer species *F. ornus* and the quercitol to the oak *Q. ilex*. In both the cases, the metabolite disposition appears to be few distinctive by the point of view of the plant organs in which the metabolite was extracted. More generally, *H. helix*, *F. ornus*, *F. drymeja*, *F. sylvatica*, *D. viscosa* for roots and *F. ornus*, *F. drymeja*, *F. sylvatica* for leaves

associate with the aforementioned mannitol and glutamic acid, sucrose, fructose, shikimic acid and raffinose.

To avoid the hiding action of carbohydrates on the other spectral regions we analyzed these in separate way with the same multivariate approach. Figure 6A and 6C showed the PCA ordination of different extracts according to their chemical composition. The PCA explained totally the 62.0% of the variance in the sample (PC1 51.7% and PC2 10.3%). In this case, we observed a marked differentiation of *A. mollis* and *H. helix* root and leaf extracts with respect to other species mainly operated by the higher content of fumaric acid. Intermediate position is instead acquired by *F. drymeja* leaves extract associated by higher content of tyrosine. The other species aggregates in same directional ordination that is given by the similarity of the spectral regions between 6.96 and 6.83 ppm. For those regions, we unassigned the signals given the high level of uncertainty. In addition, residuals signals from carbohydrates and amino acidic compounds interfere in the interpretation of the spectra limiting our multivariate approach. For the PCA of carbohydrate region (Figure 6C and 6D), is observed a specular disposition of the data to those of the comprehensive PCA (Figure 5A and 5C). Indeed, the PCA of carbohydrate region explain 72.6% of the variance with respect to the 71.4% explained by the PCA of the overall dataset. For the PCA of the aliphatic region, lower level of explained variance was reported (Figure 6E and 6F). Nonetheless, *A. mollis* and *Q. ilex* leaves and *D. viscosa* roots differentiate for quinic acid and proline contents rather than other species that differentiate for the contents of threonine and signals of rhamnose, acetic acid/GABA and residual signals from polar portion of fatty acids.

PCA from polar extract showed lower level of explained variance (Total of PC 46.6%, Figure 5C and 5D). However, we observed a clearer segregation of root extract with respect leaves extracts, that make only exception for *F. ornus* and *Q. ilex* leaves. Interestingly, *Q. ilex* root extracts dispose with leaves extract of other species in complete opposite way respect the general behaviors in the samples. So far, root chemical differentiation from apolar extracts is given by the presence of palmitic acid, 2,6,6,9,2’,6’,6’,9’-Octamethyl-[8,8’]bi[tricyclo[5.4.0.0(2,9)]undecyl], linoleic acid and behenic

acids. Oppositely, leaves are higher in contents of cerotic acid, montanic acid, eptacosane, lignoceric acid, cholestane and arachidic acid.

Conclusions

The large number of biological questions on plant metabolism requires that answers should be sought using the most versatile techniques available. Our goal was the development of a method capable to detect most classes of organic compounds.[27] Unfortunately, the huge variety of chemical compounds found in plants did not allow the use of a single technique to this aim.

The approach we developed to evaluate the chemical composition of eight Mediterranean plant species was based on spectroscopic techniques and multivariate data analysis. It was appropriate for a fast and comprehensive analysis of primary and secondary metabolites, allowing the determination of a metabolomic fingerprint of each species and the evaluation of the different distribution of the metabolites in two parts (leaves and roots) of the plant.

Q. ilex and *F. ornus* contained large amounts of specific metabolites, quinic acid, quercitol and mannitol, usually produced from plants during stress conditions. Besides being involved in osmotic stress, quercitol has been recently used as a building block in the synthetic strategy for antidiabetic compounds. [28] *D. viscosa* was characterized by a high content of aromatic compounds at the expense of carbohydrate production in the polar fraction, and a considerable content of sesquiterpenes in the apolar fraction. The separation of *A. mollis* from the other species was due to the presence of betaine and sucrose in leaves and raffinose in roots.

Moreover, fatty acids were present in all the analyzed species, especially in *T. latifolia* leaves and in *A. mollis* and *F. drymeja* roots. Fatty acids were almost exclusively present as saturated fatty acids.

In conclusion, the approach we developed proved to be suitable for a rapid investigation of different plant species containing a wild range of organic molecules.

