

基于主成分分析和聚类分析的野生战骨红外指纹图谱比较研究

韦记青¹, 孔德鑫¹, 唐 辉¹, 黄庶识², 王满莲¹, 邹 蓉¹

(1. 广西壮族自治区 广西植物研究所, 广西 桂林 541006; 2. 广西科学院 生物物理实验室, 南宁 530003)

摘 要: 利用傅里叶变换红外光谱仪检测不同产地战骨的红外光谱, 并通过主成分分析模型, 聚类分析以及主成分载荷因子等化学计量学方法, 对不同产地样本进行鉴别, 找出不同产地样本化学组成的差异。结果表明: (1) 在主成分分析模型检测产地中广西平果县和扶绥县产地相距最近, 在三维图上的距离也最近, 广西天峨县与扶绥县和田阳县相距较远, 在散点图上的距离也较远; 云南和贵州地理位置处于广西境外, 在散点图上的距离则偏离广西各地; (2) 对主成分分析模型结果通过聚类分析验证表明二者的分析结果较一致, 与主成分三维散点图相比, 聚类分析图能在各距离水平上对研究对象进行准确的划归; (3) 通过提取 PCA 模型中 PC1 的载荷因子, 发现不同产地战骨样品化学组成的差异主要体现在柚皮素、香草酸、丁香酯素、多糖、苷类 5 种物质成分含量的不同。

关键词: 战骨; 主成分分析; 聚类分析; 红外指纹图谱; 化学成分

中图分类号: Q657.33 **文献标识码:** A **文章编号:** 1000-3142(2011)04-0559-05

Comparative study on FTIR fingerprint of wild *Premna fulva* based on principal component analysis and cluster analysis

WEI Ji-Qing¹, KONG De-Xin¹, TANG Hui¹, HUANG Shu-Shi²,
WANG Man-Lian¹, ZOU Rong¹

(1. Guangxi Institute of Botany, Guangxi Zhuang Autonomous Region and the Chinese Academy of Sciences, Guilin 541006, China; 2. Lab of Biophysics, Guangxi Academy of Sciences, Nanning 530003, China)

Abstract: In order to identify and in-depth excavate the differences of chemical compositions in *Premna fulva* from different regions, infrared spectra of multi-region sampled individuals were detected using the fourier transform infrared spectroscopy. After that, chemometrics methods including principal component analysis, clustering analysis and principal component loading factors were utilitted to pretreat original spectra data. The results showed that: (1) in 3 dimensional scatter maps, among the different samples, the distance between Pingguo and Fusui of Guangxi was nearest, so the natural consequence, the distances of the samples from these two places were nearest, while the real distance from Tiane to Fusui and Tianyang was further in Guangxi, the distances of the samples among these three places were also further in 3 dimensional scatter maps. In addition, geographical position of Yunnan and Guizhou were far from Guangxi, the results showed that scatter plot chart on distance were also deviated from Guangxi all around.

收稿日期: 2010-12-23 **修回日期:** 2011-04-07

基金项目: 广西自然科学基金(0991227); 广西科技攻关项目(0815005-2-3); 广西科技创新能力建设项目(0992028-10) [Supported by the Natural Science Foundation of Guangxi(0991227); Key Technologies Research and Development Program of Guangxi(0815005-2-3); Scientific and Technological Innovation Capacity-Building Foundation of Guangxi(0992028-10)]

作者简介: 韦记青(1968-), 女, 广西桂林人, 副研究员, 主要从事经济植物的引种驯化研究工作, (E-mail) weijq@gxib.cn.

(2) The principal component analysis model results was verified by clustering analysis and had similar consequences, while, as far as the two methods were concerned, they had individual characteristic, compared with the method of principal component analysis, clustering analysis chart could be accurately classified the research objects on each distance level, but it was difficult to directly reflect the interrelationship among research objects. (3) Through extracting the loading factor based on principal component analysis, the result indicated that the differences of quality of *P. fulva* among different regions mainly represented in the different contents of chemical components including naringenin, vanillic acid, syringaresiol, polysaccharide and glycosides.

