

红外光谱阵列相关系数法在丹参鉴别中的应用

储德韧, 周 群, 郁 露, 孙素琴*

清华大学化学系, 生命有机磷化学及化学生物学教育部重点实验室, 北京 100084

摘 要 采用红外光谱阵列相关系数比对法对中国传统中药——丹参的产地与生长环境进行了鉴别研究。结果表明, 对特定谱段范围内设置较高相关系数阈值进行比对能快速鉴别丹参产地; 对 R5—R7 谱段范围内设置较低相关系数阈值的设置能对丹参进行生长环境分类识别。该方法对丹参的聚类识别具有快速、简便、准确的特点。

关键词 傅里叶变换红外光谱; 阵列相关系数; 丹参; 聚类识别

中图分类号: O657.3 文献标识码: A 文章编号: 1000-0593(2007)09-1706-04

引 言

丹参为唇形科鼠尾草属植物丹参(*Salvia miltiorrhiza* Bge.)的干燥根茎, 全国各地均有出产。一直以来, 产地的多元性给丹参的鉴别工作提出了挑战。国内外很多学者利用紫外(UV)、色谱(HPLC\TLC\GC)、质谱(MS)等多种方法对丹参加以检测^[1-4]。但丹参作为一种复杂的混合物体系, 通过各组分之间的相互作用达到其疗效, 因此不能简单的以某一两种成分含量作为质量标准。红外光谱(IR)能够对于中药全组分进行测定, 较为完整地反映整个体系的信息^[5-10]。鉴于同种药材的红外光谱具有较大的相似性, 若对谱图整体考虑则会引进较多的不可比数据, 而单一波段光谱比对在丹参鉴别中专一性又较差^[11]。本文采用多波段谱图比对软件将整个中红外区分为10段, 基于阵列相关系数比对方法, 对11个省50批次的425份丹参谱图样品进行分析, 凭借各段和总体相关系数差异对丹参进行快速鉴别。

1 实验部分

1.1 样品来源

本实验中使用的药材均由各地正规渠道提供的真品丹参。

根据产地的地域性差异, 先将50种样品分为, 北方组(黄河流域各省), 包括山东、山西、河南、河北、甘肃省, 共计37种样品; 南方组(长江流域各省), 包括四川、浙江、云

南、广西、贵州省、重庆市, 共计13种样品(见表1)。

Table 1 The Producing areas & environments of *Salvia* samples

丹参产地	样品编号	附注
山东省	D1—D12(共12批次)	野生及栽培的不同批次丹参
山西省	D13—D20(共8批次)	野生及栽培的不同批次丹参
河南省	D21—D33(共13批次)	野生及栽培的不同批次丹参
四川省	D34—D37(共4批次)	栽培的不同批次丹参
浙江省	D38—D41(共4批次)	栽培的不同批次丹参
河北省	D42—D44(共3批次)	栽培的不同批次丹参
云南省	D45—D46(共2批次)	野生的不同批次丹参
广西省	D47(共1批次)	产地不详
甘肃省	D48(共1批次)	野生不同种丹参
贵州省	D49(共1批次)	购买的丹参饮片
重庆市	D50(共1批次)	野生不同种丹参

1.2 样品处理

丹参固体样品粉碎后掺入KBr研磨后压片。

1.3 仪器设备和分析软件

FTIR 红外光谱仪为 Perkin Elmer 公司的 Spectrum GX 光谱仪, 中红外 DTGS 检测器, 光谱分辨率 4 cm⁻¹, 扫描范围 4 000~ 400 cm⁻¹, 扫描信号累加 16 次。

自行设计的多波段谱图比对软件——阵列相关系数比对软件。

1.4 阵列相关系数比对原理

阵列相关系数用于度量两条光谱曲线的线性相关程

收稿日期: 2006-05-30, 修订日期: 2006-10-12

基金项目: 科技部“十五”重大科技攻关, “重大技术标准研究”科技基础性工作专项“中药筛选红外光谱标准方法建立”基金项目(2002BA906A29-4)和清华大学 SRT 项目(052T 0240)资助

作者简介: 储德韧, 1986 年生, 清华大学化学系 2004 级本科生 * 通讯联系人 e-mail: sunsq@tsinghua.edu.cn

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

度^[12] (见图 1)。对于两条数据曲线, $x = (x_1, x_2, \dots, x_N)$, $y = (y_1, y_2, \dots, y_N)$, N 为数据维数; 其阵列相关系数定义为

$$R(X, Y) = \frac{\sum_{i=1}^p (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^p (x_i - \bar{x})^2 \right] \times \left[\sum_{i=1}^p (y_i - \bar{y})^2 \right]}}$$

