Journal of Pharmaceutical and Biomedical Analysis 114 (2015) 462–470

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

An integrated strategy of marker ingredients searching and near infrared spectroscopy rapid evaluation for the quality control of Chinese eaglewood

Guoyu Ding[a], Yan Nie[a], Yuanyuan Hou[a,*], Zenghui Liu[b], Aina Liu[a], Jiamin Peng[a], Min Jiang[a], Gang Bai[a,**]

[a] State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy, Tianjin Key Laboratory of Molecular Drug Research, Nankai University, Tianjin 300071, People's Republic of China
[b] Tianjin Institute of Agricultural Quality Standard and Testing Technology, Tianjin 300381, People's Republic of China

ARTICLE INFO

Article history:
Received 10 March 2015
Received in revised form 17 June 2015
Accepted 18 June 2015
Available online 22 June 2015

Keywords:
UPLC–Q/Tof
Near infrared spectroscopy
Principal component analysis
Chinese eaglewood
Quality control

ABSTRACT

The application of near infrared spectroscopy (NIRS) in traditional Chinese medicine (TCM) has usually been limited by its blindness to qualitative or quantitative multivariate analysis and because its chemical significance is easily ignored. Here, an integrated technology of ultra-performance liquid chromatography-quadrupole/time-of-flight mass spectrometry (UPLC–Q/Tof) and NIRS was proposed to set up a systematic quality control of Chinese Eaglewood (CE). UPLC–Q/Tof combined with principle compound analysis (PCA) was used to identify the marker ingredients of CE. Four types of highly oxidized 5,6,7,8-tetrahydro-2-(2-phenylethyl) chromones (THPECs) were then identified as potential markers to distinguish the authenticity of the CE. Based on the marker ingredients, the definite wavenumber intervals and spectral pretreatment pattern of NIRS were selected to act as an alternative evaluation technology directed against CE powder samples. Calibration equations were developed from the contents of the four markers, as detected by high performance liquid chromatography (HPLC) and NIRS data, using synergy interval partial least squares (siPLS) algorithm with Leave-One-Out (LOO) cross-validation. Using siPLS regression, satisfactory calibration statistics were obtained for the prediction of the marker ingredients. The correlation coefficient (r) between the predicted and reference results for the test set was used as an evaluation parameter for the models ($r > 0.9$). Hierarchical cluster analysis and partial least squares-discriminant analysis (PLS-DA) were applied to further analyze the quantitative results from NIRS. From this systematic method, 50 CE samples were divided into superior, qualified, unqualified, and fake samples, displaying a more elaborate division than PLS-DA, which is only based on whole NIR spectra or HPLC. This tandem technique of UPLC–Q/Tof and NIRS assessment presented in this work can be used as a rapid evaluation approach for the quality control of complicated herbal medicines.

Abbreviations: NIRS, near infrared spectroscopy; TCM, traditional Chinese medicine; UPLC–Q/Tof, ultra-performance liquid chromatography-quadrupole/time-of-flight mass spectrometry; CE, Chinese eaglewood; PCA, principle compound analysis; THPEC, 5,6,7,8-tetrahydro-2-(2-phenylethyl) chromone; HPLC, high performance liquid chromatography; siPLS„ synergy interval partial least squares; LOO, Leave-One-Out; PLS-DA, partial least squares-discriminant analysis; HCA, hierarchical cluster analysis; OPLS-DA, orthogonal partial least squares discriminant analysis; MDA, monoester diterpenoidaconitine; DDA, diesterditerpenoidaconitine; LS-SVM, least squares-support vector machine; RBF, radial basis function; PEC, 2-(2-phenylethyl)chromone; ME, mean normalization; AUTO, auto-scaling; MSC, multiple-scatter correction; SNV, standard normal variable transformation; DT, de-trend; one-DC, one-dimension convolution; two-DC, two-dimension convolution; RMSECV, root mean square error of cross validation; RMSEP, root mean square error of prediction; MW, molecular weight; KS, Kennard-Stone; FS, fake sample; SS, superior sample; QS, qualified sample; US, unqualified sample; LV, latent variable.

* Corresponding author. Fax: +86 022 23506930.
** Corresponding author. Fax +86 022 23506792.
E-mail addresses: houyy@nankai.edu.cn (Y. Hou), gangbai@nankai.edu.cn (G. Bai).

http://dx.doi.org/10.1016/j.jpba.2015.06.029
0731-7085/

1. Introduction

The fact that the effects of Traditional Chinese Medicine (TCM) are brought about by their complicated chemical constituents has created a critical demand for powerful analytical tools to perform chemical analysis, ensuring their efficacy, safety and quality. Ultra performance liquid chromatography–quadrupole/time-of-flight (UPLC–Q/Tof), based on chemical profiling, is well-suited to analyze the multi-components in complex herbal matrices due to its enhanced retention time, reproducibility, high resolution, and low separation time [1]. UPLC–Q/Tof technology combined with PCA or OPLS-DA has been developed to rapidly find potential chemical markers and characterize the distinction of similar biological materials for quality control in TCM [2,3].

The use of UPLC–Q/Tof along with multivariate statistical analysis is an effective and efficient approach to determine the potential chemical markers between similar biological materials. However, it is not suitable as a routine analysis because of its high application and maintenance costs. As a novel quality control technology, near infrared spectroscopy (NIRS) has its advantages over other analytical techniques in that it is a fast, easily used, and low cost method that records the spectra for solid and liquid samples with simple pretreatment in only a few minutes and allows for the quantification of multiple components [4–6]. A majority of research strategies before about pattern recognition were mainly aimed at the dimension detection of the spectral information[7–11]. However, unlike traditional chromatographic fingerprints, NIR spectra often consist of many overlapped bands and lack characteristic peaks. Therefore, it is difficult to assign specific absorption bands to specific functional groups, let alone to chemical components [12]. After the dimension detection method, it became more confused to help us to understand the detailed chemical significance. Therefore, So far, almost all researchers have faced a bottleneck in that an excellent cluster result was obtained by some complicated chemometrics, but without providing the crucial chemical markers to account for the reasons why TCM can be categorized by its different attributes [13,14].

