中 国 药 科 大 学 学 报
Journal of China Pharmaceutical University 1990; 21(2): 91 – 4

Classification of Chinese Traditional Medicine–Shanyao (*Dioscorea opposita* Thunb.) By Pyrolysis / High Resolution Gas Chromatography / Pattern Recognition

Fang Xingchun, Yin Xia[1], Xiang Bingren, An Dengkui

Department of Pharmaceutical Analysis; [1]Analysis and Computer Center

Abstract Pyrolysis / high resolution gas chromatography / pattern recognition was used to develop a potential technique for discriminating Chinese Traditional Medicine–Shanyao (*Dioscorea opposita* Thunb.). About 1 mg of crude drug powder was pyrolyzed in the furnance pyrolyzer and pyrolytical products were directly carried into gas chromatograph with FSOT column (L28 m × 0.28 mm, coated with SE–30, df 0.26 μm). The Py–GC data were subjected to multivariate statistical analysis. Principal component analysis provided the discrimination of five varieties of 19 samples purchased from market. Between Shanyao variety (included in Chinese Pharmacopoeia)–*D. opposita* and other four varieties–*D. japonica*, *D. persimilis*, *D. alata* and *D. fordii* were differentiated from 19 samples purchased from market.

Key words Shanyao; Chinese Traditional Medicine; Py–GC–PaRc; *D. opposita; D. alata; D. persimilis; D. fordii; D. japonica*

The Shanyao (*Dioscorea opposita* Thunb.) –a Chinese crude drug has been used in Chinese Traditional Medicine for hundreds of years. It has been proved to cure diabetes, chronic inflamation of intestine, chronic hepatitis, asthma and cough by modern pharmacology. Shanyao variety included in Chinese Pharmacopeia must be *D. opposita*. The chief constituents contained in Shanyao are allantoin, diosgenin, batatasine, abscisin, trace elements, amino acids, etc. Shanyao is also used for food in folks because of its nutritive value. Other varieties of crude drug in *Dioscorea* may be sold as *D. opposita* on the market. So, the identification of Shanyao samples purchased from the markets is important. The identification of "Shanyao", completed by qualified, experienced pharmacognosist, is usually complicate, difficult and time comsuming. The accuracy of microscopic identification may be mainly dependent upon the operator's experience. Pyrolysis–gas chromatography (Py–GC) is a rapid and reliable method for distinguishing biological organism[1–3]. The use of Py–GC as a powerful analytical technique in Chinese crude drug has also been demonstrated in our laboratory[4]. Py–GC is a degradative technique in which the molecular fabric of plant is thermally fragmented in the absence of oxygen. The resulting pyrolyzate components are then chromatographed by GC. Pyrograms of crude drug powder are very complex and overall patterns of variation are not easily discernible by visual inspection. Therefore, multivariate pattern recognition method is utilized to analyze the pyrograms and elucidate the chemical features (pyrogram peaks) that could be discriminated between plants[5]. In this paper principal component analysis is applied to data obtained from Py–GC profiles of Shanyao purchased from the market in order to differentiate Shanyao and Shanyao substitutes.

Materials and Methods

Sample preparation: 19 Shayao samples purchased from the market were ground into powder (60 mesh). The varieties of samples were identified to be *D. opposita, D. alata, D. persimilis, D. japonica and D. fordii* by pharmacognosist. (see Tab 1.)

Pyrolysis gas chromatographic analysis: About 1 mg of crude drug powder was placed on the platinum boat, then sample rod with platinum boat was introduced into PY–2A furnance pyrolyzer connected directly to the inlet of chromatograph at furnance temperature 500℃ and stayed for 1 min in furnace. The resulting pyrolysis products were resolved on a 28 m (0.28 mm I. D., 0.26 μm df) SE–30 fused silica open tubular column using Shimadzu GC–9A gas chromatograph equipped with a flame ionization detector. Highly pure Nitrogen (99.999%) was used as the carrier gas. Column pressure was 1 kg / cm^2. The split ratio was 1 : 177. The injector and detector oven was maintained at 250℃. The temperature programming was: initial temperature 50℃ which was maintained for 3 min; from 50℃ to 150℃ at a rate of 5℃ / min; from 150℃

Received Sept. 14, 1989

to 200℃ at a rate of 3℃ / min. All analyses were done in duplicate or triplicate. The experimental order of arrangement was listed in Tab 1.

