Journal of Chromatography, 514 (1990) 287–292
Elsevier Science Publishers B.V., Amsterdam

CHROM. 22 511

Application of pyrolysis–high-resolution gas chromatography–pattern recognition to the identification of the Chinese traditional medicine Mai Dong

XINGCHUN FANG
Department of Pharmaceutical Analysis, China Pharmaceutical University, Nanjing 210009 (China)
BOYANG YU
Department of Pharmacognosy, China Pharmaceutical University, Nanjing 210009 (China)
and
BINGREN XIANG and DENGKUI AN*
Department of Pharmaceutical Analysis, China Pharmaceutical University, Nanjing 210009 (China)
(First received January 23rd, 1990; revised manuscript received April 12th, 1990)

ABSTRACT

Pyrolysis–high-resolution gas chromatography–pattern recognition (Py–HRGC–PaRe) was used to develop a potential technique for identifying the Chinese traditional medicine Mai Dong. About 1 mg of crude drug powder was pyrolysed in a furnace pyrolyser and the products were directly carried into a gas chromatograph with an FSOT capillary column (30 m × 0.265 mm I.D.) coated with DB-1701 (d_f 0.25 μm). The Py–HRGC data were analysed by non-linear mapping PaRe. The results showed that Mai Dong samples could be classified into two categories: *Ophiopogon japonicus* (L.f.) Ker-Gawl (included in the Chinese Pharmacopoeia) and *Liriope spicata*.

INTRODUCTION

Mai Dong is a Chinese traditional drug widely used for the treatment of dipsosis, dry throat, tussiculation bloody sputum, palpitations and anxiety. Mai Dong as specified in the Chinese Pharmacopoeia must be the tuberous root of the plant *Ophiopogon japonicus* (L.f.) Ker-Gawl, but a recent survey of the drug revealed that the general name Mai Dong could involve the tuberous roots of several plants of different origins. They belong mainly to the genuses *Ophiopogon* and *Liriope*. Other varieties of crude drug may also be sold as Mai Dong. These unofficial products make up around 40% of Mai Dong samples purchased from the markets, hence the identification of Mai Dong samples sold in the markets is very important. Usually, the botanical source of a Mai Dong sample can be identified morphologically by a qualified, experienced pharmacognosist, but this process is complicated and time consuming. The accuracy of the morphology may also be affected by environmental factors.

0021-9673/90/$03.50

Analytical pyrolysis–gas chromatography (Py–GC) is a rapid and reliable method for distinguishing biological organisms, polymers, etc.[1–4]. Py–GC is a degradative technique in which the molecular fabric of a plant is thermally fragmented in the absence of oxygen. The resulting pyrolysate components are then subjected to GC. Pyrograms of crude drug powders are very complex and overall patterns of variation are not easily discernible by visual inspection. Therefore, a multivariate pattern recognition (PaRe) method has been utilized to analyse the pyrograms and elucidate the chemical features (pyrogram peaks) that could be used to discriminate plants[5]. The use of Py–high-resolution (HR) GC–PaRe as a powerful technique for the identification of Chinese drugs has also been demonstrated in our laboratory[6,7]. In this work, non-linear mapping[8] was applied to data obtained from Py–HRGC profiles of Mai Dong samples purchased from the market in order to distinguish Mai Dong and its substitutes.

EXPERIMENTAL

Sample preparation

Seventeen Mai Dong samples purchased from the market were ground into a fine powder (40 mesh). The varieties present were identified to be *O. japonicus* and *L. spicata* (see Table I).