Acknowledgements

NMR spectra and GC-MS analyses were recorded at the “Laboratorio di Analisi strumentale”, Dipartimento di Farmacia, Università di Napoli Federico II; the assistance of Paolo Luciano is gratefully acknowledged.

Notes

The authors declare no competing financial interest.

References

1. Dixon RA. Natural products and plant disease resistance. *Nature* 2001; 411: 843–847.

2. Hartmann T .The lost of origin of chemical ecology in the late 19th century. *Proc. Natl. Acad. Sci. USA* 2008; 105: 4541–4546.

3. Springob K, Kutchan TM. Introduction to the different classes of natural products. In: *Plant-derived Natural Products.* Springer Science; 2009. 3 p.

4. Akula R, Ravishankar, GA. Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signal Behav*. 2011; 6(11): 1720-1731.

5. Fabricant DS, Farnsworth NR. The value of plants used in traditional medicine for drug discovery. *Environ Health Perspect* 2001; 109 (suppl 1): 69-75.

6. Koul O, Dhaliwal GS. *Phytochemical biopesticides*. CRC Press 2003 (Vol. 1).

7. Gibbons S. Plants as a source of bacterial resistance modulators and anti-infective agents *Phytochem Rev* 2005; 4(1): 63-78.

8. Rates SM. Plants as source of drugs. *Toxicon* 2001; 39: 603-613.

9. Hamburger M, Hostettmann K. Bioactivity in plants: the link between phytochemistry and medicine. *Phytochem* 1991; 30 (12): 3864-3874.

10. Hostettmann K, Potterat O, Wolfender J L. The potential of higher plants as a source of new drugs. *Chimia* 1998; 52(1-2): 10-17.

11. Wittstock U, Gershenzon J Constitutive plant toxins and their role in defense against herbivores and pathogens. *Curr Opin Plant Biol* 2002; 5:1–8

12. Hammerschmidt R. Phytoalexins: what have we learned after 60 years? *Annu Rev Phytopathol* 1999; 37: 285–306

13. Morrisey JP, Osbourn AE. Fungal resistance to plant antibiotics as a mechanism of pathogenesis. *Microbiol Mol Biol R* 1999; 63: 708–724.

14. Heil M. Ecological costs of induced resistance. *Curr Opin Plant Biol* 2002; 5: 1–6

15. Motti R, Ippolito F, Bonanomi G. Folk phytotherapy in paediatric health care in Central and Southern Italy: a review. *Human Ecology* 2018; 46(4): 573-585.

16. de Falco B, Incerti G, Pepe R, Amato M, Lanzotti V. Metabolomic fingerprinting of romaneschi globe artichokes by nmr spectroscopy and multivariate data analysis: metabolomics of artichokes by nmr and chemometrics. *Phytochem Anal* 2016; 27: 304–314.

17. de Falco B, Incerti G, Bochicchio R, Phillips, TD, Amato M, Lanzotti, V. Metabolomic analysis of *Salvia hispanica* seeds using NMR spectroscopy and multivariate data analysis. *Ind Crops Prod* 2017; 99: 86-96.

18. de Falco B, Fiore A, Rossi R, Amato M, Lanzotti V. Metabolomics driven analysis by UAEGC-MS and antioxidant activity of chia (*Salvia hispanica* L.) commercial and mutant seeds. *Food Chem* 2018; 254: 137–143.

19. de Falco B, Fiore A, Bochicchio R, Amato M, Lanzotti V. Metabolomic analysis by UAE-GC MS and antioxidant activity of *Salvia hispanica* (L.) seeds grown under different irrigation regimes. *Ind Crops Prod* 2018; 112: 584–592.

20. Stoop JMH, Williamson JD, Mason P. Mannitol metabolism in plants: a method for coping with stress. *Trends Plant Sc*i 1996; 1(5): 139-144.

21. Wilson DJ, Patton S, Florova G, Hale V, Reynolds KA (1998) The shikimic acid pathway and polyketide biosynthesis. *J Ind Microbiol Biotechnol* 1998; 20: 299–303.

22. Sardans J, Gargallo-Garriga A, Pérez-Trujillo M, Parella TJ, Seco R, Filella I, Penuelas J.). Metabolic responses of Q uercus ilex seedlings to wounding analysed with nuclear magnetic resonance profiling. *Plant Biol.* 2014; 16(2): 395-403.