Tab 1. Data on the Shanyao samples studied

Sample No.	Repetitive No.	Botanical sources	Markets (crude drug company)
1	1, 2, 3	*D.opposita*	Hunan
2	4, 5, 6	*D.fordii*	not known
3	7, 8, 9	*D.opposita*	Shijiazhang
4	10, 11, 12	*D.opposita*	Anhui
5	13, 14,	*D.opposita*	Jiangsu
6	15, 16, 17	*D.opposita*	Zhejiang
7	18, 19, 20	*D.japonica*	not known
8	21, 22, 23	*D.opposita*	Henan
9	24, 25, 26	*D.persimilis*	Nanjing
10	27, 28, 29	*D.alata*	Nanjing
11	30, 31, 32	*D.opposita*	Henan
12	33, 34, 35	*D.opposita*	Hebei
13	36, 37, 38	*D.opposita*	Yunnan
14	39, 40, 41	*D.opposita*	Guangxi
15	42, 43, 44	*D.opposita*	Jianxi
16	45, 46, 47	*D.opposita*	Hunan
17	48, 49, 50	*D.opposita*	Zhejiang
18	51, 52, 53	*D.opposita*	Sichuan
19	54, 55, 56	*D.opposita*	Guangdong

For data processing, an off-line computer system was used. The gas chromatographic peak areas were obtained from a Shimadzu CR-3A integrator. PY-GC profile for Shanyao sample was showed in Fig 1.

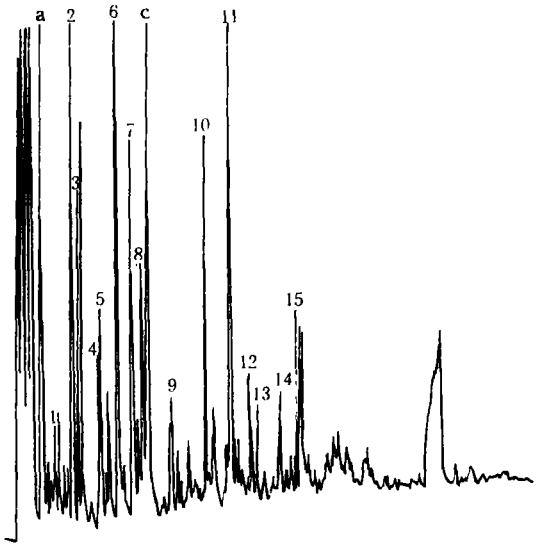

Fig 1. Py-GC profile of Shanyao crude drug powder. a, c are the nominated reference peaks

Numerical methods: The most precise method for the presentation of pyrolysis gas chromatograms to the computer is to give the retention time and peak area of individual thermal degradation product for all pyrograms available. The Py-GC data is subjected to some variability in that retention times exhibit small changes, due to progressive changes in column efficiency and operating parameters. For automatic comparisons, arrangement rather than a specific retention time must be taken to ensure that corresponding peaks are matched correctly. This problem can readily be overcome by a standardization procedure of the retantion times[6]. Two reference peaks are chosen from those in the pyrogram. These should be present in all pyrograms, be prominent, well resolved, and easily identified and also have approximately the same ratio to each other from sample to sample. The standardized retention times of all peaks are calculated by equation 1.

$$t_b^s = t_a^s + \frac{(t_b^r - t_a^r)(t_c^s - t_a^s)}{(t_c^r - t_a^r)} \quad (1)$$

where t^r indicates the raw retention times and t^s the standardized retention time of components a, b, c. a and c are the nominated reference peaks, b indicates the peak to be calculated.