Pyrolysis–high-resolution gas chromatography

A quartz tube was heated at 500°C using a Shimadzu Py-2A furnace pyrolyser. About 1 mg of crude drug powder was placed on the platinum boat, then the sample rod with the platinum boat was introduced into the quartz tube connected directly in the inlet of the chromatograph and remained there for 1 min. The resulting pyrolysate products were separated on a J&W Scientific DB-1701 fused-silica capillary column (30 m × 0.265 mm I.D.; d_f 0.25 μm) using a Shimadzu GC-9A gas chromatograph equipped with a flame ionization detector. High-purity nitrogen (99.999%) was used as the carrier gas and peak areas were obtained from a Shimadzu CR-3A Chromatopak integrator. The injector and detector temperature was maintained at 250°C. The gas chromatograph oven was maintained at 50°C for 3 min, then heated

TABLE I

BOTANICAL SOURCES AND LOCATIONS OF MAI DONG SAMPLES STUDIED

Sample No.	*Botanical source*	*Location*[a]
1–10	*O. japonicus*	Zhejiang[b], Yunan, Jiangxi, Sichuan HubeiXianning, Guizhou, Hunan
11–17	*L. spicata*	Hubei[c], Fujian, Henan, Guangxi[d], Shanxi

[a] Location of collection (market).
[b] Four batches of samples.
[c] *L. spicata* (Thunb.) Lour. var. *prolifera* Y. T. Ma, called Hubei Mai Dong.
[d] Three batches of samples.

from 50 to 150°C at 5°C/min and from 150 to 200°C at 3°C/min. A typical pyrolysis gas chromatogram of the sample is shown in Fig. 1.

Numerical method

The resulting pyrograms (retention times *vs.* percentage of peak areas) are represented by data vectors $X_l(x_{l1}, x_{l2}, x_{l3}, \ldots, x_{lj}, \ldots, x_{lm})$, where component x_{lj} is the percentage area of the jth peak in the lth sample. The feature selction is important for pattern recognition because more features give rise to difficulties in the classification between the plants. If we have two classes, class r and class k, the feature selection is as follows[9]:

$$W_j = \frac{|\bar{x}_j^k - \bar{x}_j^r|}{s_j^{rk}} \tag{1}$$

where W_j is weighing factor, $\bar{x}_j^k$ is the average percentage area of the jth peak in the kth class:

$$\bar{x}_j^k = \frac{1}{N_k} \sum_{i=1}^{N_k} x_{ij} \tag{2}$$

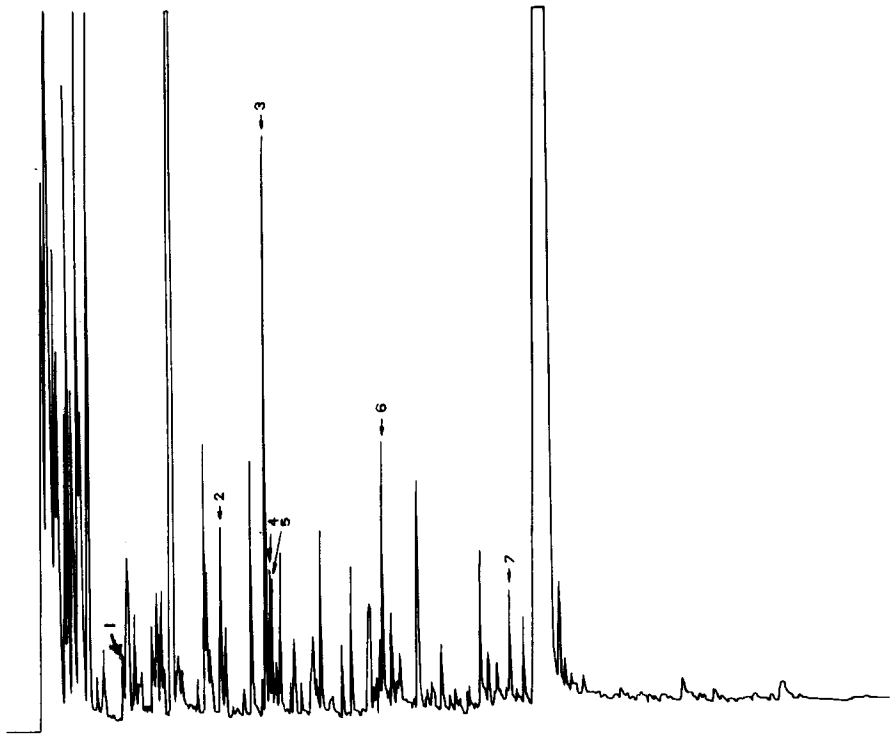

Fig. 1. Typical Py–GC profile of the sample.