式中 $\bar{x}$ 和 $\bar{y}$ 分别为 A、B 两样本吸光度的平均值。

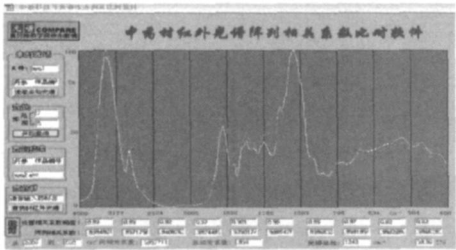


Fig 1 Array of correlation coefficient comparison software

1.5 丹参谱图分析及比对设置中阵列相关系数表中波数区间划分^[13 14]

根据丹参谱图(图 2)可以确定根据丹参的红外峰分布特点为, 在(1 630~ 1 636), (1 415~ 1 417), (1 264~ 1 266), (1 141~ 1 149), (1 050~ 1 060)和 799 cm^{-1} 处谱图都较为相似, 但相对峰强度有所不同, 较难直接用肉眼鉴定。由此可确定阵列相关系数表中波数区间划分基准: 在保证丹参的共有峰落在一个特定区间的前提下, 尽量使各区间长度均匀。

具体的划分波段和相关系数见实验结果部分。

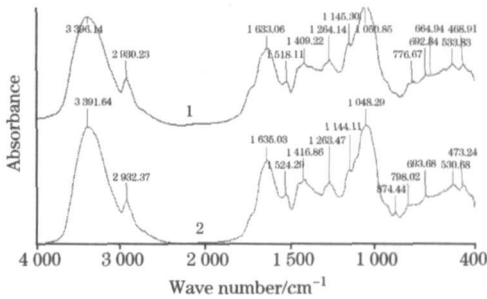


Fig 2 IR spectra of different kinds of Salviae sample

1: Wild samples from Yishui, Shangdong;
2: Cultivated samples from Lingbao, Henan

2 结果与讨论

2.1 相同产地的丹参的相关性分析

以山东沂水泉庄乡 李庄野生丹参(DS010101)为标准, 将 R1—R10 的阵列相关系数的阈值分别设为 0.800, 0.800, 0.770, 0.950, 0.930, 0.950, 0.990, 0.890, 0.760 和 0.850, 恰好可以找出所有 4 个样品的最高阈值(见表 2)。

研究对象, (1) 山东沂水泉庄乡 采挖野生丹参(DS010101); (2) 山东沂水泉庄乡 采挖栽培 1 年丹参(DS030101); (3) 山东沂水泉庄乡 采挖栽培 2 年丹参(DS040101); (4) 山东沂水购买野生丹参(DS050101)。

Table 2 Array of Correlation Coefficient for Salviae form Yishui, Shangdong

相关系数阈值	R1 0.800	R2 0.800	R3 0.770	R4 0.950	R5 0.930	R6 0.950	R7 0.990	R8 0.890	R9 0.760	R10 0.850	R 总/ R(2 000~ 500 cm^{-1})
DS010101	1	1	1	1	1	1	1	1	1	1	1/1
DS030101	0.998 3	0.985 4	0.916 1	0.997 5	0.966 5	0.975 8	0.995 9	0.995 9	0.847 6	0.858 5	0.967 7/0.964 5
DS040101	0.993 8	0.985 8	0.778 4	0.994 2	0.950 7	0.992 5	0.996 1	0.953 7	0.764 2	0.930 1	0.957 5/0.946 3
DS050101	0.993 5	0.996 3	0.984 8	0.988 7	0.969 1	0.996 1	0.995 8	0.979 4	0.880 2	0.975 1	0.988 9/0.974 4
波数区间	4 000~ 3 176	3 177~ 2 523	2 524~ 2 004	2 005~ 1 591	1 592~ 1 264	1 265~ 1 004	1 005~ 797	798~ 633	634~ 503	504~ 400	4 000~ 400

分析对比表 2 可以看出同一产地的真品丹参红外谱图表现出了极大的相似性:

(1) 重点考察 R4, R5, R6, R7, R8 五个阵列相关系数值都在 0.950 以上, 各组的 R(2 000~ 500 cm^{-1}) 值和 R 总都分别稳定在 0.967 和 0.946 以上。

(2) 除了以上重点考察的之外, 各实验组的其他相关系数的值不是十分稳定, 有的甚至低于 0.800, 这是由于指纹区的变动或批次不同导致的噪声, 弃去这些无关数据有利于对真品丹参聚类的分析。