23. Passarinho JAP, Lamosa P, Baeta JP, Santos H, Ricardo CPP. Annual changes in the concentration of minerals and organic compounds of *Quercus suber* leaves. *Physiol Plant.* 2006; 127: 100–110.

24. Koskimaki JJ, Hokkanen J, Jaakola L, Suorsa M, Tolonen A, Mattila S, Pirttila AM, Hohtola A. Flavonoid biosynthesis and degradation playa role in early defence responses of bilberry (*Vaccinium myrtilus*) against biotic stress. *Eur J Plant Pathol* 2009; 125: 629–640.

25. Adams MA, Richter A, Hill AK, Colmer TD. Salt tolerance in *Eucalyptu*s spp.: identity and response of putative osmolytes. *Plant Cell Environ* 2005; 28: 772–787.

26. Heldt HW, Piechulla B, Heldt F. *Plant Biochemistry*. Translation of the 4th German edition. 2011

27. de Falco B, Lanzotti V. NMR spectroscopy and mass spectrometry in metabolomics analysis of Salvia. *Phytochem Rev* 2018; *17*(5): 951-972.

28. Worawalai W, Sompornpisut P, Wacharasindhu S, Phuwapraisirisan P. Quercitol: from a taxonomic marker of the genus *Quercus* to a versatile chiral building block of antidiabetic agents. *J Agric Food Chem*. 2018; 66(23 5741-5745.

FIGURE CAPTIONS

Figure 1. ^{1}H-NMR at 500 MHz in D_2O of leaves (L) and roots (R) of Mediterranean species.

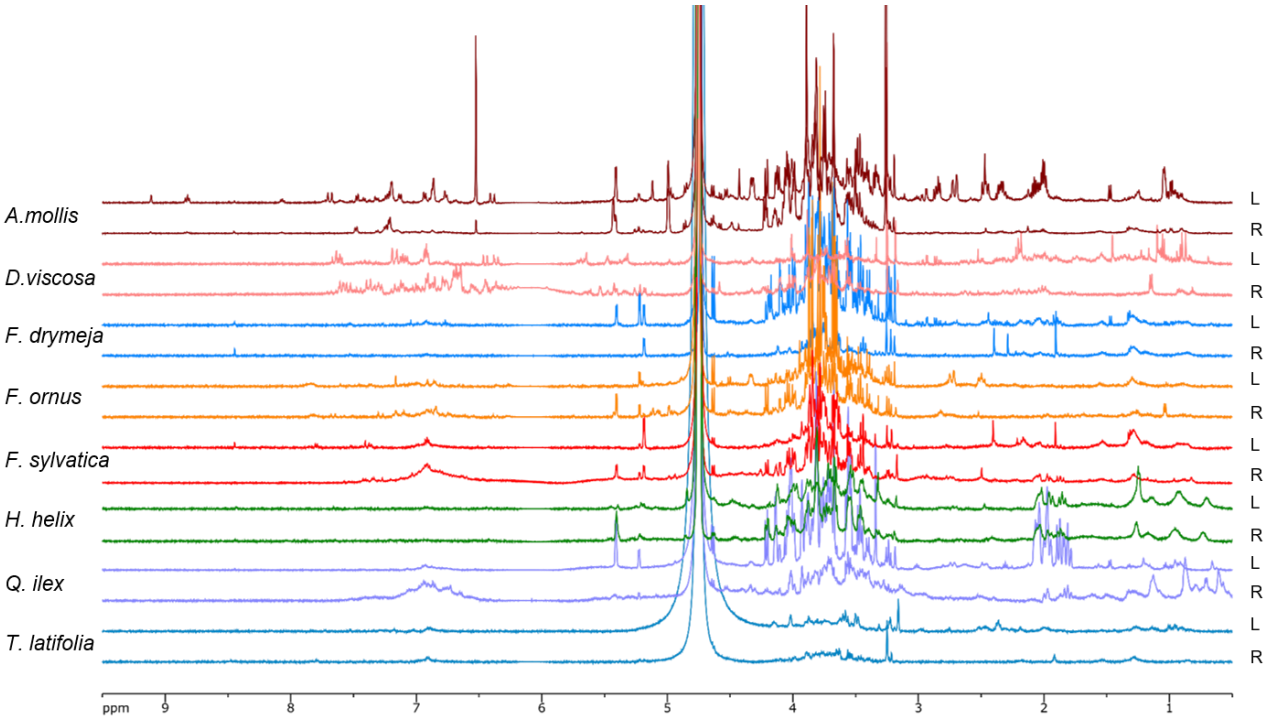
Figure 2. ^{1}H-NMR spectra at 500 MHz in D_2O of *A. mollis* and *Q. ilex* leaves (L) and roots (R).

