Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 205 (2018) 637–648

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

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

Comprehensive quality assessment of *Dendrubium officinale* using ATR-FTIR spectroscopy combined with random forest and support vector machine regression

Ye Wang [a,b], Heng-Yu Huang [b], Zhi-Tian Zuo [a,*], Yuan-Zhong Wang [a,*]

[a] *Institute of Medicinal Plants, Yunnan Academy of Agricultural Sciences, Kunming 650200, PR China*
[b] *College of Traditional Chinese Medicine, Yunnan University of Traditional Chinese Medicine, Kunming 650500, PR China*

A R T I C L E I N F O

Article history:
Received 17 May 2018
Received in revised form 25 July 2018
Accepted 30 July 2018
Available online 01 August 2018

Keywords:
Dendrobium officinale
Random forest
Support vector machine regression
Harvesting period
High-performance liquid chromatography
Attenuated total reflectance mid-infrared spectroscopy

A B S T R A C T

Dendrobium officinale, as a tonic herb, has attracted more and more consumers to consume in daily life. In order to protect the wild resource, the herb has made great progress though cultivation in vitro. However, the quality is fluctuated in Chinese herbal medicine market due to influence such as cultivated areas and harvesting period. Therefore, the herbal samples from different cultivated locations were evaluated with high-performance liquid chromatography with diode array detector (HPLC-DAD) in terms of two chemical components, quercetin and erianin. In addition, two markers in leaf and stem also were used for support vector machine regression (SVMR) prediction. Samples from different harvesting periods were also classified using attenuated total reflectance mid-infrared spectroscopy coupled with random forest model. The results indicated that Pu'er and Menghai in Yunnan Province were suitable places for the herb cultivation and the leaf of the herb was also an exploitable resource just in light of the content of two components. What's more, combination of suitable spectra pretreatment and grid search method efficiently improved the prediction performance of the regression model. The results of random forest model indicated that important variables combination between stem and leaf was an effective tool to predict the harvesting time of the herb with 94.44% accuracy in calibration set and 97.92% classification correct rate in validation set. The results of combination were better than the models using individual stem and leaf spectra. In addition, the suitable harvesting time (December) could be classified efficiently. Our study provides a reference for quality control of raw materials from *D. officinale* herb.

1. Introduction

Traditional Chinese medicine (TCM) is sustainable utilization resources as healing herbal materials. There are >11, 000 medical plant species used as TCM in China [1]. According to the statistics in above literature, the number of wild origins is approximately twice as that of artificial cultivation. Moreover, 600 Chinese medicines herbal species are added around 100,000 prescriptions with different proportion for different symptoms. However, two thirds of these 600 herbs species are from the output of wild herb collection. The increasing market demand leads to a situation that long-term and over-excavation wild herbal materials are collected by human driven by commercial benefits although ecology pressure couldn't recover again in a short period [2]. Therefore, human has to find an effective measurement to protect the wild resources. In vitro propagation of plants has become an important and commercial protocol for plant regeneration. The propagation depends on totipotency of cells that plant cells give rise to plants when providing comfortable nurseries and adequate nutrition. Compared with traditional seeding cultivation and vegetative means, tissue culture offers many key advantages such as rapid multiplication of valuable genotypes, production of disease-free plants and germplasm conservation [3]. The mature micro-propagation is also an effective measure to rescue the dangerous species when they are over collected or their inhibitions are destroyed by human impact.

Dendrobium officinale (Chinese: tie-pi-shi-hu) is the most popular and expensive species as a tonic herb, which is also incorrectly named as *Dendrobium candidum* in many published literatures previously [4]. As an epiphyte, whose stems mainly grow on the surface of bark or rock in the form of cluster where they require an annual average temperature of 12–18 °C and 1100–1500 mm annual rainfall [5]. Therefore, it is mainly distributed in the mountain of Yunnan, Guizhou, Guangxi and Fujian provinces where have appropriate environment for the growth of the species plant. In addition, seed germination of *D. officinale* depending on mycorrhizal fungi is also a great obstacle limiting the development and cultivation of *D. officinale* productions [6]. The harsh living environment of the plant and rigorous germination

* Corresponding authors at: Institute of Medicine Plants, Yunnan Academy of Agricultural Science, 2238, Beijing Road, Panlong District, Kunming 650200, PR China.
E-mail addresses: yaaszztian@126.com (Z.-T. Zuo), boletus@126.com (Y.-Z. Wang).

https://doi.org/10.1016/j.saa.2018.07.086
1386-1425/

condition of its seed lead a decreasing of production in application of clinic and ethnopharmacology research. Based on our prior knowledge, the resource of *D. officinale* is in production with mature protocol from tissue culture to offset obstacle in nature living condition.

In addition, the medical history of the plant could trace back to Tang dynasty in China, it was recorded in Taoist Canon, a classical book of Taoist school, where it was list the first choice among the nine tonic medical herbs in ancient times of China. Other eight herbs are dries flower of *Saussurea involucrata* from Tianshan Mountains, around 150 g dried root and rhizome of *Panax ginseng*, dried earthnut of *Polygonum multiflorum*, dried sclerotium of *Poriacocos*, succulent stem of *Cistanche deserticola*, dried sporocarp of *Ganoder malucidum* and *Ganoder masinense*, Margarita form deep ocean and *Cordyceps*. These nine classical herbals are also used in daily life in China as functional food to improve body condition [7]. Among the nine classical herbs, *D. officinale* is mainly responsible for nourishing *Yin*, clearing heat, and supplementing the stomach as well as promoting the regeneration of body fluids [8].