Key words: *Premna fulva*; principal component analysis; cluster analysis; infrared fingerprint; chemical composition

战骨药材为马鞭草科植物黄毛豆腐柴(*Premna fulva*)的干燥茎,别名土霸王、穿云箭,是一种常用的壮族药,具有活血散淤、强筋健骨、祛风止痛之功效,主要分布在广西、云南和贵州等省区,是广西道地药材(韦记青等,2010)。战骨药材主要含有木栓酮、木栓醇、 β -谷甾醇、胡萝卜苷、柚皮素、香草酸、硬脂酸、卢-谷甾醇(II)、熊果酸(III)、卢-胡萝卜苷(IV)、丁香酯素(II)等多种化学成分(戴春燕等,2006,2007)。在抗脂质过氧化的药理模型中表明战骨药材中黄酮类化合物柚皮素、洋芹素有不同程度的抗氧化活性(黄琼等,2009)。此外,陈丽芬等(2008)对战骨药材总黄酮含量进行了检测,蒋才武等(2004)和黄祥远等(2007)分别研究了战骨药材的生药学特征。然而,目前战骨药材来源仍然以采集野生资源为主,而野生资源分布较零散,再加上受不同产地生态环境因子的影响,药材质量必然存在较大差异。因此,开展战骨种质资源研究,深入探讨不同产地药材资源化学成分的差异,筛选化学成分含量高的药材,对开展战骨人工引种驯化、良种选育及野生资源的有效保护和可持续利用均有重要意义。

红外光谱法(fourier transform infrared spectroscopy, FTIR)有整体特征性强、取样量小、简便迅速、准确等特点。随着科技的发展,将红外光谱的指纹特征与化学计量学相结合进行药材研究分析,已经被广泛地应用于各个研究领域(程存归等,2006;徐永群等,2002;张志祥等,2008)。这种方法可以挖掘样品潜在的化学成分信息,能够克服由于红外光谱固有的复杂性及中药所含化合物的多样性致使不同产地或不同栽培条件下中药材样本光谱的细微差别肉眼不能有效识别的困难。本文利用傅里叶变换红外光谱仪检测不同产地战骨的红外光谱,对不同产地样本进行鉴别,找出样本化学组成和产地差异的信息,为战骨植物资源的评价提供简单有效的方法。

1 材料与方法

1.1 材料来源及前处理

战骨样本均为野生,分别于2008年9月采自贵州、云南以及广西扶绥、平果、田阳和天峨县等,材料经广西植物研究所韦霄研究员鉴定为战骨(*P. fulva*)。取各产地战骨的茎,用毛刷去除泥沙,清洗干净晾干,在55℃下干燥48h至恒重,采用中药粉碎机粉碎后过200目筛,制成细粉储存于干燥器内备用。

1.2 光谱采集及处理

采用美国 Thermo Nicolet 公司生产的 Nicolet 5700 型傅里叶变换红外光谱仪,分辨率为 4 cm^{-1} ,扫描范围为 $4\ 000\sim 400\text{ cm}^{-1}$ 。实验过程保持在室内温度 $25\text{ }^{\circ}\text{C}$ 左右,空气相对湿度 30% 左右,每个片累积扫描次数 32 次,扫描时扣除 H_2O 和 CO_2 的干扰。精确称取各样品 1.0 mg 并分别与 200 mg 溴化钾(碎晶)混合研磨充分均匀,压制厚度约 1 mm 的透明锭片,然后放入红外光谱仪测定。每个产地制备 15 个样品片,每个样品片随机扫描 3 个不同的位置,获得 3 个光谱图,取其平均谱图作为最后的样品谱图,原始光谱数据先经过软件 OMNIC 8.0 软件进行基线校正,将预处理后的光谱数据导入软件 Unscrambler 9.7 中,一阶导数处理后,进行标准归一化(Standard normal variate, SNV)处理,以去除不同样本称量的差异(孔德鑫等,2011)。

2 结果与分析

2.1 不同产地战骨茎红外光谱带的提取及其比较分析

从光谱归一化处理后不同产地战骨茎典型的红外光谱特征可以看出,不同产地战骨茎的红外谱图较为相似(图 1)。按照吴瑾光(1994)和翁诗甫(2005)对谱带进行归属和分析: $3\ 395\text{ cm}^{-1}$ 附近为战骨中木栓醇、 β -谷甾醇等-OH 伸缩振动;此外,此