Figure 3. Heat-map of the relative concentration (%) of metabolites in leave and root polar extract of each species.

Figure 4. a) Total count of organic compounds in apolar leaf (left) and root (right) extracts, and; b) relative concentration (%) for class of organic compounds in apolar leaf (left) and root (right) extracts.

Figure 5. Principal component analysis (PCA) ordination of eight Mediterranean plant leaves and roots based on ^{1}H-NMR resonance spectra from polar (A and B) and apolar (B and C) fractions. A and C: variable loadings; B and D: factorial scores of resonance intervals of 0.01 ppm and retention time value. Explained variance of principal components is reported on the axis labels. Plants in loading plots are numbered as: 1. *A. mollis*, 2. *D. viscosa*, 3. *F. drimejia*, 4. *F. ornus*, 5. *F. sylvatica*, 6. *H. helix*,7. *Q. ilex*, 8. *T. latifolia*.

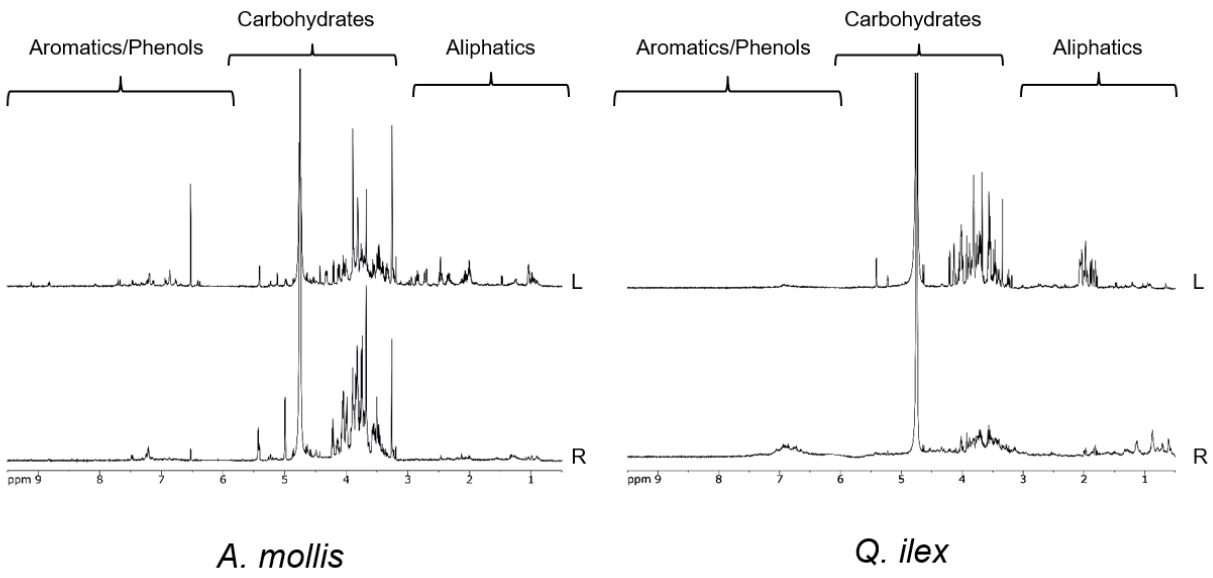
Figure 6. Principal component analysis (PCA) ordination of ^{1}H-NMR resonance intervals: (A and B) from δ 10.5 to 5.5; (C and D) from δ 5.5 to 3.0; (E and F) from δ 3.0 to 0.5. Left: variable loadings; right: factorial scores of resonance intervals of 0.01 ppm. Explained variance of principal components is reported on the axis labels. Plants in loading plots are numbered as: 1. *A. mollis*, 2. *D. viscosa*, 3. *F. drimejia*, 4. *F. ornus*, 5. *F. sylvatica*, 6. *H. helix*, 7. *Q. ilex*, 8. *T. latifolia*.



