Food Chemistry 152 (2014) 391–398

Contents lists available at ScienceDirect

Food Chemistry

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

Analytical Methods

An integrated approach utilising chemometrics and GC/MS for classification of chamomile flowers, essential oils and commercial products

Mei Wang [a], Bharathi Avula [a], Yan-Hong Wang [a], Jianping Zhao [a], Cristina Avonto [a], Jon F. Parcher [a], Vijayasankar Raman [a], Jerry A. Zweigenbaum [d], Philip L. Wylie [d], Ikhlas A. Khan [a,b,c,*]

[a] National Center for Natural Products Research, University of Mississippi, MS 38677, USA
[b] Department of Pharmacognosy, School of Pharmacy, University of Mississippi, MS 38677, USA
[c] Department of Pharmacognosy, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia
[d] Agilent Technologies, 2850 Centerville Rd., Wilmington, DE 19808-1610, USA

ARTICLE INFO

Article history:
Received 14 August 2013
Received in revised form 6 November 2013
Accepted 21 November 2013
Available online 4 December 2013

Keywords:
Matricaria chamomilla
Chamaemelum nobile
Chrysanthemum morifolium
Sample class prediction model
Chemometric analysis

ABSTRACT

As part of an ongoing research program on authentication, safety and biological evaluation of phytochemicals and dietary supplements, an in-depth chemical investigation of different types of chamomile was performed. A collection of chamomile samples including authenticated plants, commercial products and essential oils was analysed by GC/MS. Twenty-seven authenticated plant samples representing three types of chamomile, *viz*. German chamomile, Roman chamomile and Juhua were analysed. This set of data was employed to construct a sample class prediction (SCP) model based on stepwise reduction of data dimensionality followed by principle component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). The model was cross-validated with samples including authenticated plants and commercial products. The model demonstrated 100.0% accuracy for both recognition and prediction abilities. In addition, 35 commercial products and 11 essential oils purported to contain chamomile were subsequently predicted by the validated PLS-DA model. Furthermore, tentative identification of the marker compounds correlated with different types of chamomile was explored.

1. Introduction

The 'wonder' plant chamomile is one of the most widely used medicinal plants in the world. In the form of herbal teas, over one million cups of this natural product are consumed each day (Srivastava & Gupta, 2010). Chamomile is preferred for its pleasant taste and calming, sedative effects, as well as its long established medicinal properties. Frequently cited medicinal effects include the relief of sleeping disorders, diarrhoea, colic, wounds, mucositis and eczema (McKay & Blumberg, 2006; Petronilho, Maraschin, Coimbra, & Rocha, 2012). Additional beneficial properties, such as anti-inflammatory, anti-spasmodic, anti-allergic and anti-bacterial, have been attributed to chamomile (Buono-Core, Nunez, Lucero, Robinson, & Jullian, 2011). Commercial chamomile products include beverages, cosmetics, hair dyes, perfumes, massage oils, soaps and shampoos among others. Chamomile flowers are considered as an official drug in the pharmacopoeia of 26 countries.

Despite, or perhaps because of, its popularity and commercial significance, no exact characterisation of 'chamomile' is universally accepted. There are several types of chamomiles described in literature; however, all these herbs are grouped under the dicot plant family *Asteraceae*. The three most common types of chamomile observed in commercial products are German chamomile (*Matricaria chamomilla* L. syn: *M. recutita* L.), Roman chamomile (*Chamaemelum nobile* (L). All. syn: *Anthemis nobilis* L.) and Juhua (*Chrysanthemum morifolium* Ramat.) (Mabberley, 2008). The similarity of medicinal practice of different types of chamomile and lack of a clear definition of 'chamomile' lead to significant issues with respect to the quality control and authentication of commercial products purported to contain chamomile as an active ingredient. Moreover, the quality, safety and efficacy of herbal medicines commonly used throughout the world are difficult to determine or control with such a poorly defined natural product. Adulteration of commercial chamomile products is one of the most significant drawbacks in the promotion of herbal chamomile products (Omidbaigi, Sefidkon, & Kazemi, 2004).

