ORIGINAL ARTICLE

CrossMark

# Feasibility of UV–VIS–Fluorescence spectroscopy combined with pattern recognition techniques to authenticate a new category of plant food supplements

Raffaella Boggia[1] · Federica Turrini[1] · Marco Anselmo[1] · Paola Zunin[1] ·
Dario Donno[2] · Gabriele L. Beccaro[2]

**Abstract** Bud extracts, named also "gemmoderivatives", are a new category of natural products, obtained macerating meristematic fresh tissues of trees and plants. In the European Community these botanical remedies are classified as plant food supplements. Nowadays these products are still poorly studied, even if they are widely used and commercialized. Several analytical tools for the quality control of these very expensive supplements are urgently needed in order to avoid mislabelling and frauds. In fact, besides the usual quality controls common to the other botanical dietary supplements, these extracts should be checked in order to quickly detect if the cheaper adult parts of the plants are deceptively used in place of the corresponding buds whose harvest-period and production are extremely limited. This study aims to provide a screening analytical method based on UV–VIS–Fluorescence spectroscopy coupled to multivariate analysis for a rapid, inexpensive and non-destructive quality control of these products.

**Keywords** Plant food supplements · Bud extracts · UV–VIS spectroscopy · Fluorescence spectroscopy · Multivariate analysis

✉ Raffaella Boggia
  boggia@difar.unige.it; boggia@dictfa.unige.it

[1] Department of Pharmacy, University of Genoa, Via Brigata Salerno, 13, 16147 Genoa, Italy

[2] Department of Agriculture, Forestry and Food Science, University of Torino, Largo Braccini, 2, 10095 Grugliasco, Turin, Italy

## Introduction

Bud extracts are a new category of natural products, obtained macerating embryonic tissues (named also gemmoderivatives or embrioextracts) of trees and plants, as buds, young sprouts, and young roots (Amaral et al. 2004; Rates 2001). In the European Community these botanical remedies are classified as plant food supplements (Calixto 2000; Gulati and Berry Ottaway 2006; Knoss and Chinou 2012; Konik et al. 2011; Dell'Agli et al. 2013).

The raw material is fresh, usually taken in the late Winter or/and in the early Spring, at the peak time of the tree or shrub's annual germination, in order to capture the different nutrients, vitamins, plant hormones and enzymes that are released during this process, and which in some cases are only present in the plant at this time (i.e. plant hormones: auxins, gibberellins, cytokines, ethylene, abscisic acid) (Espín et al. 2007; Scalbert et al. 2005; Bosco et al. 2013). The quality of these preparations can be defined by the genotype of the considered species and varieties, the environmental characteristics of the sampling-sites, the phenological stage of the buds and the applied agrotechniques (Donno et al. 2014a, b; Vegvari et al. 2008; Dvaranauskaitė et al. 2009).

There are some methods to prepare the extracts of plant embryonic tissues, usually as Glyceric Macerates (GM) or as Mother Tinctures (MT), macerating the whole embryonic tissues in a mixture of water/glycerine/ethanol (96%) and water/ethanol (96%), respectively.

Nowadays these products are still poorly studied, even if the use of buds for healthy purposes is very ancient and gemmotherapy, a fast emerging branch of complementary medicine, is presently undergoing considerable development on the market (Fowler 2006; Gurib-Fakim 2006; Cordell 2011; Donno et al. 2012).

Because of the increasing interest in these economically-valuable products, there is a demand for an efficient quality control to ensure the authenticity of their botanical source and content in order to avoid mislabelling and frauds (Chandra et al. 2001; Donno et al. 2013, 2016). Relevant cases of substitutions, changes or adulterations are reported (Kesting et al. 2010; Nicoletti 2012).

Since a quality control tool should be simple, viable, comprehensible and low cost, a comprehensive spectroscopic fingerprint method has been proposed without a chromatographic separation of the phytocomplexes. Even if a non-selective qualitative tool is not exhaustive (Marston 2007), it is non-destructive, inexpensive and time-saving (Lian et al. 2004; Ovalle-Magallanes et al. 2013).

Thus this study proposes a high speed and easy-to-use shortcut based on UV–VIS spectroscopy and chemometrics recently proposed by some of the authors (Boggia et al. 2013) in order to highlight different spectral trends depending on the different botanical origin of the extracts and to evaluate its ability to quickly screen the possible adulteration of gemmoderivatives with adult parts derivatives of the same plant species too. In addition, a fluorimetric fingerprint has been proposed in order to check the results (Reid et al. 2006; Sadeckà and Tòthovà 2007).

Thus the samples were described by a vector of UV–VIS absorbance or Fluorescence emissions, which can be considered as multivariate non-targeted signals of the corresponding sample. The analysis of the spectral data was performed using Principal Component Analysis (PCA), as unsupervised patter recognition technique in order to extract useful analytical information, to group the samples with similar characteristics and to perform a dimensionality reduction using a limited number of significant PCs as new variables to describe the samples (Donno et al. 2014b; Gad et al. 2013; Zhu et al. 2014). PCA has been preferred respect to other unsupervised pattern recognition techniques such as clustering, since it allows not only to identify small subsets of similar samples, but also to perform a dimensionality reduction of the data set by projecting the original high-dimensional data into the low-dimensional linear subspace spanned by the leading eigenvectors of the data's covariance matrix. It is a way to compress a set of signals (i.e. spectroscopic fingerprints) allowing to subsequently use a limited number of significant PCs as new variables to describe the samples. The final goal was to authenticate this new category of plant food supplements using mathematical models, built with the measured spectroscopic variables, in order to verify the trueness of what declared on their label. The chemometric techniques that built such models are the class modeling techniques (CMTs) (Esteban-Díez et al. 2007; Marini et al. 2006). CMTs allow to model each class separately whose bounderies separate a particular class from the rest of the hyperspace (Oliveri and Downey 2012). Then two of the most commonly used class-modeling tools in chemometrics (Massart et al. 1998) were investigated: UNEQ (unequal class space) and SIMCA (soft independent modelling of class analogy) in order to build a model of the studied categories (i.e. "embryonic phenological stage" category or "adult phenological stage category"). The two different class-modelling techniques (SIMCA, UNEQ) have been used to evaluate and compare their performances in highlighting the differences between gemmoderivatives and extracts coming from plants in adult phenological stage. Models were evaluated on a separate validation sample set (external evaluation set). UNEQ and SIMCA are probabilistic and distance-based modelling techniques, respectively. Both UNEQ and SIMCA techniques consider each class separately.

UNEQ is the modelling version of Quadratic Discriminant Analysis (QDA) proposed by Derde and Massart (1986). It is based on the assumption of multivariate normality for each class population and each class model is represented by the class centroid and the category space is defined on the basis of the Mahalanobis distance from this barycenter, corresponding to a desired confidence level. In the present paper, the class spaces have been built as the 95% confidence level hyperellipsoids around each centroid. With respect to SIMCA, UNEQ requires the data matrix to present a specific ratio between the number of samples and the number of variables in each category ($N_{samples}/N_{vars} > 3$) so a variable reduction step was necessary. PCA has been used to reduce the number of variables: a limited number of significant PCs scores have been used as new variables to describe the samples.