392

393 Figure 1. ^{1}H-NMR at 500 MHz in D_2O of leaves (L) and roots (R) of Mediterranean species.

394



395

396 Figure 2. ^{1}H-NMR spectra at 500 MHz in D_2O of *A. mollis* and *Q. ilex* leaves (L) and roots (R).

397

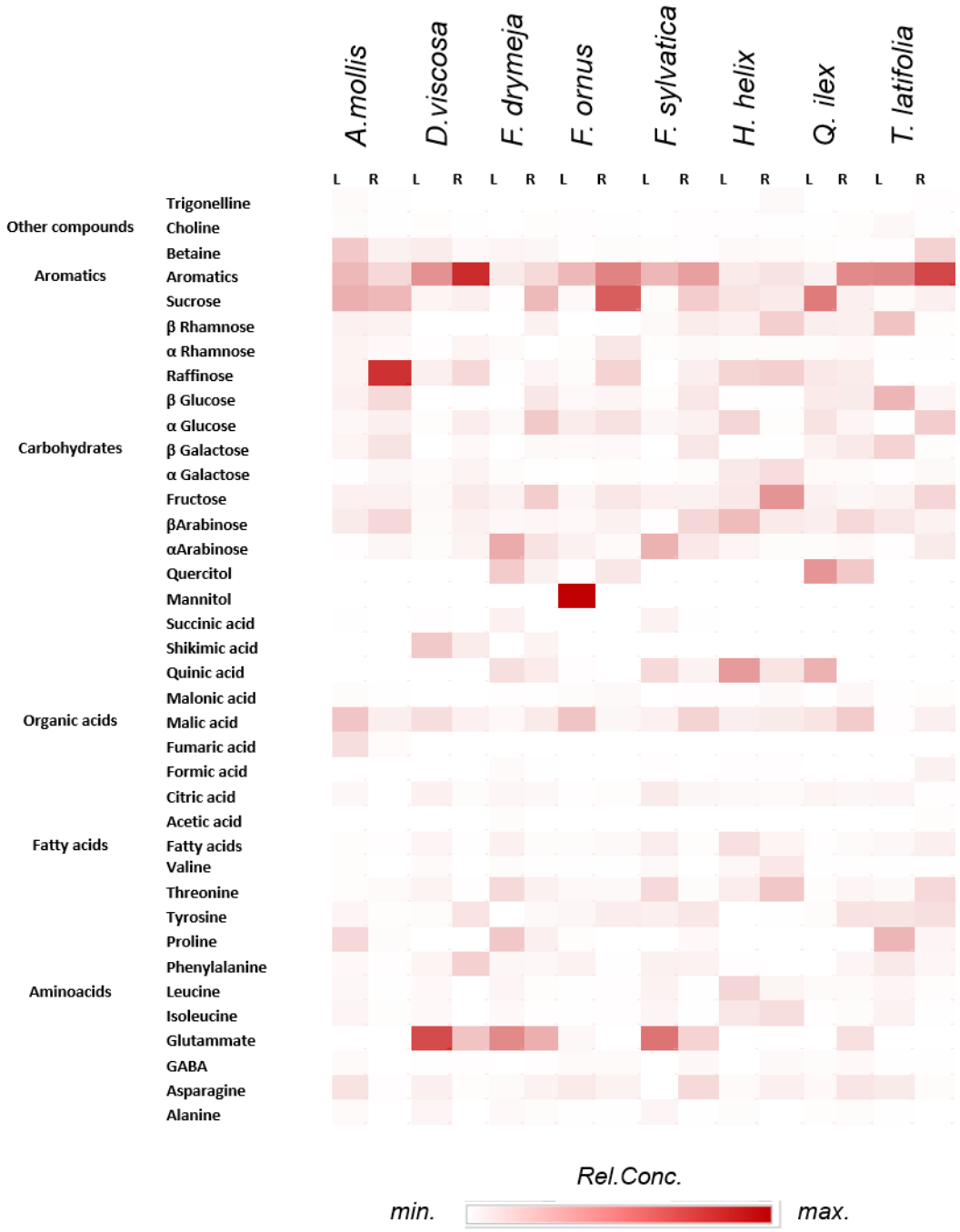


Figure 3. Heat-map of the relative concentration (%) of metabolites in leave and root polar extract of each species.