Chinese eaglewood (CE) is the resinous heartwood of the *Aquilariasinensis* (Lour.) Gilg, a genus taxonomically belonging to the Thymelaeceae. The high-priced CE with sandalwood-ambergris tonalities has been used for centuries as incense in traditional religious ceremonies and plays an important role in TCM because of its effects, such as sedative, analgesic, carminative and digestive [15]. However, wild *Aquilaria* trees have decreased drastically to a point of near extinction, leading to a significant amount of adulterated and inferior CE to appear on the market. Chemical analyses of eaglewood show that 52% of its abundant components are 2-(2-phenylethyl) chromones (PECs) [16]. It is worth mentioning that highly oxidized 5,6,7,8-tetrahydro-2-(2-phenylethyl) chromones (THPECs) have not been detected in healthy (non-injured) *Aquilaria* trees and appear to be unique to eaglewood [15]. In the past several years, qualitative and quantitative methods, such as GC, HPLC–UV and tandem mass spectrometry, have been developed for CE quality appreciation [17–19]. However, these methods are time-consuming and expensive for commercial use; therefore, a simple and reliable method for rapid quality control remains called for.

In this study, citing CE as an example, an integrated strategy was developed to distinguish the authenticity of CE and estimate its quality based on the merits and demerits of UPLC–Q/Tof and NIRS techniques. UPLC–Q/Tof coupled with PCA was used to identify the marker ingredients from a small batch of complicated CE samples which has been graded to qualified and fake sample according to the 2010 edition of Chinese pharmacopeia. The loading plot of PCA guides us to find out the crucial marker ingredients for classification. Then, the synergy interval partial least squares (siPLS) algorithm was used to search the suitable quantitative wavebands of NIR for the marker ingredients, and a calibration model was built to quantify these crucial marker ingredients. Based on the crucial marker ingredients' quantitative results, we furtherly only apply the unsupervised pattern recognition to provide the elaborate distinguish of the whole CE samples. Finally, a simple and reliable method was provided for the quality control of solid CE samples.

2. Materials and methods

2.1. Chemicals, reagents, and materials

Agarotetrol was purified according to reported studies with some modifications, and the purity was determined to be more than 98% by HPLC–UV analysis [17]. The detailed information is listed in the Supplementary data (Figs. S1 and S2). Acetonitrile of HPLC grade was purchased from Merck (Darmstadt, Germany). Ultrapure water was prepared with a Milli-Q purification system (Millipore, Bedford, MA, USA). All other reagents were of analytical grade and purchased from Yifang S&T (Tianjin, China). The first batch of CE samples, containing 12 authentic and 3 fake samples were provided and confirmed by Zhongxin Pharmaceutical Group, Medicinal materials company (Tianjin, China) following the 2010 edition of Chinese pharmacopeia. The second batch of 35CE samples of unknown properties was purchased from the Hainan and Guangdong provinces of China. Detailed information relative to these samples are listed in Table S1.

2.2. Sample preparation

All of the CE samples were crushed to a powder by a cyclone mill, and the particle size of the powder was verified to be below 60 meshes. The dried powder (1 g) was extracted with 10 mL of alcohol-water (1:1, v/v) solution using an ultrasonic extraction apparatus (40 kHz, 500 W, Ningbo, China) for 30 min at room temperature. The same solution was used to replenish the extraction system upon solvent loss due to volatilization. The extracts were centrifuged at 12,000 rpm for 10 min, and the supernatant was used for UPLC–Q/Tof and HPLC analysis. The standard stock solution of agarotetrol was prepared by dissolving 5.11 mg in 5 mL alcohol–water (1:1, v/v) solution.

2.3. UPLC–Q/Tof analysis

A Waters Acquity UPLC system (Waters Co., Milford, Ma, USA) equipped with a photodiode array detector (PDA) and a Waters Q/Tof Premier Mass Spectrometer with an electrospray ionization system (Water MS Technologies, Manchester, UK) were used for sample analysis. The data acquisition was supported by the MassLynx V4.1 software. The separations were performed on a Waters Acquity UPLC BEH C_{18} column (2.1 mm × 100 mm, 1.7 μm) at 30 °C. The mobile phase system was acetonitrile (A) and water with 1% formic acid (B) at a flow rate of 0.4 mL/min and ran a gradient as follows: 0 min, 10% (v/v) A; 20 min, 55% (v/v) A; 22 min, 100% (v/v) A; 23 min, 100% (v/v) A. The compounds were detected by a PDA scanning from 200 nm to 400 nm. The sample injection volume for analysis was 5 μL. The ESI–MS spectra were acquired in the positive ion mode. The conditions for ESI–MS analysis were as follows: the capillary voltage was set to 3.5 kV, the sample cone voltage was set to 30 V, the desolvation gas flow was set to 600 L/h at 350 °C, the cone gas was set to 50 L/h, and the source temperature was 110 °C. The Q/Tof Premier acquisition rate was 0.1 s with a 0.02 s inter-scan delay. The MS spectra are acquired from 100 to 1000 Da. Leucine enkephalinamide acetate was used as the lock mass (*m/z* 555.2931 in ESI^+) at a concentration of 200 ng/mL and a flow rate