SIMCA is a powerful distance-based method introduced by Wold and Sjöström in 1977. SIMCA models are based on PCs, which are, by definition, the orthogonal directions of maximum variance in a multivariate data space; boundaries of the SIMCA models are determined by a critical distance obtained using Fisher statistics (Wold and Sjöström 1977).

The extracts characterization was performed monitoring the chain production from harvesting to packaging of eight different vegetable species. The eight different vegetable extracts at different phenological stages were initially studied (first set of samples numbered 1–23 in Table 1); then the attention was focalized on three of them considering commercial samples too (second set of samples: numbered 24–35 in Table 1).

## Materials and methods

### Raw materials

Several plant leaves (*Acer pseudoplatanus*, *Aesculus hippocastanum*, *Fagus sylvatica*, *Ribes nigrum*, *Rosa canina*, *Salix caprea*, *Tilia tomentosa*, *Vitis vinifera* spp.) were

**Table 1** First set of samples (1–23) and second set of samples (24–35)

| | Identification Code | Vegetable species | Family | Order | Part | Product |
|---|---|---|---|---|---|---|
| 1 | APALET | *Acer pseudoplatanus* | *Sapindaceae* | *Sapindales* | ADULT | MT |
| 2 | APGEET | *Acer pseudoplatanus* | *Sapindaceae* | *Sapindales* | EMBRYONIC (bud) | MT |
| 3 | APGGET[a] | *Acer pseudoplatanus* | *Sapindaceae* | *Sapindales* | EMBRYONIC (young sprout) | MT |
| 4 | AHALET[a] | *Aesculus hippocastanum* | *Sapindaceae* | *Sapindales* | ADULT | MT |
| 5 | AHGEET | *Aesculus hippocastanum* | *Sapindaceae* | *Sapindales* | EMBRYONIC (bud) | MT |
| 6 | FSALET[a] | *Fagus sylvatica* | *Fagaceae* | *Fagales* | ADULT | MT |
| 7 | FSGEET[a] | *Fagus sylvatica* | *Fagaceae* | *Fagales* | EMBRYONIC (bud) | MT |
| 8 | RN1GEET | *Ribes nigrum* | *Grossulariaceae* | *Saxifragales* | EMBRYONIC (bud) | MT |
| 9 | RN2GEET | *Ribes nigrum* | *Grossulariaceae* | *Saxifragales* | EMBRYONIC (bud) | MT |
| 10 | RN3GEET | *Ribes nigrum* | *Grossulariaceae* | *Saxifragales* | EMBRYONIC (bud) | MT |
| 11 | RN4GEET | *Ribes nigrum* | *Grossulariaceae* | *Saxifragales* | EMBRYONIC (bud) | MT |
| 12 | RNALET[a] | *Ribes nigrum* | *Grossulariaceae* | *Saxifragales* | ADULT | MT |
| 13 | RCALET[a] | *Rosa canina* | *Rosaceae* | *Rosales* | ADULT | MT |
| 14 | RCGEET[a] | *Rosa canina* | *Rosaceae* | *Rosales* | EMBRYONIC (bud) | MT |
| 15 | RCGGET[a] | *Rosa canina* | *Rosaceae* | *Rosales* | EMBRYONIC (young sprout) | MT |
| 16 | SCALET[a] | *Salix caprea* | *Salicaceae* | *Salicales* | ADULT | MT |
| 17 | SCGEET | *Salix caprea* | *Salicaceae* | *Salicales* | EMBRYONIC (bud) | MT |
| 18 | TTALET[a] | *Tilia tomentosa* | *Malvaceae* | *Malvales* | ADULT | MT |
| 19 | TTGEET | *Tilia tomentosa* | *Malvaceae* | *Malvales* | EMBRYONIC (bud) | MT |
| 20 | TTGGET | *Tilia tomentosa* | *Malvaceae* | *Malvales* | EMBRYONIC (young sprout) | MT |
| 21 | TTGQET | *Tilia tomentosa* | *Malvaceae* | *Malvales* | EMBRYONIC (quiescent bud) | MT |
| 22 | VVALET | *Vitis vinifera* | *Vitaceae* | *Vitales* | ADULT | MT |
| 23 | VVGEET[a] | *Vitis vinifera* | *Vitaceae* | *Vitales* | EMBRYONIC (bud) | MT |
| 24 | RNMGAGR | *Ribes nigrum* | *Grossulariaceae* | *Saxifragales* | EMBRYONIC | GM homemade sample |
| 25 | RNMGPM | *Ribes nigrum* | *Grossulariaceae* | *Saxifragales* | EMBRYONIC | GM commercial sample |
| 26 | RN1ADCO | *Ribes nigrum* | *Grossulariaceae* | *Saxifragales* | ADULT | MT commercial sample |
| 27 | TTMGAGR | *Tilia tomentosa* | *Malvaceae* | *Malvales* | EMBRYONIC | GM homemade sample |
| 28 | TTCFMMG | *Tilia tomentosa* | *Malvaceae* | *Malvales* | EMBRYONIC | GM commercial sample |
| 29 | TTCND | *Tilia tomentosa* | *Malvaceae* | *Malvales* | EMBRYONIC | GM commercial sample |
| 30 | TTCNOMG | *Tilia tomentosa* | *Malvaceae* | *Malvales* | EMBRYONIC | GM commercial sample |
| 31 | TTCTM | *Tilia tomentosa* | *Malvaceae* | *Malvales* | ADULT | MT commercial sample |
| 32 | VVMG | *Vitis vinifera* | *Vitaceae* | *Vitales* | EMBRYONIC | GM homemade sample |
| 33 | VVCBMG | *Vitis vinifera* | *Vitaceae* | *Vitales* | EMBRYONIC | GM commercial sample |
| 34 | VVCBTM | *Vitis vinifera* | *Vitaceae* | *Vitales* | ADULT | MT commercial sample |
| 35 | VVCOTM | *Vitis vinifera* | *Vitaceae* | *Vitales* | ADULT | MT commercial sample |

[a] This sample was analysed twice at distance of time to check repeatability

collected at two different phenological stages: embryonic stage (meristematic tissues: quiescent bud, bud, young sprout) and adult stage (adult leaves). The collection was performed in two years (2013/2014), from February to April, from their natural habitat in Val Chisone, Val Pellice and Val Germanasca, Turin Province (Italy). The manufacturing of the corresponding extracts (Table 1) was performed both in an Italian company (GEAL PHARMA- Turin, Italy) and by the authors in the analytical laboratory of the University of Turin (DISAFA) during 2013 and 2014.

Both embryonic and adult parts were used fresh to prepare herbal preparations. MT and GM were prepared according to the European Pharmacopoeia 8th edition (Pharmaciens 1965), following the procedure deriving from the French Pharmacopoeia.

## Chemicals

All chemicals were purchased from Sigma-Aldrich (Steinheim, Germany) and from VVR Chemicals. High purity water produced (HPW) with Millipore Milli-Q system was used throughout.