N_k is number of sample in class k.

$$s_j^{rk} = \{2(s_j^r)^2 (s_j^k)^2 / [(s_j^k)^2 + (s_j^r)^2]\}^{\frac{1}{2}} \tag{3}$$

s_j^r is the standard deviation of percentage area of the jth peak of all samples in class r.

Seventeen samples were divided into classes (see Table I). We selected seven features (peaks) as the original data vector of pattern recognition based on the rule of Max W_j. Each sample data vector was then normalized to sum 100 over the seven peaks. Thus, a 17 × 7 data matrix constituted the empirical material for the non-linear mapping (NLM).

NLM by error minimization according to Sammon[10] is a display method that is commonly used. The method starts by calculating the m variable means $\bar{x}_j$ (j = 1–m) as

$$\bar{x}_j = \frac{1}{N} \sum_{i=1}^{N} x_{ij} \quad (j = 1,2, \ldots, m) \tag{4}$$

where m is the number of variables (peaks) and N is the total number of points (objects). Then the covariance matrix C is generated, each element C_{ij} of which compares variables i and j as

$$C_{ij} = \sum_{l=1}^{N} (x_{li} - \bar{x}_i)(x_{lj} - \bar{x}_j) \tag{5}$$

Next, the eigenvalues λ_j and eigenvectors μ_j for j = 1–m are calculated by solving

$$C\mu_j = \lambda_j \mu_j \tag{6}$$

The basis vectors μ_1 and μ_2, which correspond to the two largest eigenvalues λ_1 and λ_2, are used as the starting configuration for NLM map. All of the m-space interpoint distances, d_{ij}^*, are calculated as

$$d_{ij}^* = \left[\sum_{k=1}^{m} (x_{ik} - x_{jk})^2 \right]^{\frac{1}{2}} \tag{7}$$

and all of the two-space interpoint distances, d_{ij}, are calculated as

$$d_{ij} = [(y_{i1}^* - y_{j1})^2 + (y_{i2} - y_{j2})^2]^{\frac{1}{2}} \tag{9}$$

where the y values are found by the rotation matrix that diagonalizes C in eqn. 6. The object here is iteratively to change the two coordinates (y_{l1} and y_{l2}) for each point y_l so as to minimize an error function E, defined as

$$E = \frac{1}{\sum_{i<j} d_{ij}^*} \sum_{i<j} \frac{(d_{ij}^* - d_{ij})^2}{d_{ij}^*} \tag{10}$$

The minimization attempts to preserve interpoint distances by finding d_{ij}s that are as close as possible to d_{ij}^*s.

In order to change iteratively the two-space coordinate and minimize E, a gradient method should be used. Sammon[10] suggested the method of steepest descent. As this method is adequately described elsewhere[10], the details will not be given here.

NLM was implemented using an IBM PC/XT microcomputer and the program in True BASIC was written by the authors.

RESULTS AND DISCUSSION

NLM was applied to the 17 × 7 data matrix in Table I. NLM results are given in Fig. 2.

Py–GC can be used to obtain a chemical fingerprint of the herbs. Efforts to use these chromatograms for the classification of different samples in the same family have often been difficult, owing to the 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. At the same time, pattern recognition techniques can be used to handle the complex chromatograms to give the correct classification of herbs.