The pyrolysis chromatogram obtained from each sample was digitized using the peak areas of 15 peaks occurring in all 56 chromatograms. Each chromatographic data vector is then normalized to the sum 100 over the 15 peaks. Thus a 56× 15 data matrix constitutes the original data matrix X(N × M); x(nm) is the value of variable m measured on experimental order n.

PCA (Principal Component Analysis) is probably the best known and most commonly used display method. The basic idea of this method is to reduce the dimensionality of multidimensional data while retaining as much of the variation in the data as possible, thus enabling the direct examination of the relative positions of the data points (pyrograms) in the high-dimensional space. New variables, called PC (principal component), were calculated as linear combinations of the original variables in such a way that the first PC takes up as much as possible of the variance present in the original variables, the second PC accounts for as much of the remaining variance as possible, and so on. Because the axes of the first two or three PCs account for most of the variance, the relative positions of data points in the high- dimensional space can be visualized by plotting the first two or three PCs.

PCA method was implemented using IBM PC / XT microcomputer and the program in True BASIC was written by the author following the equations of ref[7].

Results and Discussion

PCA was applied to the entire data set in Table 1. PCA results were shown in Fig 2. Four varieties of herbs, *D. fordii* (4, 5, 6), *D. persimilis* (24, 25, 26), *D. alata* (27, 28, 29) and *D. japonica* (18, 19, 20) were clearly separated each other also separated from *D. opposita* which is included in Chinese Pharmacopoeia. The result of PCA indicated that this data set is best described utilizing two principal component directions, although no assumptions about similarities or dissimilarities between varieties were made. The difference between *D. opposita* and four varieties of *Dioscorea* was shown. Only after we make comparisons with their pharmacological activities, can we decide whether the Shanyao purchased from market can be considered as genuine Shanyao used in Chinese traditional medicine. The false product takes up 20% of

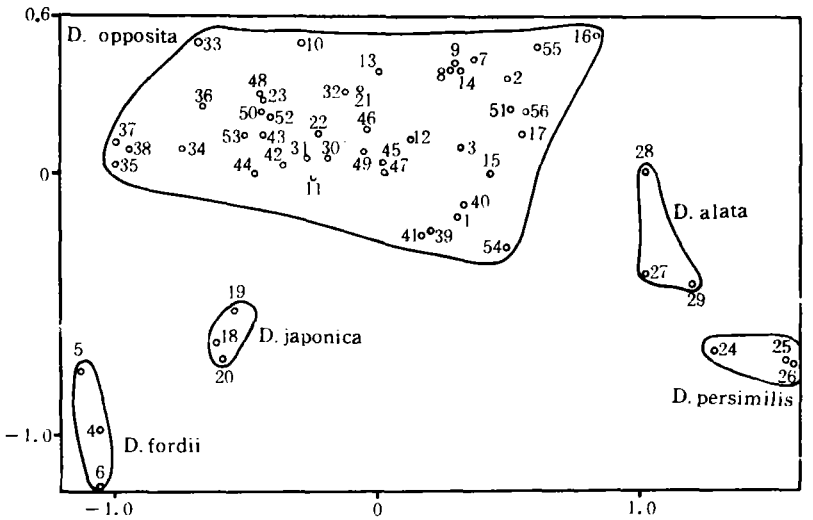

Fig 2. First and second PCs from PCA of Shanyao samples in Table 1.

Shanyao samples purchased from the markets.

Pyrolysis–gas chromatography can be used to obtain a chemical fingerprint of the herbs in terms of chromatograms. The use of these chromatograms for the classification of different samples in same family is often difficult, owing to their characteristic peaks of different varieties in same family. On the other hand, the large apparent variability of repetitive chromatograms was measured on the same type of samples. Of course, part of the variability between repetitive chromatograms measured on the same type of herbs was systematic. This systematic variability could be overcome automatically by means of the normalization of peak areas and pattern recognition. In the same time, pattern recognition technique can be used to handle the complex chromatogram to give the correct classification of herbs.