## Samples

A first group of samples (Table 1: samples numbered from 1 to 23), containing the same plant materials at different phenological stages, was prepared using the protocol of Mother Tinctures (MT) in ethanol 96% v/v with a 1:10 ratio between plant and solvent. Briefly, for each fresh plant species, about 1 kg of stuff was treated, and the amount of solvent was added so as to obtain a weight ratio of 1:10 between plant and solvent (water/ethanol). Bioactive compounds were extracted through a cold maceration process for 21 days, in 60/40 (by weight) water/ethanol (96%), followed by a first filtration (Whatman filter paper, hardened ashless circles, 185 mm diameter), a manual pressing and, after 2 days of decanting, a second filtration (Whatman filter paper, hardened ashless circles, 185 mm diameter). The samples were stored in dark bottles at normal atmosphere (N.A.), at 4 °C and 95% relative humidity until analysis.

Later, in order to verify the proposed strategy also on samples prepared using the protocol of Glyceric Macerates (GMs), the study was focalized on the following three vegetable species: *Ribes nigrum*, *Tilia tomentosa*, *Vitis vinifera* spp. Concerning them, the study was enlarged to a second group of samples (Table 1: samples numbered from 24 to 35): a part prepared by the authors using the protocol of Glyceric Macerates (GMs) and another part taken by trade (both GMs and MTs). The GMs protocol is analogous to the one described for MTs, but the extraction solvent is a mixture of water/glycerol/ethanol (50/30/20 by weight) and it was adjusted so as to obtain a weight ratio of 1:20 between plant and solvent.

## UV–VIS and Fluorescence spectroscopy

Absorption spectra in the ultraviolet and visible regions were recorded in the range 190–1100 nm using an Agilent 8453 spectrophotometer with 1 nm resolution. The cells were rectangular quartz cuvettes with 0.1 cm path length. Before being analysed, the samples were centrifuged at 5000 rpm for 15 min and diluted at a ratio of 1:10 with three different solvents: water, water: ethanol 60:40, and ethanol, respectively (codified in the final letter of the sample name as "A", "M" and "E" respectively). This was performed in order to check the spectral variations due to the change of solvent. For each sample the spectrum was collected at room temperature in duplicate and the results were averaged. The replicates were performed at two months away.

Conventional Fluorescence Spectroscopy, using a Shimadzu RF-5301PC spectrofluorophotometer, was performed exciting a sub-set of samples at a fixed wavelength λex (430 nm), and collecting the corresponding emission spectra (in the range 450–800 nm) (Sadeckà and Tòthovà 2007). Measuring emission 20 nm beyond the excitation wavelength prevented Rayleigh scatter. Xenon lamp source was used for excitation, the acquisition interval and integration time were set to 1 nm and 0.1 s, respectively. The widths of both excitation and emission slit were set to 5 nm. The same dilution of the samples at the ratio of 1:10 with solvent was used to have the same samples preparation as before. Thus the traditional right angle fluorescence spectroscopic technique was applied. The total absorbance of samples has been always lower than 0.1 a.u. avoiding spectral distortions. Each sample was analyzed in triplicates, and the results were reported as mean values.

## Multivariate analysis

The multivariate repeatability of the spectra was preliminarily checked by replicating twice at a distance of time the UV–VIS analysis of at least one sample for each vegetable extracts (Table 1: samples marked with an letter "a") and analyzing the score plot obtained by Principal Component Analysis (PCA) (Massart et al. 1998) on the autoscaled data as raw data and after 1st derivatization. In detail, the replicates were structured in a data matrix, namely matrix $R_{22,571}$, with 22 rows (the 11 replicate samples) and 571 columns (variables: the absorbances at different wavelengths in the range 230–800 nm since the two intervals 190–230 nm and 800–1100 nm were preliminary removed since the signals were saturated or without interesting absorptions respectively).

Then the UV–VIS spectra were acquired for the first twenty-three MTs of Table 1 (samples numbered from 1 to 23) each diluted in the three different solvents (water, water: ethanol 60:40, and ethanol) for a total of 69 samples plus the 11 replicates. They were organized into a datamatrix named $A_{80,571}$: consisting in how many rows as the number of samples (80) and how many columns as the recorded absorbance (the 571 absorbances at different wavelengths in the range 230–800 nm).

Then PCA was performed as unsupervised patter recognition technique both on the original spectral data matrix ($A_{80,571}$) and on the corresponding 1st derivatives respectively ($D_{80,571}$). In every case the data were previously divided into two sets: training set and external test set. The objects of the external set are about the 10% of the total ones, excluding replicates; they were selected by Kennard–Stone algorithm (Kennard and Stone 1969). Thus the PCA was performed on the training set and the samples of the test set were projected onto the PCA hyperplane calculated for the training data.

Then with the aim of both classifying samples on the basis of their phenological stage (i.e. "embryonic phenological stage" class or "adult phenological stage" class) and defining statistical models capable of determining the authenticity of unknown samples, two class-modelling techniques were performed: UNEQ and SIMCA respectively.

For each class, UNEQ defines the mathematical model and the category space around it as the confidence hyperellipsoid of the category, according to the Mahalanobis distance from the centroid. Confidence intervals can be built at different levels of significance; in this study a 95% level of confidence was considered. Since UNEQ can be used only with well-defined data matrixes, i.e., those having a high ratio between number of objects and number of variables inside each category, the scores on the first eight PCs were used as descriptive variables (Oliveri et al. 2011).

Then Soft independent modeling of class analogy (SIMCA) was applied to the original data sets. This method builds class models based on PCA. The models are defined by the range of the scores on a selected number of the first components (in this study the number of components was set to 5). Therefore the models are hyper-parallelepipeds referred to as the multidimensional boxes of the SIMCA inner space. The other dimensions are the outer space, the space of noise. An object is accepted by a model when its distance from the model is less than a critical distance, defined by means of Fisher statistics. Nonerror-rate (the percentage of correctly classified objects), sensitivity (the nonerror-rate for a class), and specificity (the percentage of objects of other classes rejected by the class model under study) of the obtained models are the criteria used to measure the classification and the modeling performances. An internal cross-validation procedure with five cancellation groups (5CV) and an external validation procedure, using an external test set were always applied. The external test set was again selected by Kennard–Stone algorithm. By this algorithm, the objects having the largest distance in the multivariate space were assigned to the test set in order to have a severe validation of the classification rules.

Later the attention was focused on the following three plant species: *Ribes nigrum*, *Tilia tomentosa*, *Vitis vinifera*

spp. expanding the study also to other samples (Table 1: second group of samples numbered 24–35) and acquiring both the UV–VIS spectra and the Fluorescence ones. In detail, the 23 extracts of these three species (Table 1: samples 8–12 and samples 18–35), after ethanol dilution (ratio 1:10), were described both by UV–VIS fingerprints (range 230–800 nm: 571 absorbances) and by fluorescence fingerprints (i.e. the 351 corresponding emissions in the range 450–800 nm). The UV–VIS spectral data and the Fluorescence spectral data were structured in the following two data matrices: $T_{23,571}$ and $F_{23,351}$ respectively and PCA was performed as already described.

Multivariate statistical analysis was performed using the R-based chemometric software of the Italian chemometrics society and the V-PARVUS 2008 software (Forina et al. 2008).