对样品较丰富的其他产地采集的四组样品进行分组对比实验, 得到了相似的结果。

2.2 不同生长环境——栽培与野生的丹参的分析与鉴别

研究对象包括三组栽培与两组野生丹参: (1) 山东沂水购买栽培丹参(DS060101); (2) 山西平陆购买栽培丹参(DS200101); (3) 河南禹州购买栽培丹参(DS330101); (4) 云

南曲靖购买野生丹参(DS450101); (5) 云南璐西购买野生丹参(DS460101)。以山东沂水购买栽培丹参为标准, 分别将 R1—R10 的阵列相关系数的阈值设为 0.990, 0.990, 0.940, 0.990, 0.950, 0.970, 0.988, 0.987, 0.830 和 0.930, 结果检出了山西平陆(DS200101)与河南禹州(DS330101)两种栽培丹参, 符合检测要求。

该检测方法的关键是对于 R5 的把握, 在对其他野生样品组做的比对数据中, 设置 R5 值在 0.950, 发现栽培丹参的 R5 列相关系数都在 0.970 以上, 而野生丹参的该系数(以栽培丹参为参比)基本都在 0.950 以下。进一步验证所有 15 组野生丹参样品, 只发现唯一的例外: 即山东沂水购买野生丹参(DS050101) R5 值在 0.976, 超过了 0.950 的阈值, 检测精确度为 1/15 约 93.3%。出现例外的原因很大程度上是由于编号 DS050101 的野生丹参与参比标准同产于山东沂水, 产地的一致性掩盖了生长条件的差异性。同时对其他产地的

野生与栽培丹参进行分析, 设置与实验 2.2 同样的阈值基准可保证从栽培丹参组中剔除野生丹参。

2.3 不同产地的丹参的分析与鉴别

实验表明丹参谱图的确表现出了区域的相似性以及省区间品种微小差异性, 南北方的丹参的 R6, R7 及 R(2 000~500 cm⁻¹) 和 R 总都稳定在 0.964 以上。同时, R5 与 R8 的明显波动性为区别不同丹参产地提供了思路。

2.3.1 不同省区栽培丹参的鉴别举例

以浙江永康购买栽培丹参(DS410101) 为标准, 经过反复比对产自浙江本地(DS380101)、山西(DS200101)、山东(DS030101)、河北(DS030101) 四省的栽培丹参样品, 将 R1—R10 的阵列相关系数的阈值分别设为 0.990, 0.990, 0.990, 0.970, 0.900, 0.970, 0.970, 0.970, 0.960 和 0.920, 结果查找到了浙江磐安深泽乡采挖栽培丹参(DS380101), R5 和 R7 值高达 0.986 和 0.989, 而其他各组数据由于 R5 值低于 0.880, R7 值低于 0.970 而被剔除。

2.3.2 不同省区野生丹参之间的鉴别举例

以山东沂水泉庄乡李庄采挖野生丹参(DS010101) 为标准, 经过反复比对产自山东本地(DS070101)、河南(DS070101)、云南(DS450101) 三省的样品, 将 R1—R10 的阵列相关系数的阈值分别设为 0.990, 0.990, 0.990, 0.975, 0.975, 0.970, 0.970, 0.970, 0.960 和 0.920, 结果查找到了山东平邑羊城乡采挖野生丹参(DS070101), R5 和 R6 值分别高达 0.985 和 0.985, 而其他各组数据由于 R5 值低于 0.965, R6 值低于 0.970 而被剔除。

针对其他各省区组合共 6 组检出试验, 发现只要是需检出组与标准组地域相距较远, 如分别位于长江流域和黄河流

域, 检出率就高达 95% 以上, 可由此将各省区的丹参按相似程度加以分类鉴别, 具体的聚类关系见表 3。但美中不足的是采用上述方法对相邻省份特别是山东、山西组合的检出率偏低, 只有约 85%; 这就要求我们考虑调整阵列相关系数的阈值设置, 以及与其他检测方法联合使用等^[15], 以期在将来取得更好的效果。

Table 3 Clusters of Salviae from different provinces of China

北方丹参系统		南方丹参系统	
北方 1 组	北方 2 组	南方 1 组	南方 2 组
山东省	河南省	浙江省	云南省
山西省	河北省		四川省
甘肃省			重庆市