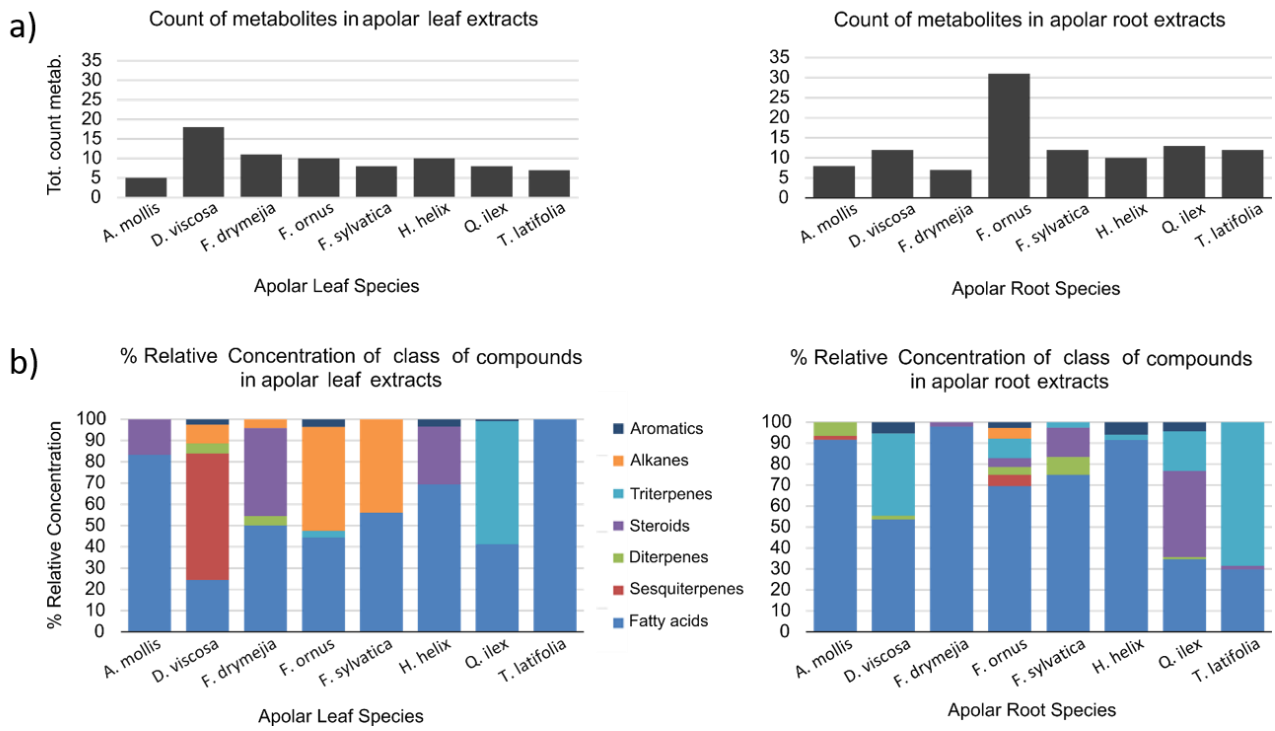


Figure 4**.** a) Total count of organic compounds in apolar leaf (left) and root (right) extracts; b) Relative concentration (%) for class of organic compounds in apolar leaf (left) and root (right) extracts