The various types of chamomile and cultivars have been extensively studied, and the chemical compositions of many 'chamomile' plants have been determined (Antonelli & Fabbri, 1998; Farkas et al., 2003; Gao, He, Yue, Zou, & Zha, 2012; Guan, Wang, Shi, Bai, & Zhou, 2007; Mulinacci, Romani, Pinelli, Vincieri, & Prucher, 2000; Omidbaigi, Sefidkon, & Kazemi, 2003; Omidbaigi

* Corresponding author. Tel.: +1 662 915 7821; fax: +1 662 915 7062.
E-mail address: ikhan@olemiss.edu (I.A. Khan).

0308-8146/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.foodchem.2013.11.118

et al., 2004; Schilcher, Imming, & Goeters, 2005; Sparkman, 2005). The pharmacological activity of chamomile is mainly associated with the essential oil and flavonoid fractions (Barnes & Anderson, 2007; Galleano, Verstraeten, Oteiza, & Fraga, 2010; McKay & Blumberg, 2006; Mladenka, Zatloukalova, Filipsky, & Hrdina, 2010). The primary substances of the essential oil extracted from German chamomile flowers are α-bisabolol and its oxides, azulenes including chamazulene and acetylene derivatives (Adams, Berset, Kessler, & Hamburger, 2009; Bucko & Salamon, 2007; International Standard, 2007; McKay & Blumberg, 2006; Orav, Raal, & Arak, 2010). The main constituents of the Roman chamomile oil have been reported to be primarily angelate, tiglate and butyrate esters. In addition, the Roman chamomile oils often contain monoterpene and sesquiterepene derivatives. Juhua has not been as well studied as German and Roman chamomiles; however, several Chinese research groups have reported that the primary substances in Juhua essential oil were borneol, verbenyl acetate, eucalyptol, eudesm-7(11)-en-4-ol and lanceol (Gao et al., 2012; Guan et al., 2007; Liu, Xing, Chen, & Wang, 2007; Sun, Hua, Ye, Zheng, & Liang, 2010). The chemical compositions of chamomile essential oils may vary dramatically for samples collected from different geographical locations, cultivars, times of harvest, and provenance. Sample collection and handling techniques can also influence the chemical fingerprints especially for volatile oil constituents.

Gas chromatography coupled to mass spectrometry (GC/MS) is commonly applied for the analysis and profiling of volatile compounds in chamomile samples. So far, more than 200 compounds have been isolated and identified in chamomile. Due to the chemical complexity of chamomile, large chromatographic and spectral data sets are usually generated by the GC/MS analysis. Manual inspection of the data and determination of the chemical marker compounds would be time consuming and also oversimplify the process of assessing the potency and the wholeness-value of the plant materials.

Data mining has become a fundamental task in chemical analysis due to the large quantity of information created by modern instruments. Effective software tools are essential to address the vast amounts of very similar data produced by the GC/MS experiments in this study. In the past few years, software tools capable of rapid data mining procedures and aligning algorithms have been applied in many research areas such as food, agriculture, pharmaceutical, herbal medicine and dietary supplements for informative, discriminative, and predictive purposes associated with safety and quality (Berrueta, Alonso-Salces, & Heberger, 2007; Khattab, Abou-shoer, Harraz, & El-Ghazouly, 2010; Vaclavik, Lacina, Hajslova, & Zweigenbaum, 2011). While principle component analysis (PCA), an unsupervised analysis, has been commonly employed to observe variance in multivariate data sets and visualise data clustering, supervised classification methods that include class information in their models have been used effectively in class determination and prediction. These techniques have been applied to a wide variety of analytical data, such as chromatographic, spectrometric and spectroscopic for the purpose of profiling, fingerprinting, authentication, detection of adulteration and data interpretation (Baumann & Aronova, 2012; Serino, 2012; Tan et al., 2012).