## Results and discussion

Figure 1 shows the UV–VIS spectral profiles recorded for the following vegetable species *Acer pseudoplatanus*, *Aesculus hippocastanum*, *Fagus sylvatica*, *Ribes nigrum*, *Rosa canina*, *Salix caprea*, *Tilia tomentosa*, *Vitis vinifera* spp. diluted with three different solvents: water, water: ethanol 60:40, and ethanol, respectively (codified in the final letter of the sample name as "A", "M" and "E" respectively) at different phenological stages (blue: adult stage; green: embryonic stage). For each plant different spectral profiles according to the different stages are highlighted in blue for the adult stage and in green for the embryonic stage respectively.

Looking at Fig. 1, it can be clearly seen that the spectra of the different vegetable species are different. This highlights as the pattern of absorbances in the UV–VIS region is strictly connected with the botanical origin of the plants. On the contrary, for each botanical specie the spectral variations due to the change of solvent, in which the samples were diluted, are practically influent respect to the differences due to the phenological stage of each plant, as shown in Fig. 1. Concerning embryonic stage both the quiescent buds and the young sprouts, when present, were taken into account respectively, besides the buds. Concerning *Vitis vinifera* (Fig. 1) is important to underline that two different adult parts of the plant have been taken into account: leaves (blue line) and stalks (blue dotted line) respectively. It is evident that the two different extracts of adult parts give origin to different spectral profiles and that both profiles are different from those of the buds (leaf buds).

In other words for each vegetable species the same plant part at different phenological stages have different spectral patterns, especially after removing two intervals: 190–230
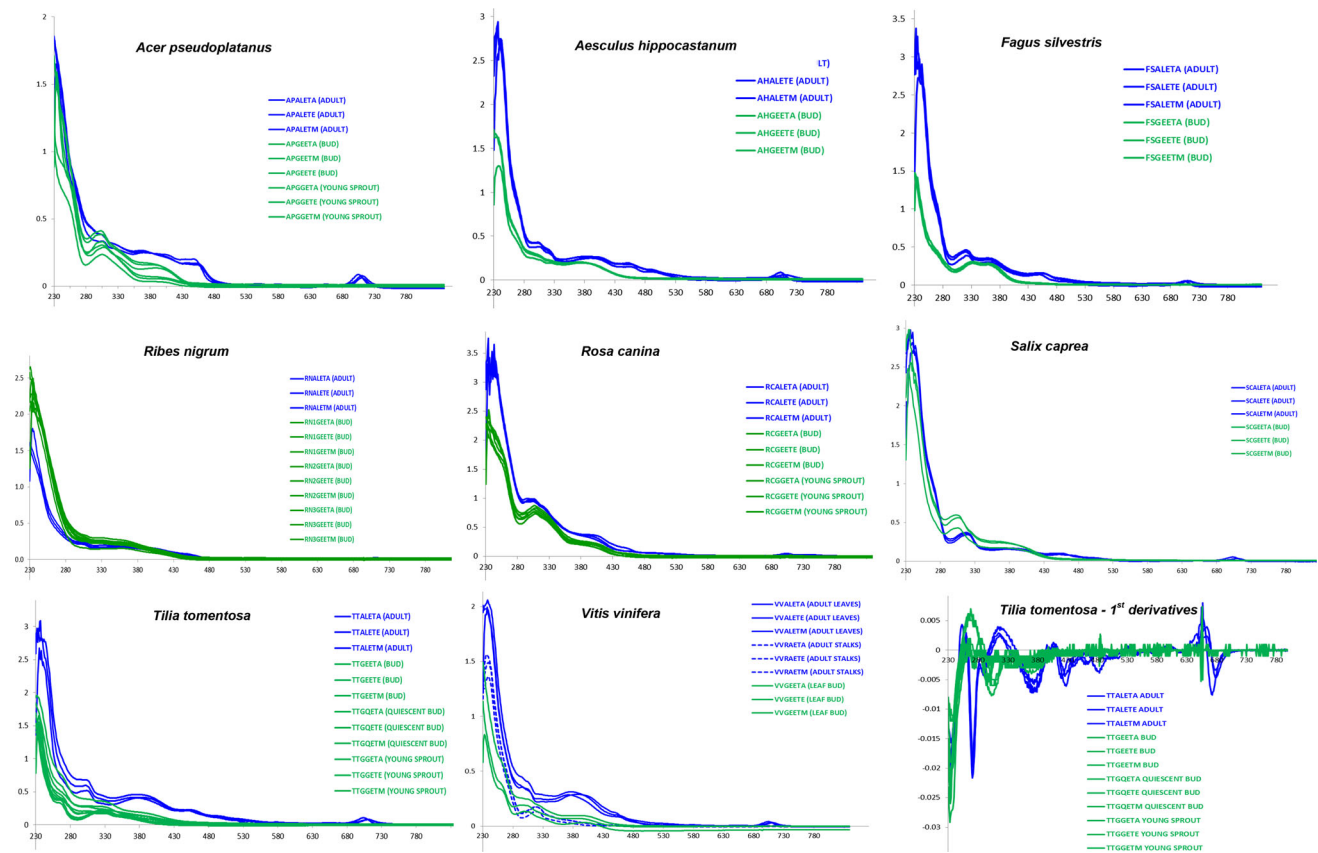
**Fig. 1** The UV–VIS spectral profiles of *Acer pseudoplatanus*, *Aesculus hippocastanum*, *Fagus sylvatica*, *Ribes nigrum*, *Rosa canina*, *Salix caprea*, *Tilia tomentosa*, *Vitis vinifera* spp. diluted with three different solvents: water, water:ethanol 60:40, and ethanol, respectively (codified in the final letter of the sample name as "A", "M" and "E" 129 respectively) at different phenological stages (*blue* adult stage; *green* embryonic stage). Concerning embryonic stage
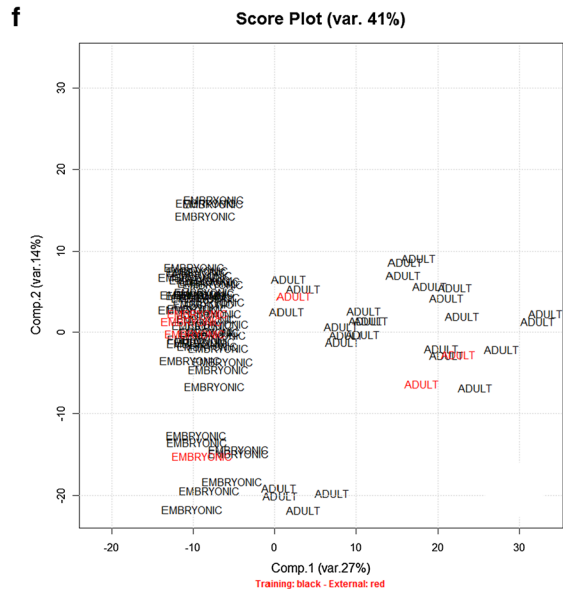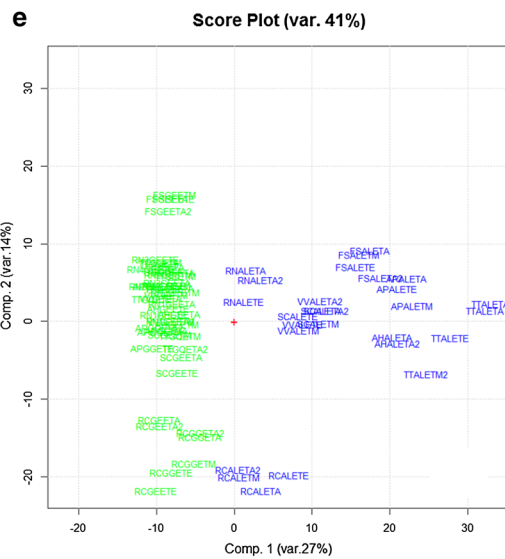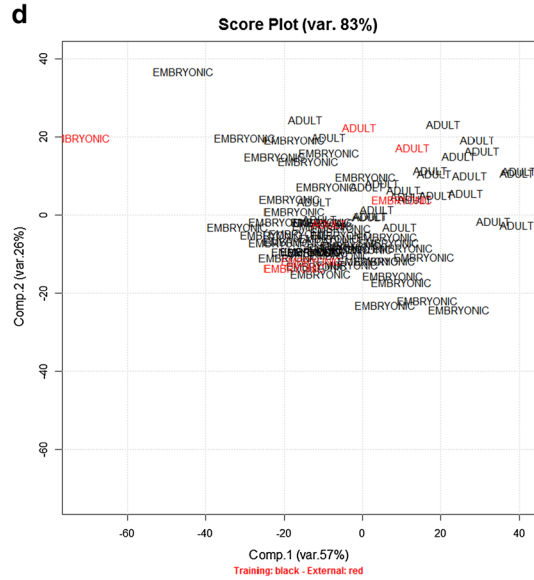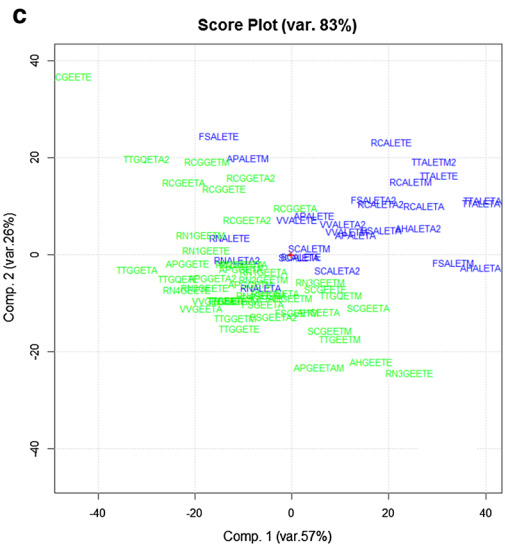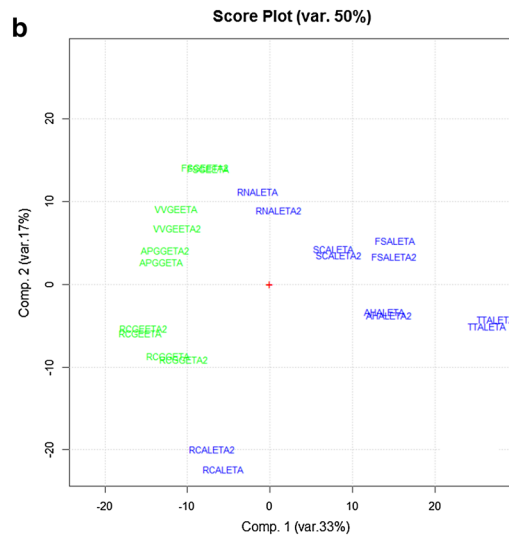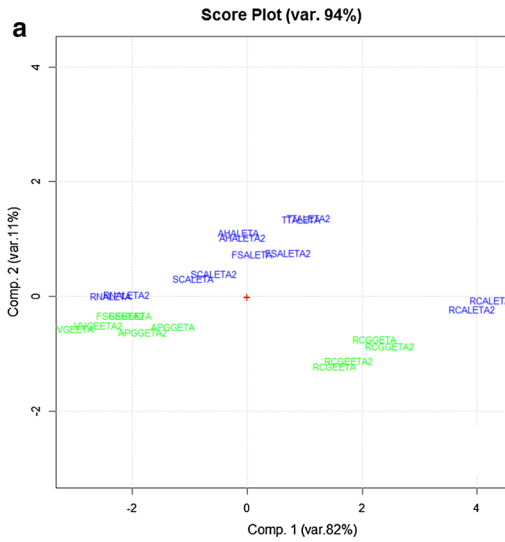
both the quiescent buds and the young sprouts, when present, were taken into account besides the buds. Concerning adult stage only for *Vitis vinifera* spp. two different adult parts of the plant have been taken into account: leaves (*blue line*) and stalks (*blue dotted line*) respectively. In the *lower right corner* there is the corresponding first-order derivative spectrum for the *Tilia tomentosa* spp (color figure online)