区间还有水分中羟基和氨基酸中-NH 的伸缩振动。2 922、1 460 和 1 375 cm^{-1} 附近为烷烃 CH_3 、 CH_2 的振动吸收,1 738 cm^{-1} 左右为酯类 $\text{C}=\text{O}$ 的振动吸收,1 653 cm^{-1} 附近为 $\text{C}=\text{C}$ 的伸缩振动,在 1 604 和 1 509 cm^{-1} 附近为芳环的振动吸收,这些与战骨茎中柚皮素、香草酸、丁香酯素等化合物的特征吸收相吻合;1 241 cm^{-1} 处为糖类、糖苷、脂类中 $\text{C}=\text{C}$ 骨架、 $\text{C}=\text{O}-\text{C}$ 及 $\text{P}=\text{O}$ 的伸缩振动峰,1 106、1 052、777 cm^{-1} 等振动峰归属为甾醇、多糖和苷类等化合物的振动吸收。比较不同产地战骨的平均光谱,发现不同产地其化学成分具有复杂性和多样性。澄清不同产地战骨化学成分差异,需要进一步研究。

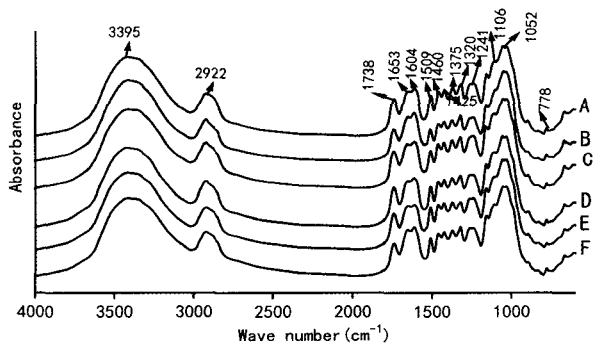


图 1 不同产地战骨茎的红外光谱图
Fig. 1 FTIR spectra of stems of *Premna fulva* from different regions
A. 广西扶绥; B. 贵州; C. 广西平果; D. 广西田阳;
E. 广西天峨; F. 云南。
A. Fusui of Guangxi; B. Guizhou; C. Pingguo of Guangxi;
D. Tianyang of Guangxi; E. Tian'e of Guangxi; F. Yunnan.

2.2 不同产地战骨茎光谱特征的主成分分析

为更加清晰地辨析不同产地战骨茎化学成分差异,选取波段 1 800~600 cm^{-1} 作为指纹特征光谱分析区域,建立主成分分析模型进行分析。图 2 是不同产地战骨茎的红外光谱经过一阶导数处理、SNV 标准归一化以及主成分分析后,主成分 PC1、PC2、PC3 得分的三维散点图。第一主成分的方差贡献率为 93.3%,是最重要的成分,前 3 个主成分的累积方差贡献率为 97.7%,表示前 3 个主成分可表达全部信息的 97.7%,只有 2.3%的信息丢失,所以,前三个主成分基本可以显示出各产地战骨化学成分之间的相似性关系。从排序图上可看出,不同产地战骨红外光谱在排序三维图上的位置关系与它们间的地理距离和相对地理位置有一定的对应性。如:在产地中,广西的平果县和扶绥县产地相距最近,在三

维图上的距离也最近,广西天峨县与扶绥县及田阳县相距较远,在散点图上的距离也较远;云南和贵州在地理上位于广西境外,在散点图上的距离也就偏离广西各产地,产地相近的产区战骨药材化学成分比较相似,而产地相距较远的地区战骨样本化学成分差异较大。

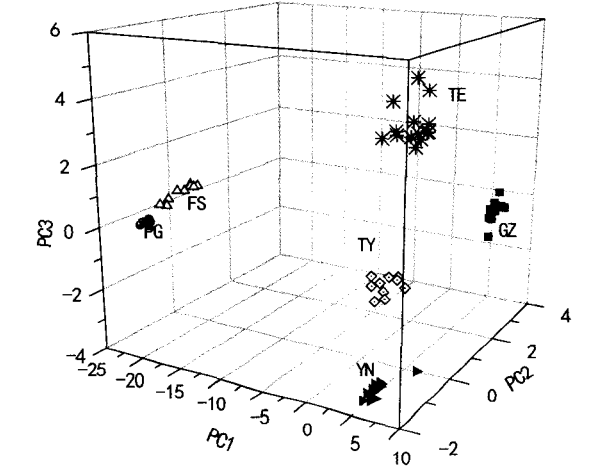


图 2 不同产地战骨样品的主成分得分散点图
Fig. 2 Scatter diagram of the main components in *P. fulva* samples from different regions