In a companion study (Avula et al., 2014), an investigation using UHPLC-UV-QToF/MS for phenolic compounds in various chamomile samples was reported. This was the first published comparison study of Roman and German chamomiles along with *Chrysanthemum*. Partial least squares discriminant analysis (PLS-DA) was used to discriminate between commercial chamomile samples.

In the current study, an analytical method was developed and applied for the non-targeted volatile, non-polar compound analysis of various chamomile samples. An automatic data processing procedure was introduced for control of input variety, alignment of retention time and data reduction by different filters using various criteria. A predictive model was constructed based on the PLS-DA for classification and discrimination of different types of chamomile from the authenticated plant samples. This model was subsequently employed to evaluate commercial samples purported to contain chamomile. Finally, a unique set of m/z data was generated for each individual subgroup of chamomile by chemometric analysis to identify the marker compounds. The objective of this portion of the research was to ascertain and address the problems of botanical classification and differentiation of chamomiles used in commercial products and dietary supplements, and to meticulously reveal quality attributes of different types of chamomile.

2. Materials and methods

2.1. Chamomile samples

The investigated samples included 27 authenticated plants, 35 solid commercial products and 11 essential oils. Specimens of all samples are deposited at the botanical repository of the National Center for Natural Products Research (NCNPR), University of Mississippi (documented with NCNPR accession code). Authenticated *Chamaemelum nobile* samples (3076, 9254 and 11577) were obtained either from the cultivated, living collection of the Maynard W. Quimby Medicinal Plant Garden, or collected by the NCNPR, University of Mississippi. Sample 13163 was an AHP-verified botanical reference standard obtained from the American Herbal Pharmacopoeias (AHP). Authenticated *Matricaria chamomilla* samples (259, 2802, 9172 and 11680) were obtained from the cultivated, living collection of the Maynard W. Quimby Medicinal Plant Garden, or collected by the NCNPR, University of Mississippi. Other authenticated *Matricaria chamomilla* samples (11781, 12182 and 12213–12221) were provided by Missouri Botanical Garden. Authenticated *Chrysanthemum morifolium* samples (9414–9421) were provided by the Research and Inspection Center of Traditional Chinese Medicine and Ethnomedicine, National Institute for the Control of Pharmaceutical and Biological Products, China. The detailed information about the authenticated plant samples used for the construction of the sample class prediction model is summarised in S1 in the supplemental material. The solid commercial samples in various forms included crude drugs, capsules, tea bags, crude drugs mixed with other plant materials, powder and extracts. The essential oil samples (Roman and German) included oils obtained from the authenticated plant materials by the steam distillation method described in the British Pharmacopoeia and commercial oils obtained by undetermined extraction techniques. All the chamomile commercial samples were purchased at food supermarkets, local retail pharmacies or online from different countries.

2.2. Chemicals

n-Hexane was purchased from Sigma–Aldrich. A mixture of series alkanes (C_9H_{20}–$C_{22}H_{46}$) was used for the determination of the retention index, and the alkane standards were purchased from Poly-Science Corporation. The analytical standard, *n*-tridecane ($C_{13}H_{28}$) was selected as the internal standard, and was obtained from Poly-Science Corporation. β-Farnesene, α-bisabolol oxide A and B were used as the reference standards for compound identification and were purchased from Sigma–Aldrich. *E* and *Z*-1,6-dioxaspiro[4.4]non-3-ene, 2-(2,4-hexadiynylidene)- were isolated at the NCNPR, University of Mississippi. The purity of the standards (>95%) was determined by ^1H and ^{13}C NMR as well as GC/MS analysis.

2.3. Sample preparation

Solid samples were ground and homogenized to obtain a uniform matrix. About 1 g of the fine powder was accurately weighed