and 800–1100 nm in which the signal was saturated or without interesting absorptions respectively. In particular it is evident as the region around 670 nm is typical of the adult stage and always lacks in the meristematic stages. This is more evident analysing the corresponding spectra in the 1st derivatives as highlighted for *Tilia tomentosa*, taken as an example, in the lower right corner of Fig. 1. This small region seems to be peculiar of the leaf adult stage and probably corresponds to absorptions of chlorophylls, which are recognized to be almost absent or present in low amount in the leaf buds (Mishra and Mahananda 2013).

However it was not possible to exclude a priori the presence of informative features along the whole spectrum of each sample, since the differences among the samples may cause very slight spectral modifications. Thus the analytical information contained in each spectrum was taken into account as a multivariate fingerprint, then PCA and class-modelling techniques were performed to explore and to classify the data set.

**Fig. 2** The PC1–PC2 scores plot of the UV–VIS absorbances data ▶ matrix $R_{22,571}$ (**a**) and after 1st order derivatization (**b**) (embryonic stage: *green samples*; adult stage: *blue samples*). The PC1–PC2 scores plot of the UV–VIS absorbances data matrix $A_{80,571}$ (**c**) and after 1st order derivatization $D_{80,571}$ (**e**) (embryonic stage: *green samples*; adult stage: *blue samples*) calculated using the 73 training set samples and the corresponding plot obtained projecting the external test set (7 samples highlighted in red and coded according Table 1) respectively (**d, f**) (color figure online)

The multivariate repeatability for the spectral fingerprints was preliminarily checked by PCA on the autoscaled data matrix $R_{22,571}$ with 22 rows (the 11 replicate samples) and 571 columns (variables: the absorbances at different wavelengths in the range 230–800. The results are summarized in Fig. 2a, b for the $R_{22,571}$ data matrix before and after derivatization (1st derivatives) respectively. As Fig. 2a, b shows, the multivariate repeatability is really good since the scores of each replicate are quite close between them on the first two PCs, accounting for 94 and

|  | UNEQ | SIMCA |
|---|---|---|
| **Classification rate (%)** | 100 | 98.6 |
| **Prediction rate (5CV) (%)** | 98.6 | 97.2 |
| **Sensitivity (%)** | 93.1 | 73.9 |
| **Specificity (%)** | 53.4 | 100 |
| **Efficiency (%)** | 70.5 | 86.0 |
| **External prediction rate (7 objects in the test set) (%)** | 100 | 100 |



**Fig. 3** UNEQ and SIMCA results of the 1st order derivatives data matrix $D_{80,57}$ with the corresponding Coomans plot for the embryonic class (*green*) and the adult class (*blue*). The external test set objects are shown in *red* (color figure online)

50% of the total variance before and after the 1st derivatization of the data respectively.