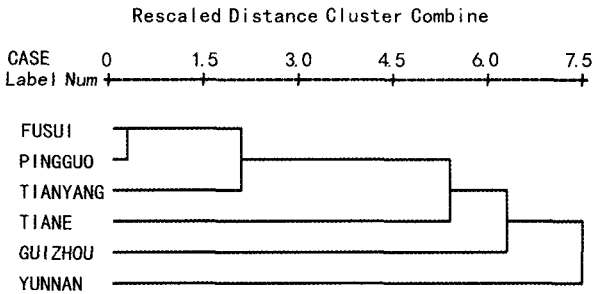


图 3 不同产地战骨样品化学成分的红外光谱聚类分析图
Fig. 3 Spectra diagram of the main components in *P. fulva* samples from different regions

2.3 不同产地战骨茎光谱特征聚类分析

取不同产地样本的平均图谱,基于欧式距离,运用 SPSS 11.5 对战骨茎的红外光谱进行分析,得到表征 6 个不同地区战骨化学成分的树系图(图 3)。6 个产地战骨之间欧式距离系数在 0~7.5 之间变动,当欧式距离系数在 2.5~5.5 间,各测试样品可划分为 4 个表征群,广西扶绥县和平果县居群间的距离系数最小,首先聚在一起,它们环境条件相似,化学成分的特征光谱较相似,然后他们再与广西田阳样本聚在一起;广西天峨、贵州、云南等产地分化

较大,分别归为一类,这种地理环境造成的数据差异与战骨茎的主成分分析结果一致。

2.4 不同产地战骨茎光谱特征的载荷分析

为从 PCA 模型中获取更多不同产地样本化学组成差异的信息,提取 PCA 模型的载荷因子,在波数 $1\ 800\sim 600\text{ cm}^{-1}$ 范围内,绘制载荷因子图谱(图 4)。因为在 PCA 的分析模型中,PC1 是主要成分,贡献率最大,达 93.3%,因此,只探讨 PC1 的载荷即可说明不同产地战骨化学成分的差异。战骨结构成分复杂,在整个波段范围内均有明显的吸收,某些不太灵敏的吸收峰被掩盖,未能在图 1 中表现出来,这些峰在图 4 中得以体现,如 $1\ 773$ 、 $1\ 623$ 、 $1\ 584$ 、 $1\ 419$ 、 $1\ 339$ 、 $1\ 283$ 、 $1\ 175\text{ cm}^{-1}$,这些峰依次对应为酰胺、芳香酸酯、烯烃、甾醇类等物质的特征吸收。此外,对 PC1 贡献较大的峰有 $1\ 735$ 、 $1\ 660$ 、 $1\ 584$ 、 $1\ 512$ 、 $1\ 497$ 、 $1\ 472$ 、 $1\ 320$ 、 $1\ 175$ 、 $1\ 014$ 、 975 cm^{-1} ,这些峰与战骨茎中柚皮素、香草酸、丁香酯素等化合物的特征吸收有密切的关系;由于 PC1 贡献率为 93.3%,对 PCA 模型的贡献最大,因而,不同产地战骨样品化学组成的差异主要体现在柚皮素、香草酸、丁香酯素、多糖、苷类 5 种物质成分含量的不同。

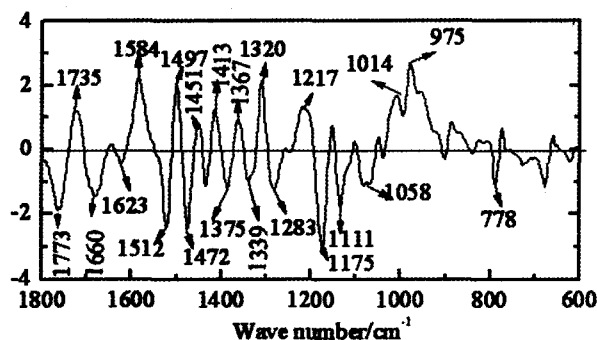


图 4 PCA 鉴别模型的载荷因子

Fig. 4 The loading factors of PCA identification model

3 结论与讨论

药用植物的红外光谱特征是其化学成分中含何种官能团的叠加,不同样品在植化组成上的差异程度可通过红外光谱特征体现出来。因此,对药用植物化学成分有影响的外界因素,如:地域、加工方法、储存等可以从各自的红外光谱上反映出其差异性。将红外光谱与现代化学计量学,如主成分分析、聚类分析、偏最小二乘法等方法相结合,可以挖掘药材原