3 结 论

综上所述, 利用阵列相关系数比对法, 借助于丹参药材粉末红外指纹谱图数据库, 通过调整重点考察的 R4, R5, R6, R7, R8 五个阵列相关系数的数据阈值, 特别是对丹参特别敏感的 R5, R6, R7 阈值, 可以做到区别栽培与野生丹参, 以及鉴别南北方不同省份丹参品种。检测正确率达 93.3% 以上。同时通过对阵列相关系数阈值的设定, 可以得到不同的比对效果: 当阈值较高时, 可以找到相近产地的丹参; 当阈值较低时, 可以对野生及栽培丹参进行鉴别。根据相关性水平可以对药材进行分类。

参 考 文 献

- [1] MO Zhijiang(莫志江). China Journal of Chinese Materia Medica(中国中药杂志), 2007, 32(2): 116.
- [2] SONG Min, HANG Taijun, ZHANG Zhengxing(宋敏, 杭太俊, 张正行). Chinese Traditional and Herbal Drugs(中草药), 2005, 36(3): 360.
- [3] HUANG Jingqun, ZHANG Xirqiao, GUO Jun, et al(黄敬群, 张秀巧, 郭军, 等). Journal of Anhui Traditional Chinese Medical College(安徽中医学院学报), 2005, 24(1): 37.
- [4] CHEN Yong, ZHANG Ling, WANG Shimin(陈勇, 张玲, 王世敏). Chinese Journal of Analytical Chemistry(分析化学), 2004, 32(11): 1421.
- [5] LIU Shihua, ZHANG Xuegong, ZHOU Qun, et al(刘沐华, 张学工, 周群, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2005, 25(6): 878.
- [6] WANG Lirun, LI Yingming, ZHANG Liping, et al(王立群, 李英明, 张丽萍, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(6): 1061.
- [7] YU Lu, SUN Srqin, ZHOU Qun, et al(郁露, 孙素琴, 周群, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(12): 2181.
- [8] A YI Girli, ZHOU Qun, DONG Xiaou, et al(阿依古丽, 周群, 董晓鸥, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(7): 1237.
- [9] LI Yan, WU Rarran, YU Baihua, et al(李燕, 吴然然, 于佰华, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(10): 1846.
- [10] SUN Srqin, ZHOU Qun, QIN Zhu(孙素琴, 周群, 秦竹). Atlas of Two Dimensional Correlation Infrared Spectroscopy for Traditional Chinese and Medicine Identification(中药二维相关红外光谱鉴定图集). Beijing: Chemical Industry Press(北京: 化学工业出版社), 2003.
- [11] XU Yongqun, ZHU Jian, QIN Zhu, et al(徐永群, 诸健, 秦竹, 等). Computers and Applied Chemistry(计算机与应用化学), 2002, 19(3): 223.

- [12] XU Yong qun, SUN Si qin, XU Jiu wen(徐永群, 孙素琴, 许锦文). Chinese Journal of Spectroscopy Laboratory(光谱实验室), 2002, 19(5): 606.
- [13] XIE Jing-xi, CHANG Jun-biao, WANG Xu-ming(谢晶曦, 常俊标, 王绪明). The Application of Infrared Spectroscopy in Organic Chemistry and Medical Chemistry(红外光谱在有机化学和药物化学中的应用). Beijing: Science Press(北京: 科学出版社), 2001.
- [14] SUN Si qin, ZHOU Qun, YU Jia-r yuan, et al(孙素琴, 周群, 郁鉴源, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2000, 20(2): 199.
- [15] GUO Ping, YUAN Ya-li, XIONG Ping(郭萍, 袁亚莉, 熊平). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2004, 24(10): 1210.

Application of FTIR Array of Correlation Coefficient to the Identification of *Slviae Miltiorrhizae* Bge

CHU De ren, ZHOU Qun, YU Lu, SUN Si qin*

Key Laboratory of Bioorganic Phosphorus Chemistry and Chemical Biology(Ministry of Education), Department of Chemistry, Tsinghua University, Beijing 100084, China

Abstract Based on the fingerprint infrared spectrum database, array of correlation coefficient has been first applied to identify traditional Chinese medicine *slviae miltiorrhizae* Bge. from different producing areas and growing environments. The result showed that the setting of high correlation coefficient in particular ranges of spectrum could differentiate the producing area of *Slviae*, while the setting of low correlation coefficient threshold of R5 to R7 ranges could identify wild or cultivated samples. This approach seems to be not only a simple but also an accurate method for identifying the character of different *Slviae*.

Keywords Fourier transform infrared spectroscopy (FTIR); Array of correlation coefficient; *Slviae miltiorrhizae* Bge.; Recognition of cluster

(Received May 30, 2006; accepted Oct. 12, 2006)

* Corresponding author