Concerning the original data set $A_{80,571}$, it was divided in two subsets by Kennard–Stone algorithm to build the chemometric rules (*training set: 73 objects*) and the other one was used to validate them (*external test set: 7 objects, representing about the 10% of the objects excluding replicates*). In details, the 7 samples of the test set were projected onto the PCA hyperplane calculated for the training data. The sample scores (of the 73 samples) on PCs 1 and 2—that explain 83% of total variance—are shown in the corresponding plot as far as training set is concerned (Fig. 2c). Figure 2d shows the results obtained projecting the test set objects (7 samples) onto the PC1–PC2 plot previously calculated using the training set. A satisfactory separation among the embryonic extracts (buds and young sprouts) and the adult ones is achieved on the first two PCs.

As far as the 1st derivatives matrix $D_{80,571}$ is concerned, the PCA results (of the training and test set) are quite improved and they are displayed in Fig. 2e, f respectively without and with the test set by the sample scores on PCs 1 and 2—that explain 41% of total variance. There is a very good separation between adult and meristematic tissues



**Fig. 4** The PC1–PC2 scores plot of the UV–VIS absorbances data matrix $T_{23,571}$ (three botanical species: *Ribes nigrum*, *Tilia tomentosa*, *Vitis vinifera*) after 1st order derivatization (embryonic stage: *green samples*; adult stage: *blue samples*) (color figure online)

extracts for each species. In particular PC1 is mainly responsible for separation between adult and young phenological stages; the spectral variables having greater importance (loading value) on this PC are represented by the regions around 600–700 nm and around 450 nm. On the other hand, PC2 seems mainly responsible for the separation among the different vegetable species and the spectral variables more involved are located at lower wavelenghts (UV region).

Then two class-modelling techniques were performed (UNEQ and SIMCA) with the aim of both classifying samples on the basis of their phenological stage (i.e. "embryonic phenological stage" class or "adult phenological stage" class) and defining statistical models. The results are summarized in Fig. 3, where validation parameters and Coomans plots are respectively reported. In details, internal and external prediction rate, sensitivity (the percentage of objects of the studied category accepted by the model) and specificity (the percentage of objects of other category rejected by the model of the class being modelled) were computed to evaluate the models performance. Looking at the class-modelling results both UNEQ

and SIMCA give satisfactory results: all test set objects are correctly predicted (external prediction rate); the SIMCA model has a better efficiency than the UNEQ one. Looking at the class modelling results it can be seen that sometimes the external prediction results are slightly more optimistic respect to that one provided by both the calibration and the internal evaluation sets, as already reported in literature (Casale et al. 2012). Since it depends how the Kennard–Stone algorithm chooses the external test set (see red objects of Fig. 2d, f). Sometimes the objects belonging to the three sets (calibration, internal and external evaluation sets) are not so distant one from each other. Thus, randomly, the algorithm can provide a slightly more optimistic prediction respect to that one provided by the other sets.

A comparison of the results of these two modeling approaches reveal that UNEQ has a greater sensitivity, this means fewer false positive results, while SIMCA has greater specificity, this means fewer false negative results.

In order to verify if the proposed procedure could be useful for the detection of adult/embryonic extracts also in samples obtained using the protocol of Glyceric Macerates (GMs), the study was focused on the following three
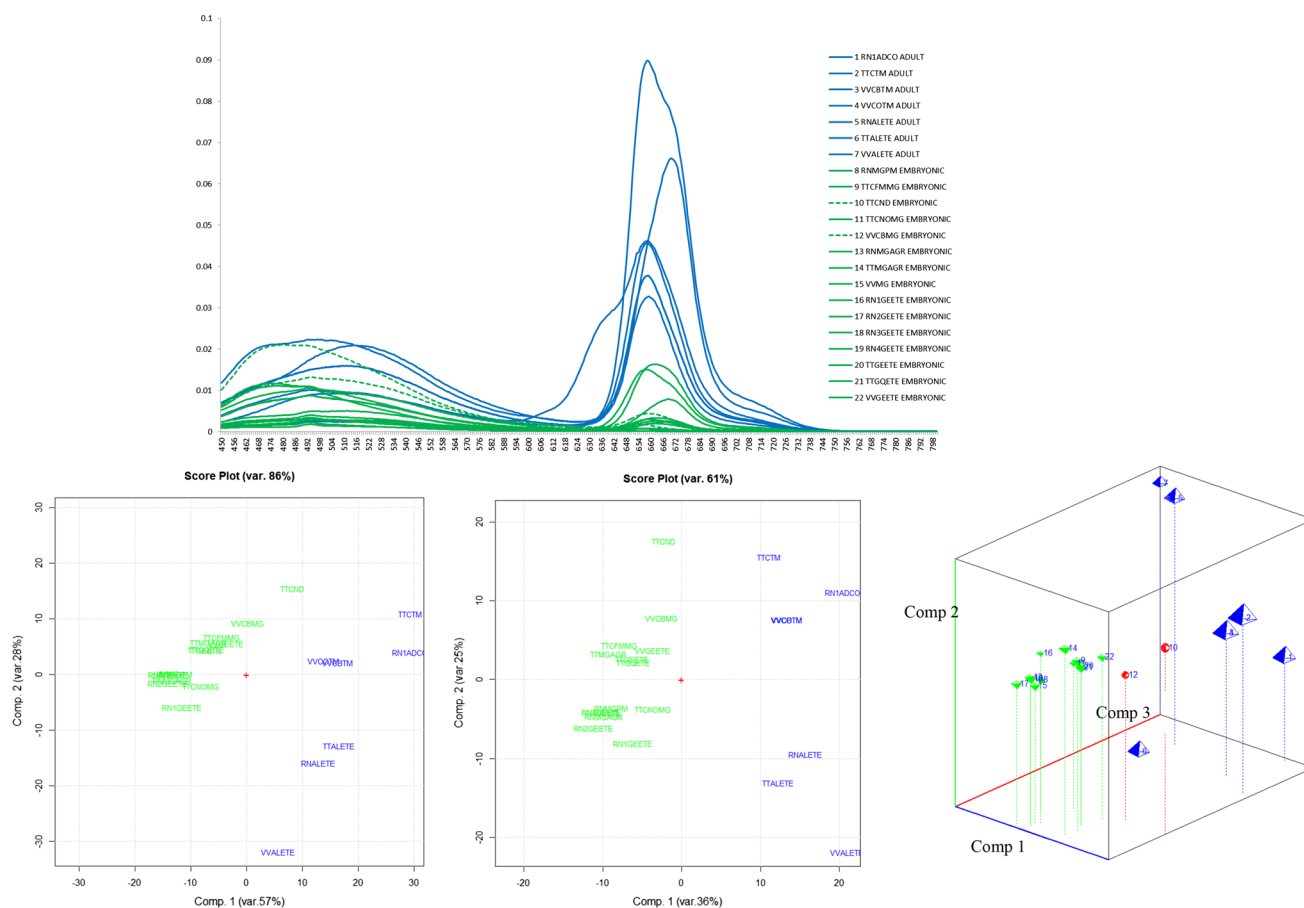