始光谱中潜在的化学成分差异信息。左康等(2007)采用主成分分析法对 7 个产地的虎杖植化组成上的相似性进行聚类分析,表明虎杖样品在 $4\ 000\sim 700\text{ cm}^{-1}$ 区域有相似的红外光谱特征,而在 $1\ 700\sim 900\text{ cm}^{-1}$ 区域不同产地的样品又存在一定的差异。徐永群等(2002)用红外指纹图谱和主成分分析法,对来自 8 个省的 74 个山药样本道地与非道地性进行了聚类分析。张志祥等(2008)应用 FTIR 法测定了浙江省 7 个不同地区香果树茎和叶的红外光谱,并应用主成分分析和聚类分析,比较了 7 个不同地理居群香果树茎和叶在红外光谱图上的差异程度,进而分析其不同产地化学成分含量差异。

在这些理论基础上,本文利用傅里叶变换红外光谱技术对不同产地战骨茎的化学成分进行鉴别、指认,发现不同产地战骨的红外光谱在三维图上的位置关系与相对地理位置有一定对应性,即:广西平果县和广西扶绥产地相距最近,在三维图上的距离也最近,广西天峨县与扶绥县和田阳县相距较远,在散点图上的距离也较远;云南和贵州地理位置处于广西境外,在散点图上的距离也就偏离广西各地;云南、河池与平果及扶绥产地相距较远,在三维散点图上的距离也较大,表明了产地相距较近的产区战骨药材化学成分比较相近,产地相距较远的产区战骨样本化学成分差异较大,这种由环境条件对药用植物生长及其化学成分累积影响的结果与左康等(2007)对不同产地虎杖研究结果相似。

与主成分分析模型相比,聚类分析虽然不如主成分三维图谱那样能清楚表达样本间的相互关系,但聚类分析图能在各距离水平上对研究对象进行准确划归。如:在三维散点排序图中,贵州与天峨和田阳的距离相差不明显,较难鉴别贵州样品与广西地区天峨县和田阳县样品间差别时,聚类分析则体现出优越性,可以表达出三者的关系。因此,将主成分分析与聚类分析相结合,可准确判断出不同产地战骨由于地理距离而引起的分化。

通过提取 PCA 模型中 PC1 的载荷因子,在波数 $1\ 800\sim 600\text{ cm}^{-1}$ 范围内,绘制载荷因子图谱。发现不同产地战骨样品化学组成的差异主要体现在柚皮素、香草酸、丁香酯素、多糖、苷类 5 种物质成分含量的不同。在此基础上,通过比较原始图谱中 $1\ 660\text{ cm}^{-1}$ 和 $1\ 512\text{ cm}^{-1}$ 等处与黄酮类物质相关的特征吸收峰峰强发现,各产地样品在上述两处的峰强在 $0.413\sim 0.591$ 间变化,广西地区以广西天峨产区样

品的峰强最高,与云南省的样品相当,其次是广西田阳和平果县,最低的是广西扶绥县。此结果表明,不同产地战骨黄酮含量以广西天峨县最高,广西扶绥的最低,这与韦记青等(2010)检测不同产地总黄酮含量结果相一致。但是,此结论是借助黄酮类物质官能团的振动特征得出的,有待结合柚皮素等黄酮类物质标准品进行进一步分析考证。

本文运用红外光谱和主成分分析模型、聚类分析及主成分分析的载荷因子,可以实现对不同产地野生战骨药材进行评价,并找出了造成这些差异的主要成分,研究结果可对不同产地战骨药材质量评价提供参考,对其化学成分的研究有一定指导意义。

参考文献:

