Talanta 119 (2014) 524–532

Contents lists available at ScienceDirect

Talanta

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

Fuzzy clustering evaluation of the discrimination power of UV–Vis and (±) ESI-MS detection system in individual or coupled RPLC for characterization of *Ginkgo Biloba* standardized extracts

Andrei Medvedovici [a,b], Florin Albu [a,b], Rodica Domnica Naşcu-Briciu [c], Costel Sârbu [c,*]

[a] University of Bucharest, Faculty of Chemistry, Department of Analytical Chemistry, Panduri Ave., No. 90, Bucharest 050663, Romania
[b] Bioanalytical Laboratory, SC Labormed Pharma SA, No. 44B, Th. Pallady Blvd., Bucharest 032266, Romania
[c] Babeş-Bolyai University, Faculty of Chemistry and Chemical Engineering, Department of Chemistry, Arany Janos Street, No. 11, Cluj-Napoca 400028, Romania

ARTICLE INFO

Article history:
Received 5 August 2013
Received in revised form
6 November 2013
Accepted 9 November 2013
Available online 27 November 2013

Keywords:
Ginkgo Biloba standardized extracts
Fuzzy hierarchical clustering
Principal components analysis
Cluster analysis
HPLC/UV/(±)ESI-MS

ABSTRACT

Aim: Discrimination power evaluation of UV–Vis and (±) electrospray ionization/mass spectrometric techniques, (ESI-MS) individually considered or coupled as detectors to reversed phase liquid chromatography (RPLC) in the characterization of *Ginkgo Biloba* standardized extracts, is used in herbal medicines and/or dietary supplements with the help of Fuzzy hierarchical clustering (FHC).
Experimental: Seventeen batches of *Ginkgo Biloba* commercially available standardized extracts from seven manufacturers were measured during experiments. All extracts were within the criteria of the official monograph dedicated to dried refined and quantified *Ginkgo* extracts, in the European Pharmacopoeia. UV–Vis and (±) ESI-MS spectra of the bulk standardized extracts in methanol were acquired. Additionally, an RPLC separation based on a simple gradient elution profile was applied to the standardized extracts. Detection was made through monitoring UV absorption at 220 nm wavelength or the total ion current (TIC) produced through (±) ESI-MS analysis. FHC was applied to raw, centered and scaled data sets, for evaluating the discrimination power of the method with respect to the origins of the extracts and to the batch to batch variability.
Results: The discrimination power increases with the increase of the intrinsic selectivity of the spectral technique being used: UV–Vis < MS(−) < MS(+), but it is strongly sensitive to chemometric transformation of data. Comparison with cluster analysis (CA) and principal components analysis (PCA) indicates that the FHC algorithm produces better classification. However, PCA and CA may be successfully applied to discriminate between the manufacturing sources of the standardized extracts, and at some extent, to monitor the inter-batch variability. Although the chromatographic dimension sensibly contributes to the discrimination power, spectral MS data may be used as the lone powerful holistic alternative in characterization of standardized *Ginkgo Biloba* extracts.

1. Introduction

Nowadays fingerprinting and pattern recognition algorithms represent valuable tools for the characterization of the complex chemical mixtures of natural origins, with promising results in various application fields such as food [1,2], beverages [3,4], agriculture [5,6], chemotaxonomy [7,8], herbal medicines [9], dietary supplements [10,11], metabolic profiling [12–14], environmental [15,16] and standardization [17]. The European Medicines Agency [18] recommends that the appropriate fingerprinting procedures should be based on chromatographic techniques. However, other techniques, such as the spectral ones, may lead to interesting and useful results [19] as the US Food and Drug Administration [20] recommends.

Evolvement of fingerprinting procedures during the last decade has been supported by the application of powerful and advanced chemometric methods, among which principal component analysis (PCA), partial least squares (PLS), cluster analysis (CA), linear discriminant analysis (LDA) and artificial neuronal networks (ANN)

Abbreviations: ANN, artificial neuronal network; a.m.u., atomic mass unit; CA, cluster analysis; DAD, diode array detection; (±)ESI, positive/negative electro spray ion source; FHC, Fuzzy hierarchical clustering; HPLC, high pressure liquid chromatography; IBJ, Backer–Jain Index; LDA, linear discriminant analysis; mAU, milli-absorption unit; MS, mass spectrometry; PCA, principal components analysis; PCN, normalized partition coefficient; PEN, normalized partition entropy; PLC, partial least squares; RAM, random access memory; RPLC, reversed phase liquid chromatography; TIC, total ion current; TOC, total organic carbon; UV-Vis, ultraviolet and visible

* Corresponding author. Tel.: +40 264 593833; fax: +40 264 590818.
E-mail address: csarbu@chem.ubbcluj.ro (C. Sârbu).

0039-9140/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.11.035

should be mentioned as the most frequently used [21–25]. These methods generally lead to very efficient classifications, but are highly sensitive to outliers, missing data, and/or inadequate linear correlation between variables due to their poor distribution, as major sources for erroneous conclusions. These disadvantages may be eliminated by using robust techniques, such as Fuzzy hierarchical clustering (FHC). Fuzzy logic is a form of many-valued logic, which deals with reasoning that is approximate rather than fixed and exact. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true (1) and completely false (0) [26]. In classical CA each object must be assigned to exactly one cluster and this is leading to ambiguity and error in cases of outliers or overlapping clusters and affords information loss, while FHC is leading to classes and subclasses representing a collection of ill-defined and not-distinct objects with undefined boundaries in which the transition from membership to non-membership in a subclass of a reference set is gradual rather than abrupt.