**Fig. 5** The Fluorescence spectral emissions (from 450 to 800 nm) of MTs and GMs of the following three botanical species: *Ribes nigrum*, *Tilia tomentosa*, *Vitis vinifera*; the PC1–PC2 and the PC1–PC2–PC3

score plots of the corresponding fluorescence emission data matrix $F_{23,351}$(embryonic stage: *green samples*; adult stage: *blue samples*) (color figure online)

species: *Ribes nigrum*, *Tilia tomentosa* and *Vitis vinifera* enlarging the study to a second group of samples. Besides the samples already mentioned (Table 1: samples 8–12 and 18–23), the homemade GMs and the commercial samples (GMs and MTs) numbered from 24 to 35 in Table 1 were also taken into account. All the 23 samples were diluted just in ethanol, since, as already mentioned, not significant variations were detected using the other dilution solvents, such as water and water:ethanol 60:40.

The 23 UV–VIS spectra were recorded as above and structured in the $T_{23,571}$ data matrix. Figure 4 reports the scores plot (PC1/PC2) obtained by the PCA on the auto-scaled 1st derivatives of this data set. Analyzing these results it is evident that the behavior of these samples is consistent with the previous results, independently from the extraction protocol. In fact PC1 is able to discriminate between the two phenological stages apart from two embryonic commercial samples (namely TTCND and VVBMG). More in details, the commercial sample of *Tilia tomentosa* GM named TTCND has scores on the first two PCs in the middle among the adult and the embryonic extracts of *Tilia tomentosa*, while the commercial sample of *Vitis vinifera* GM named VVBMG has scores quite close to those of the corresponding samples in adult stage.

In order to check the UV–VIS results, fluorescence fingerprints for these same 23 samples were collected and structured in the $F_{23,351}$ data matrix (i.e. the corresponding emission spectra in the range 450–800 nm obtained as reported above).

Figure 5a–c reports: the fluorescence spectra, the corresponding PC1/PC2 scores plots obtained by the PCA on the autoscaled data before and after derivatization respectively. The fluorescence results confirm the previous UV–VIS ones: PC1 primarily discriminates between the two different phenological stages.

Nevertheless, as far as the already mentioned commercial embryonic samples named TTCND and VVBMG are concerned, their scores are again slightly moved towards the adult samples. This is evident on the third principal component (PC3) whose projections are highlighted in Fig. 5d together with the PC1/PC2 ones. Thus for these two commercial samples further analysis should be necessary in order to evade the suspect of a non-declared mixing of adult and embryonic extracts.

## Conclusion

A quality control of plant-derived preparations is urgently needed in order to avoid mislabelling and frauds. Methods based on the characterization of the phytocomplex of herbal products, despite their high level of accuracy and specificity, are often expensive and time consuming for industry.

In this contest the present study provides an analytical shortcut based on UV–VIS–Fluorescence spectroscopy coupled to unsupervised (i.e. PCA) and supervised (i.e. class-modelling) pattern recognition techniques for a rapid identification of plants extracts coming from meristematic tissues, whose limited productions could lead to adulterations with adult parts derivatives or even with undeclared vegetable species. The overall results demonstrate that the UV–VIS–Fluorescence absorptions of an extract, coming from a given plant tissue of a given species, are useful in the identification of both the plant species and its tissues. These absorptions are strictly correlated to the whole phytocomplex, which is genetically determined and whose molecules are able to absorb in this spectral range.

Thus the non-specific analytical procedure reported in this study may allow to detect suspicious samples, that deserve further investigations, avoiding to extensively analyse a large number of samples with the promising prospect for analysts to directly address the difficult problems in quality control of complex herbal products, in particular for composition assessment and fraud detection.

## References

Amaral JS, Seabra RM, Andrade PB, Valentao P, Pereira JA, Ferreres F (2004) Phenolic profile in the quality control of walnut (*Juglans regia* L) leaves. Food Chem 88:373–379. doi:10.1016/j.foodchem.2004.01.055

Boggia R, Casolino C, Hysenaj V, Oliveri P, Zunin P (2013) A screening method based on UV–Visible spectroscopy and multivariate analysis to assess addition of filler juices and water to pomegranate juices. Food Chem 140:735–741. doi:10.1016/j.foodchem.2012.11.020

Bosco R, Caser M, Vanara F, Scariot V (2013) Development of a rapid LC-DAD/FLD method for the simultaneous determination of auxins and abscisic acid in plant extracts. J Agric Food Chem 61:10940–10947. doi:10.1021/jf4034305

Calixto JB (2000) Efficacy, safety, quality control, marketing and regulatory guidelines for herbal medicines (phytotherapeutic agents). Braz J Med Biol Res 33:179–189. doi:10.1590/S0100-879X2000000200004

Casale M, Oliveri P, Casolino C, Sinelli N, Zunin P, Armanino C, Forina M (2012) Characterisation of PDO olive oil Chianti Classico by non-selective (UV–visible, NIR and MIR spectroscopy) and selective (fatty acid composition) analytical techniques. Anal Chim Acta 712:56–63

Chandra A, Rana J, Li Y (2001) Separation, identification, quantification, and method validation of anthocyanins in botanical supplement raw materials by HPLC and HPLC − MS. J Agric Food Chem 49:3515–3521. doi:10.1021/jf010389p

Cordell GA (2011) Phytochemistry and traditional medicine—a revolution in process. Phytochem Lett 4:391–398. doi:10.1016/j.phytol.2011.05.005

Dell'Agli M, Di Lorenzo C, Badea M, Sangiovanni E, Dima L, Bosisio E, Restani P (2013) Plant food supplements with anti-

inflammatory properties: a systematic review (I). Crit Rev Food Sci Nutr 53:403–413. doi:10.1080/10408398.2012.691916

Derde MP, Massart DL (1986) UNEQ: a disjoint modelling technique for pattern recognition based on normal distribution. Anal Chim Acta 184:33–51. doi:10.1016/S0003-2670(00)86468-5

Donno D, Beccaro GL, Mellano GM, Cerutti AK, Canterino S, Bounous G (2012) Effect of agronomic and environmental conditions on chemical composition of tree-species buds used for herbal preparations. Int J Plant Res (VEGETOS) 25:21–29

Donno D, Beccaro GL, Mellano MG, Cerutti AK, Marconi V, Bounous G (2013) Botanicals in *Ribes nigrum* bud-preparations: an analytical fingerprinting to evaluate the bioactive contribution to total phytocomplex. Pharm Biol 51:1282–1292. doi:10.3109/13880209.2013.786101