- 吴瑾光. 1994. 近代傅里叶变换红外光谱技术及应用(上卷)[M]. 北京:科技文献出版社:573—625
- 翁诗甫. 2005. 傅里叶变换红外光谱仪[M]. 北京:化学工业出版社:239—262
- Chen LF(陈丽芬), Huang Q(黄琼), Lin CW(林翠梧), et al. 2008. Determination of total flavonoids content in *Premna fulva* Craib by ultraviolet spectrophotometry(紫外分光光度法测定战骨中总黄酮的含量)[J]. *Guangxi J Light Industry*(广西轻工业), (1):27
- Chen CG(程存归), Li DT(李丹婷), Chen JH(陈建华), et al. 2006. HATR-FTIR spectroscopy combined with PCA: application to identification of *Radix Paeoniae Rubra* and its confusable varieties (HATR-FTIR 结合主成分分析应用于赤芍的真伪鉴别)[J]. *Chin Pharm J*(中国药理学杂志), 41(8):581—582
- Dai CY(戴春燕), Chen GY(陈光英), Zhu GY(朱国元), et al. 2007. Study on chemical component of the stem of *Premna fulva*(战骨茎的化学成分研究)[J]. *Chin Trad Herb Drug*(中草药), 38(1):34
- Dai CY(戴春燕), Chen GY(陈光英), Zhu GY(朱国元), et al. 2006. Study on the chemical constituents from petroleum ether fraction of the stem of *Premna fulva* craib(战骨茎石油醚部分的化学成分研究)[J]. *J Hainan Nor Univ: Nat Sci Edit*(海南师范学院学报·自然科学版), 19(1):50
- Huang XY(黄祥远), Zhang KF(张可锋). 2007. Studies on pharmacognosy of *Premna fulva* Craib(黄毛豆腐柴的生药学研究)[J]. *Chin J Eth*(中国民族民间医药), 88(5):275
- Huang Q(黄琼), Chen LF(陈丽芬), Lin CW(林翠梧), et al. 2009. Flavonoids in *Premna fulva* Craib and their anti-oxidative activities *in vitro*(药用植物战骨中黄酮类化学成分及抗氧化作用研究)[J]. *Lishizhen Med Mat Med Res*(时珍国医国药), 20(11):2 076—2 078
- Jiang CW(蒋才武), Zeng CH(曾春晖), Liu SY(刘寿养). 2004. The description and microscopy identification of *Premna fulva* Craib(壮族药战骨的性状与显微鉴别)[J]. *Chin Med Mat*(中药材), 7(11):811
- Kong DX(孔德鑫), Huang XY(黄夕洋), Li F(李锋), et al. 2011. Analysis and evaluation taxol contents in different parts between *Taxus media* and *Taxus mairei* based on fourier transform infrared spectroscopy(基于 FTIR 的曼地亚和南方红豆杉不同部位紫杉醇含量分析与评价)[J]. *Spec and Spec Ana*(光谱学与光谱分析), 31(3):657
- Wei JQ(韦记青), Shi YC(史艳财), Jiang YS(蒋运生), et al. 2010. Investigation on wild germplasm resources of *Premna fulva* Guangxi(广西战骨种质资源调查研究)[J]. *Lishizhen Med Mat Med Res*(时珍国医国药), 21(8):2 014—2 044
- Xu YQ(徐永群), Sun SQ(孙素琴), Yuan ZM(袁子民), et al. 2002. Discrimination of *bvtrueborn* tuber *dioscoreae* by fingerprint infrared spectra and principal component analysis(红外光谱结合主成分分析鉴别道地山药)[J]. *Chin J Anal Chem*(分析化学), 30(10):1 231—1 233
- Zhang ZX(张志祥), Liu P(刘鹏), Kong HJ(康华靖), et al. 2008. A study of the diversity of different geographical populations of *Emmenopterys henryi* using FTIR based on principal component analysis and cluster analysis(基于主成分分析和聚类分析的 FTIR 不同地理居群香果树多样性分化研究)[J]. *Spec Spec Ana*(光谱学与光谱分析), 28(9):2 081—2 086
- Zuo K(左康), Li DT(李丹婷), Guo SL(郭水良), et al. 2007. FTIR spectra principal component analysis of roots of *Polygonum cuspidatum* from different areas(不同产地虎杖根 FTIR 主成分分析研究)[J]. *Spec and Spec Ana*(光谱学与光谱分析) 27(10):1 989—1 992