In Fig. 2, *O. japonicus* was clearly separated from *L. spicata*. The botanical source of Mai Dong included in the Chinese Pharmacopoeia must be the tuberous root of the plant *O. japonicus* (L.f.) Ker-Gawl. The family of *L. spicata* contains two varieties, *L. spicata* (Thunb.) Lour and *L. spicata* (Thunb.) Lour. var. *prolifera* Y. T. Ma (from Hubei, also called Hubei Mai Dong) which is debated by

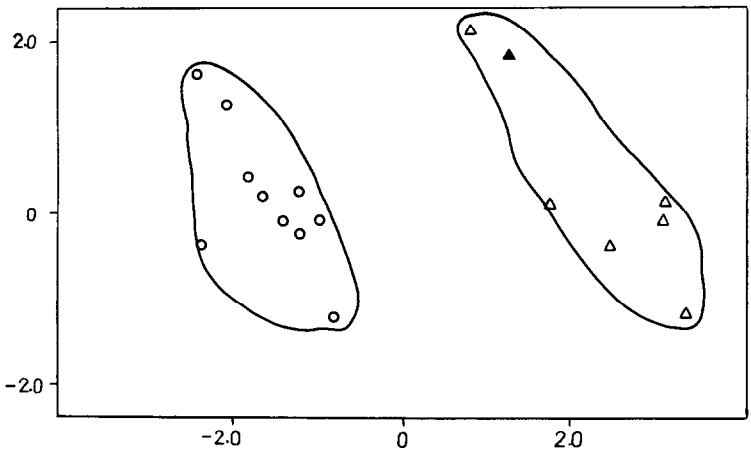

Fig. 2. Classification of Mai Dong samples obtained by NLM. ○ = *O. japonicus* (L.f.) Ker-Gawl; △ = *L. spicata* (Thunb.) Lour.; ▲ = *L. spicata* (Thunb.) Lour. var. *prolifera* Y. T. Ma. Horizontal axis represents y_1, vertical axis represents y_2.

pharmamognosists; some think that Hubei Mai Dong is same as *L. spicata* (Thunb.) Lour. In this paper, we use Py–HRGC–PaRe to attempt a chemical classification of these two type of samples; this method was not able to distinguish between the different varieties in the same family.

The combination of Py–HRGC and NLM (Py–HRGC–PaRe) gave a good classification with the present example of 17 Mai Dong samples chosen to illustrate this potential method. This method also has simplicity, rapidity and reliability. Without any prior chemical treatment, micro-amounts were sufficient to complete the identification, which makes the technique especially valuable for use with expensive or rare herbs. The identification of herbs by pharmacognosists is very important, hence this work must be seen as an illustration of the possibilities of the methodology being developed for the physical and chemical identification of Chinese drugs and not as a final method for a working identification of herbs. However, the application of Py–HRGC–PaRe to the identification of herbs will provide an independent means of classifying these herbs.

REFERENCES

1 W. J. Irwin, *Analytical Pyrolysis: a Comprehensive Guide*, Marcel Dekker, New York, 1982.
2 G. Blomquist, E. Johansson, B. Söderström and S. Wold, *J Chromatogr.*, 173 (1979) 19.
3 E. Kullik, M. Kaljurand and M. Koel, *J. Chromatogr.*, 126 (1976) 249.
4 R. Valcaece and G. G. Smith, *Chemometr. Intell. Lab. Systems*, 6 (1989) 157.
5 B. R. Xiang and D. K. An, *Guowai Yixue Yaoxue Fence*, 10 (1983) 204.
6 X. C. Fang, B. R. Xiang and D. K. An, *Acta Pharm. Sin.*, 25 (1990) in press.
7 X. C. Fang, X. Yin, B. R. Xiang and D. K. An, *J. China Pharm. Univ.*, 21 (1990) 91.
8 D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte and L. Kaufman, *Chemometrics: a Textbook*, Elsevier, Amsterdam, 1988, p. 366.
9 I. Juricskay and G. E. Veress, *Anal. Chim. Acta*, 171 (1985) 61.
10 W. J. Sammon, Jr., *IEEE Trans. Comput.*, C-18 (1969) 401.