The Fuzzy theory is basically a theory of graded concepts. It is an extreme generalization of ordinary set theory and is basically designed to handle the concept of partial truth or fuzziness. It provides an adequate conceptual framework as well as a mathematical tool to model the real world problems which are often vagueness and indistinct [27–29]. Most FHC algorithms are based on objective functions and determine an optimal classification by minimizing them. In objective function based clustering each cluster is usually represented by a cluster prototype, which consists in a cluster center (whose name already indicates its meaning) and maybe some additional information about the size and the shape of the cluster. The affiliation to a specific cluster is given by the degrees of membership, computed as the distance from the data point to the cluster center defined by the fuzzy means. The closer a data point lies to the center of a cluster, the higher is its degree of membership to this cluster. Hence, the problem to divide a data set into clusters can be stated as the task to minimize the distances of the data points to the cluster centers, since, of course, the principal target represents the maximization of membership degrees [30,31].

The leading concepts of the Fuzzy theory may successfully support not only concrete fingerprinting procedures of complex natural mixtures, but also a holistic characterization of complex mixtures, with reduced risks relating to the misinterpretation of the primary data.

Ginkgo Biloba is a medicinal plant frequently used as dried vegetal material or refined extracts in herbal medicines or dietary supplements. The composition of *Ginkgo Biloba* derived materials is complex [32], usually requiring powerful analytical tools for a comprehensive characterization [33–36].

The herein presented approach aims to deliver responses to the following main questions: (i) Do the spectrometric techniques (UV/Vis and $\pm$ MS) have, taken separately, enough potential to produce experimental data with the required discriminating power to assess the origins and batch to batch variability of complex natural samples? (ii) Is the chromatographic dimension essential for the correct assessment of the origins and reproducibility in the production stages of standardized natural extracts? (iii) Which of the chemometric methods (HFC, PCA and CA) is the most suitable one for informational discrimination assessment of the analytical experimental data? (iv) Is chemometric data treatment affecting the discrimination power?

Our starting decision was to focus on standardized extracts and not on dried vegetal materials, considering this approach as a "worst" case application (discrimination is tested on materials meeting strictly specified quantitative criteria).

The opportunity of our approach is based on the increasing success of dietary supplements from natural source, on the medical markets, which are characterized by affordable prices.

This market is, however, less strictly regulated than the one of classic medicines (drugs) and the variability of the raw active materials is naturally higher than for synthetic products. An evident contradiction appears between the need of efficient production with the lowest expenditure and the objective necessity of deep analytical characterization of the complex active raw materials of natural origins and their inherent quality control, involving expensive resources. Holistic approaches as those investigated in the present manuscript may represent straightforward and relatively inexpensive alternatives to the assessment of the composition variability with respect to origins and batch to batch variability of complex mixtures from natural sources. The information produced through application of chemometric methods to spectral or chromatographic data should not be considered as a substitute for the quantitative assays involved in the characterization of standardized materials according to official monographs, but only an easier and rapid way of controlling the source (manufacturer) and batch to batch reproducibility.

2. Materials and methods

2.1. Chemicals

Acetonitrile and methanol were HPLC gradient grade from Merck (Darmstadt, Germany). Formic acid (extra pure grade) from Merck was also used during experiments. Water for chromatography (resistivity of minimum 18.2 MΩ and residual total organic carbon content – TOC – of maximum 30 ng mL^{-1}) was produced within the laboratory by means of a TKA Lab HP 6UV/UF instrument (TKA Instruments as part of Thermo Fischer Scientific, Niederelbert, Germany).

2.2. Samples

Seventeen *Ginkgo Biloba* standardized extracts from six different manufacturers (A–F) were used during experiments. Samples 1–3 are batches produced at 1 year distance by manufacturer A. Samples 4–11 are eight different batches from manufacturer B, produced over 2 years interval (samples 7–11 are consecutive batches). Samples 12 and 13 are batches from manufacturer C, while samples 14 and 15 are consecutive batches from manufacturer D. Samples 16 and 17 are from suppliers E and F. All analyzed batches were placed within their declared shelf life period at the moment of the analysis. All standardized extracts are declared by manufacturers to comply with requirements of the official monograph of European Pharmacopoeia for *Ginkgo Biloba* dry extract, refined and quantified [37].

2.3. Equipment

Experiments were performed with an Agilent 1200 SL series LC/MSD (Agilent Technologies) system consisting of the following modules: degasser (G1379B), binary pump (G1312B), automated injector (G1367C and its corresponding thermostat G1330B), column thermostat (G1316B), diode array detector (G1315C) fitted with a semi-micro 5 μL flow cell (G1314-60011), ESI standard source (G1948B), and triple quadrupole mass spectrometric detector (G2571A). System control, data acquisition and interpretation were made with the Agilent Mass Hunter software version B 04.01 (B4114 Patch 1) incorporating both qualitative and quantitative packages.

2.4. MS parameters

The parameters controlling the ESI ion source were as following: drying gas (N_2); temperature (350 °C); drying gas flow (13 L min^{-1}); pressure of the nebulizing gas (60 psi); capillary voltage (4000 V).