Donno D, Beccaro GL, Masnari F, Mellano GM, Cerutti AK, Bounous G (2014a) Castanea spp as phytochemical source for herbal preparations. II Eur Congr Chestnut 1043:75–82

Donno D, Beccaro GL, Mellano MG, Bonvegna L, Bounous G (2014b) Castanea spp buds as a phytochemical source for herbal preparations: botanical fingerprint for nutraceutical identification and functional food standardization. J Sci Food Agric 94:2863–2873. doi:10.1002/jsfa6627

Donno D, Boggia R, Zunin P, Cerutti AK, Guido M, Mellano MG, Prgomet Z, Beccaro GL (2016) Phytochemical fingerprint and chemometrics for natural food preparation pattern recognition: an innovative technique in food supplement quality control. J Food Sci Technol 53(2):1071–1083. doi:10.1007/s13197-015-2115-6

Dvaranauskaitė A, Venskutonis PR, Raynaud C, Talou T, Viškelis P, Sasnauskas A (2009) Variations in the essential oil composition in buds of six blackcurrant (*Ribes nigrum* L) cultivars at various development phases. Food Chem 114:671–679. doi:10.1016/j.foodchem.2008.10.005

Espín JC, García-Conesa MT, Tomás-Barberán FA (2007) Nutraceuticals: facts and fiction. Phytochemistry 68:2986–3008. doi:10.1016/j.phytochem.2007.09.014

Esteban-Díez I, González-Sáiz JM, Sáenz-González C, Pizarro C (2007) Coffee varietal differentiation based on near infrared spectroscopy. Talanta 71:221–229. doi:10.1016/j.talanta.2006.03.052

Forina M, Lanteri S, Armanino C, Casolino C, Casale M, Oliveri P (2008) V-PARVUS 2008. Dip Chimica e Tecnologie Farmaceutiche e Alimentari University of Genova Available (free, with manual and examples) from authors or at <http://www.parvusunigeit>

Fowler MW (2006) Plants, medicines and man. J Sci Food Agric 86:1797–1804. doi:10.1002/jsfa2598

Gad HA, El-Ahmady SH, Abou-Shoer MI, Al-Azizi MM (2013) Application of chemometrics in authentication of herbal medicines. Rev Phytochem Anal 24:1–24. doi:10.1002/pca.2378

Gulati OP, Berry Ottaway P (2006) Legislation relating to nutraceuticals in the European Union with a particular focus on botanical-sourced products. Toxicology 221:75–87. doi:10.1016/j.tox.2006.01.014

Gurib-Fakim A (2006) Medicinal plants: traditions of yesterday and drugs of tomorrow. Mol Asp Med 27:1–93. doi:10.1016/j.mam.2005.07.008

Kennard RW, Stone LA (1969) Computer aided design of experiments. Technometrics 11:137–148. doi:10.1080/00401706.1969.10490666

Kesting JR, Huang J, Sorensen D (2010) Identification of adulterants in a Chinese herbal medicine by LC-HRMS and LC-MS-SPE/NMR and comparative in vivo study with standards in a hypertensive rat model. J Pharm Biomed Anal 51:705–711. doi:10.1016/j.jpba.2009.09.043

Knoss W, Chinou I (2012) Regulation of medicinal plants for public health—European community monographs on herbal substances. Planta Med 78:1311–1316. doi:10.1055/s-0031-1298578

Konik EA, Jungling RC, Bauer BA (2011) Herbs and dietary supplements in the European Union: a review of the regulations with special focus on Germany and Poland. J Diet Suppl 8:43–57. doi:10.3109/19390211.2010.547243

Lian YZ, Xie P, Chan K (2004) Quality control of herbal medicines. J Chromatogr B 812:53–70. doi:10.1016/j.jchromb.2004.08.041

Marini F, Bucci R, Magrì AL, Magrì AD (2006) Authentication of Italian CDO wines by class-modeling techniques. Chemometr Intell Lab Syst 84:164–171. doi:10.1016/j.chemolab.2006.04.017

Marston A (2007) Role of advances in chromatographic techniques in phytochemistry. Phytochemistry 68:2786–2798. doi:10.1016/j.phytochem.2007.08.004

Massart DL, Vandeginste BGM, Buydens LMC, De Jong S, Lewi PL, Smeyers-Verbek J (1998) Handbook of chemometrics and qualimetrics: part B. Elsevier, Amsterdam

Mishra KN, Mahananda MR (2013) Chlorophyll content studies from inception of leaf buds to leaf fall stage of Teak (*Tectona grandis*) of Kapilash forest Dhenkanal, Odisha. J Glob Biosci 2:26–30

Nicoletti M (2012) Nutraceuticals and botanicals: overview and perspectives. Int J Food Sci Nutr 63:2–6. doi:10.3109/09637486.2011.628012

Oliveri P, Downey G (2012) Multivariate class modeling for the verification of food-authenticity claims. TrAC Trends Anal Chem 35:74–86. doi:10.1016/j.trac.2012.02.005

Oliveri P, Di Egidio V, Woodcock T, Downey G (2011) Application of class-modelling techniques to near infrared data for food authentication purposes. Food Chem 125:1450–1456. doi:10.1016/j.foodchem.2010.10.047

Ovalle-Magallanes B, Rivero-Cruz I, Mata R (2013) Quality control tests for the crude drug of *Conyza filaginoides*. Pharm Biol 52:117–123. doi:10.3109/138802092013816972

Pharmaciens OND (1965) Pharmacopée Française, Codex Medicamentarius Gallicus, Codex Français: Monographie, Préparations Homéopathiques VIII edn, Paris

Rates SMK (2001) Plants as source of drugs. Toxicon 39:603–613. doi:10.1016/S0041-0101(00)00154-9

Reid LM, O'Donnell CP, Downey G (2006) Recent technological advances for the determination of food authenticity. Trends Food Sci Technol 17:344–353. doi:10.1016/j.tifs.2006.01.006

Sadeckà J, Tòthovà J (2007) Fluorescence spectroscopy and chemometrics in the food classification—a review. Czech J Food Sci 25:159–173

Scalbert A, Johnson IT, Saltmarsh M (2005) Polyphenols: antioxidants and beyond. Am J Clin Nutr 81:215S–217S

Vegvari G, Brunori A, Sandor G, Jocsak I, Rabnecz G (2008) The influence of growing place on the ruitn content on *Fagopyrum esculentum* and *Fagopyrum tataricum* varieties seeds. Cereal Res Commun 36:599–602

Wold S, Sjöström M (1977) Chemometrics: theory and applications. In: Kowalski BR (ed) Chemometrics: theory and application, ACS symposium series 52. American Chemical Society, Washington, pp 243–282

Zhu JQ, Fan XH, Cheng YY, Agarwal R, Moore CMV, Chen ST, Tong WD (2014) Chemometric analysis for identification of botanical raw materials for pharmaceutical use: a case study using *Panax notoginseng*. PLoS ONE 9(1):e87462. doi:10.1371/journal.pone.0087462