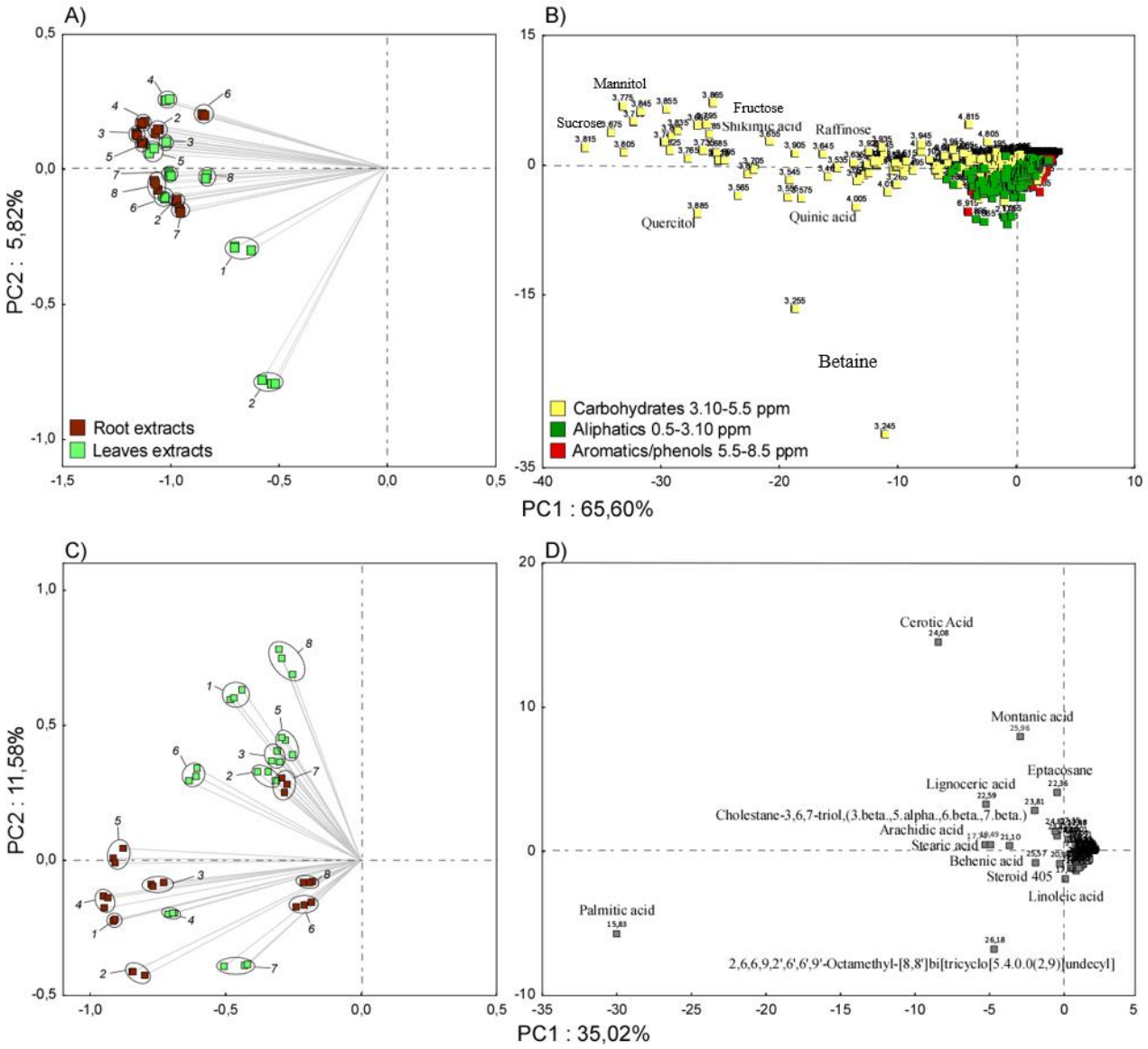


Figure 5. Principal component analysis (PCA) ordination of eight Mediterranean plant leaves and roots based on ^{1}H-NMR resonance spectra from polar (A and B) and apolar (B and C) fractions. A and C: variable loadings;B and D: factorial scores of resonance intervals of 0.01 ppm and retention time value. Explained variance of principal components is reported on the axis labels. Plants in loading plots are numbered as: 1. *A. mollis*, 2. *D. viscosa*, 3. *F. drimeja*, 4. *F. ornus*, 5. *F. sylvatica*, 6. *H. helix*,7. *Q. ilex*, 8. *T. latifolia*.