The combination of Py–GC and PCA (Py–GC–Pr) gives a good classification in the present example of 19 Shanyao samples chosen to illustrate this potential method. This method also has its simplicity, rapidity and reliability. Without any prior chemical treatment, micro quantity of sample was sufficient to complete the identification. Therfore, it was valuable especially for expensive herbs. The identification of herbs by pharmacognosist is very important, hence this paper must be seen as an illustration of the possibilities of the methodology, not as a final method describing a working identification of herbs.

Key words Shanyao; Chinese Traditional Medicine; PyGC–PaRc; *D. opposita; D. alata; D. persomolis; D. fordii; D. japonica*

Acknowledgements: The authors wish to thank Ms. Hang Yueyu for contributing part of samples and identifying the botanical source of all samples.

References

1 Blomquist G, Johansson E, Sodersterom, *et al.* Classification of fungi by means of pyrolysis–gas chromatography–pattern recognition. *J Chromatogr* 1979; **173**: 19–32

2 Kullik E, Kaljurand M, Koel M. Analysis of Pyrolysis gas chromatograms using pattern recognition technique. *J Chromatogr* 1976; **126**: 249–56

3 Valcarce R, Smith GG. Chemical characterization of honey bees by curie-point pyrolysis–gas chromato-

graphy—pattern recognition. *Chemometrics and Intelligent Laboratory System.* 1989; 6 : 157–66
4 房杏春，相秉仁，安登魁. 裂解气相色谱模式识别在牡丹皮分析中的应用. 药学学报 (待发表)
5 相秉仁，安登魁. 模式识别及其在药物分析中的应用. 国外医学 药学分册 1983; 10(4) : 204–21
6 Irwin WJ. *Analytical Pyrolysis: A Comprehensive Guide.* New York: Marcel Dekker, 1982; 237–90
7 Thielemans A, Lewi PJ, Massart DL. Similarities and differences among multivariate display techniques illustrated by belgian cancer mortality distribution data. *Chemometrics and Intelligent Laboratory Systems* 1988; 3 : 277–300

裂解-高分辨气相色谱-模式识别在山药鉴别分析中的应用

房杏春 殷 霞[1] 相秉仁 安登魁

(药物分析研究室; [1]分析计算中心)

应用裂解-高分辨气相色谱-模式识别技术鉴别分析中药山药，从裂解-高分辨气相色谱图中选取 15 个特征峰作为分类指标，应用主成分分析法进行分类识别，有效地区分了"山药"样品中药典规定正品和非药典正品，结果与生药学方法鉴定一致. 应用本法分析中药操作简易，分析快速，无需化学预处理，所得数据经模式识别处理，结果客观可靠，有可能成为中药鉴定分析的方法.

关键词 山药; 中药; 裂解-高分辨气相色谱-模式识别; 薯蓣; 参薯; 褐苞薯蓣; 山薯; 日本薯蓣

国际著名化学计量学家 Massart 教授来我校讲学

欧洲化学计量学协会主席、比利时自由大学医药学院副院长 D.L.Massart 教授应邀于 2 月 14 日到 3 月 6 日来校讲学. 在此期间 Massart 教授为我校师生和中国药学会委托我校举办的"计算机在药物分析上的应用学习班"的学员共 60 余人全面、系统地讲授了计算机在化学上的应用 (Chemometrics). 其中包括数理统计应用、实验方法设计和优化到因子分析、聚类分析、模式识别和专家系统等. 有些方法，如回归分析中的 Robust 法、单中位数法、重复中位数法、多重权重法等以及实验设计中的 Dochlert 均匀设计、不完全中心组成设计等都是较新的方法，使国内同行更新了概念，开阔了视野. Massart 教授还对一些国内不常用的而国外应用较多的方法，如非参数检验、模糊回归、主成分回归以及相关因子分析等也都作了较为详细的讨论. Massart 教授的系统讲演为我校计算药物分析这门新兴学科的建设起了支持和促进的作用.

(漫 绿)