However, the quality of the herbal medicine is fluctuant with their productive area, harvesting time and plant sources. Therefore, it has become a critical step for their quality control using suitable techniques. Since 1960s, the infrared spectra combined various accessories has become a popular technique for quality control purpose of traditional medicine [9], because spectroscopic method emphasized the integrative and holistic characteristics of herbal medicine [10]. Compared with other kinds of infrared spectroscopy, attenuated total reflection infrared (ATR-IR) spectroscopy has more advantages such as convenience, simplicity and effectiveness. Furthermore, sample needn't tedious preparation and specialized sample holder such as KBr for transmission analysis. On the one hand, the trait could eliminate the additional spectra of these holders and structural changes associated with chemical preparation of samples. On the other hand, the detected samples keep in their natural state, whose spectrum could reflect their original chemical information of various metabolites. Based on the existing literatures, ATR-IR spectra combined with chemometrics have been widely applied for quality assessment and authentication of herbal medicines [11,12], especially of food [12–18]. As far as we concerned, there were not published literatures concerning harvesting period discrimination and chemical component prediction of *D. officinale* using ATR-IR spectra.

In the absence of prior knowledge, samples of *D. officinale* from three cultivated areas in Yunnan Province were evaluated using high-performance liquid chromatography and attenuated total reflection infrared (ATR-IR) spectroscopy combined with support vector machine regression. The same botanical origins harvested from January to December in the same year were classified via random forest model. Further, the present research primarily focused on advantages of ATR-IR spectroscopy in the application of harvesting period discrimination and marks content prediction in *D. officinale*.

2. Materials and Methods

2.1. Samples and Chemicals

Both of two cultivated species of *D. officinale* were obtained by a protocol that has applied in production from protocorm to mature plant in Yuxi City of Yunnan Province (Fig. 1). One species was transplanted in greenhouse in Pe'er, Puwen and Menghai cities located in the southwest of Yunnan to select the most suitable cultivated location. The distribution of three cities in Yunnan Provinces was displayed in Fig. 2. The other was left in original greenhouse of Yuxi to collect 10 individuals every month at the same time point from January to December. Total 29 samples from three areas and 120 individuals of different harvesting time were collected for further analysis. All of these samples were air-dried at room temperature and smashed into powder through a 60-mesh sieve for HPLC and ATR-IR analysis.

Two reference standards, quercetin and erianin with >98% purities, were purchased from National Institutes for Food and Drug Control (Beijing, China) for content analysis evaluating the influence caused by growth environment. Analytical grade methanol as extraction solvent was obtained from Tianjin Fengchuan Chemical Reagent Technologies Co., Ltd. (Tianjin, China). MeOH and acetonitrile of HPLC grade were obtained from Thermo Fisher Scientific (Fair Lawn, NJ, USA), and formic acid of HPLC grade was obtained from Dikmapure (Lake Forest, CA, USA). Ultrapure water (2.22 μs/cm) conductivity in the whole HPLC analysis was provided by Hangzhou Wahaha Group (Hangzhou, China).

2.2. Chromatographic Analyses

1 mg sample powder was weighed precisely with an electronic balance (Precisa, Switzerland), and extracted with 80% methanol (1.5 ml) solution by ultrasound-assisted method (500 W) for 30 min after 30 min of immersion at room temperature. The extracts were filtrated through a filter paper after replenishing volatilization solution. Filtrate was though a 0.22 μm millipore filter (Millipore, USA) before injecting into the auto-sampler vials. The mobile phase was combined with two solvents: the water phase with 0.1% formic acid (A) and organic phase

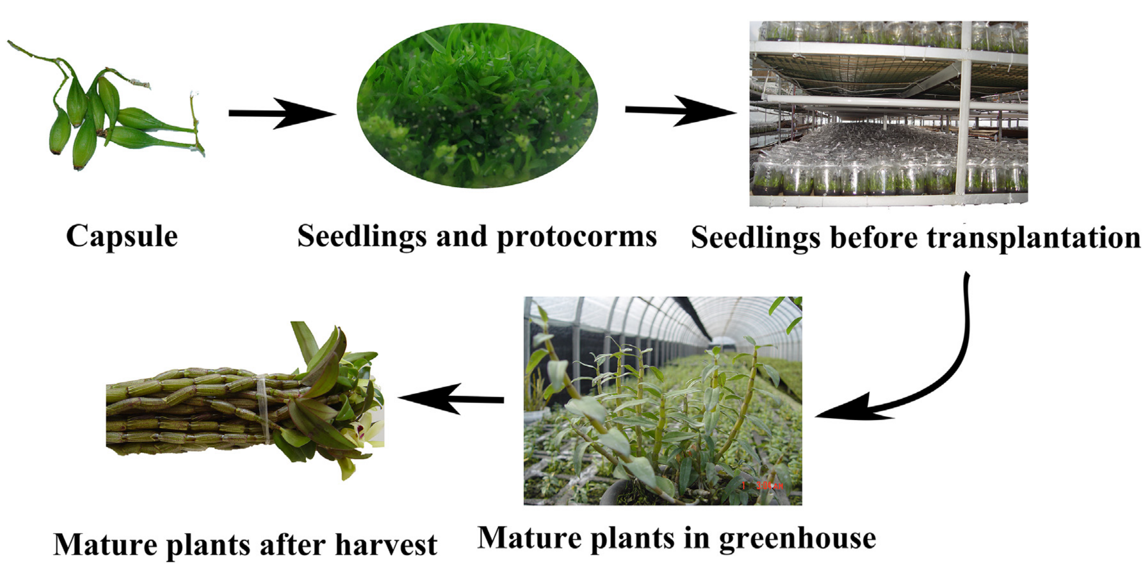

Fig. 1. The tissue culture protocol of *D. officinale*.