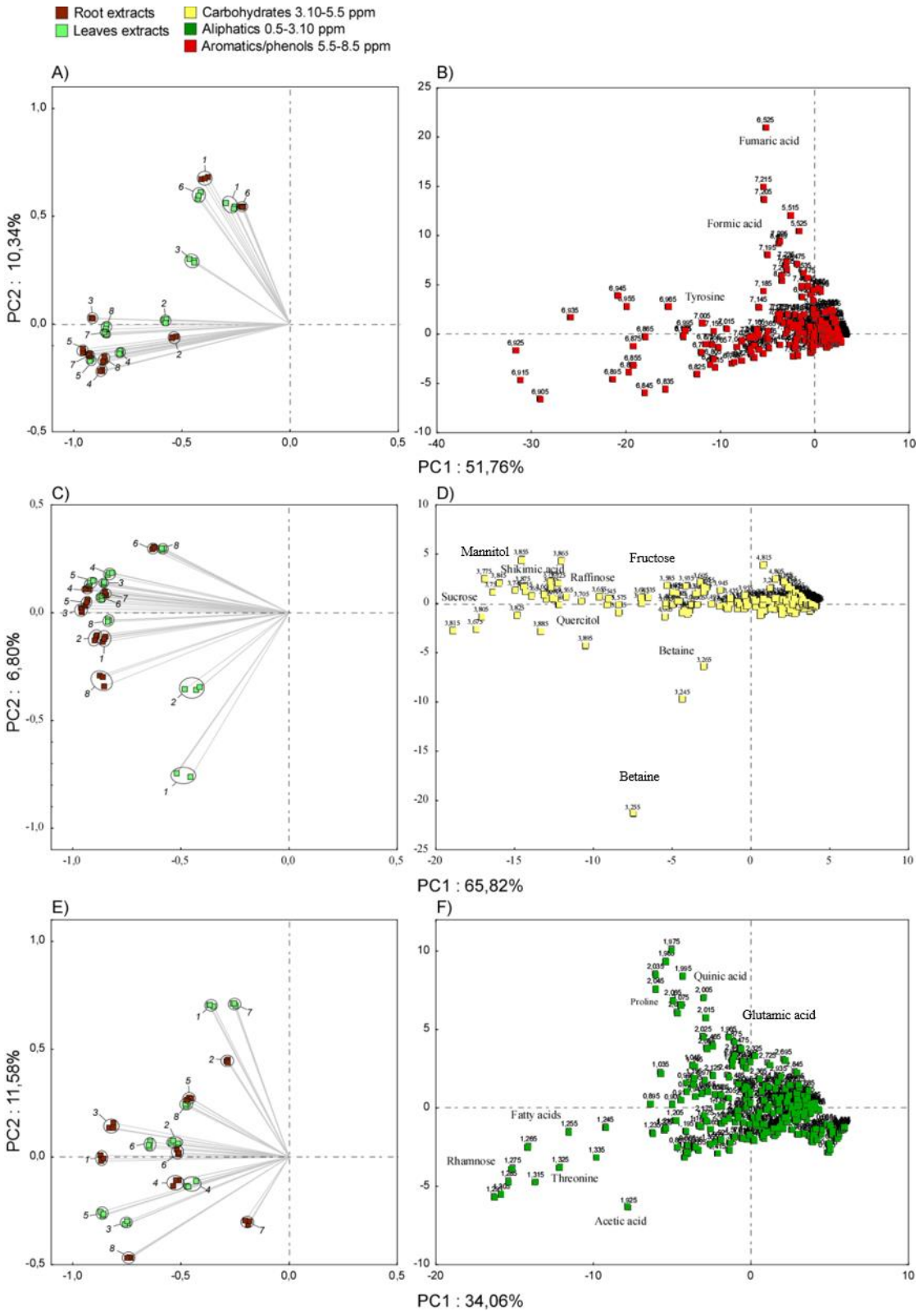


Figure 6. Principal component analysis (PCA) ordination of ^{1}H-NMR resonance intervals: (A and B) from δ 10.5 to 5.5; (C and D) from δ 5.5 to 3.0; (E and F) from δ 3.0 to 0.5. Left: variable loadings; right: factorial scores of resonance intervals of 0.01 ppm. Explained variance of principal components is reported on the axis labels. Plants in loading plots are numbered as: 1. *A. mollis*, 2. *D. viscosa*, 3. *F. drimeja*, 4. *F. ornus*, 5. *F. sylvatica*, 6. *H. helix*, 7. *Q. ilex*, 8. *T. latifolia*.