

北柴胡气相指纹图谱数据的化学计量学分析

王启帅, 肖功胜, 杨云*, 郝鹏飞(河南中医学院药学院, 郑州 450008)

摘要: 目的 对北柴胡挥发性成分的气相指纹图谱进行化学计量学研究, 提取出指纹图谱中的特征数据信息。方法 运用 Agilent 7890A 气相色谱仪, 安捷伦 HP-5 毛细管气相色谱柱(30 m×0.25 mm, 0.25 μm), 载气为氮气, 流速为 1.03 mL·min⁻¹, FID 检测温度为 280 ℃, 对不同产地北柴胡样品进行分析。运用指纹图谱处理软件对指纹图谱数据进行相似度分析、聚类分析及主成分分析。结果 建立了北柴胡挥发性成分的气相指纹图谱共有模式, 获得了 28 个共有峰。依据化学计量学分析, 将样品分成了 3 类。结论 结果表明化学计量学分析方法适用于指纹图谱数据的分析, 可为定量研究提供依据, 有助于柴胡药材的质量评价研究。

关键词: 北柴胡; 气相指纹图谱; 化学计量学; 主成分分析

中图分类号: R284.1; R917.101

文献标志码: B

文章编号: 1007-7693(2010)11-1097-06

Chemometric Analysis on the GC Fingerprint Data of the Root of *Bupleurum chinense* DC

WANG Qishuai, XIAO Gongsheng, YANG Yun*, HAO Pengfei(College of Pharmacy, Henan University of Traditional Chinese Medicine, Zhengzhou 450008, China)

ABSTRACT: OBJECTIVE To study the gas chromatographic fingerprint of the volatile constituents from the root of *Bupleurum chinense* DC. and obtain the characteristic data based on chemometric analysis. **METHODS** Samples of *B. chinense* DC. from different habitats were determined by Agilent 7890A GC system with HP-5 capillary column (30 m×0.25 mm, 0.25 μm). Nitrogen was used as carrier gas and flow rate was 1.03 mL·min⁻¹, FID temperature was 280 ℃. The methods of similarity, cluster analysis and principal component analysis were used to process the data from which was obtained gas chromatographic fingerprint. The assisted computational tool was "Chromafinger" solution software. **RESULTS** The common pattern of GC fingerprint was established. Twenty-eight co-possessing peaks were selected as the fingerprint peaks of *B. chinense* DC.. Samples were divided into three categories based on the analysis of chemometrics. **CONCLUSION** The results show that chemometrics is suitable for the analysis of fingerprint data, which can provide a basis for quantitative research. Chemometric analysis contributes to research the quality evaluation for *B. chinense* DC..

KEY WORDS: *Bupleurum chinense* DC.; GC fingerprint; chemometrics; principal component analysis

由于中药成分非常复杂,即使是单味中药的药效也是多种活性成分之间的协同作用。这就要求在中药的研究方法中,不能仅仅以单个成分的定性或定量标准来衡量中药的质量,还应从整体出发对中药的质量进行阐述。色谱指纹图谱着眼于宏观的规律性的特征分析,能够全面地、系统地反映中药的整体特征,因此在中药质量控制的应用中显现了强大的优越性。WHO 草药评价指南、英国草药典、美国 FDA 以及德国药用植物学会均将指纹图谱列入其中,要求

草药制剂需提供制剂产品的指纹图谱,以保证草药品种的真实性及产品的稳定性。但同时,由于指纹图谱本身所携带的信息量过于庞大,不仅包含了大量的有效信息,诸如指纹图谱所反映的成分特征及组分相对含量等,还包含了诸如色谱基线漂移等随机干扰信息。因此,如何将色谱指纹图谱中所携带的信息化繁为简,并从中提炼出能反映药材内在质量规律的特征信息,已成为近年来的热点话题。化学计量学作为一门新兴的化学分支学科,它运用数学、统计学、计算机科学

基金项目: 国家十一科技支撑计划资助项目(2006BAI06A10)

作者简介: 王启帅,男,硕士生 Tel: 13838550076
(0371)65680605 E-mail: yyun@china.com.cn

E-mail: iliveinhainan@126.com

*通信作者: 杨云,女,教授,硕士

Tel:

以及其他相关学科的理论与方法,优化化学量测过程,并从化学量测数据中最大限度地提取有用的化学信息,为阐明物质的成分、结构及其性能之间的复杂关系提供解决方案。化学计量学的最大特征是将多变量分析方法引入化学研究^[1]。它根据采集到的光谱或色谱数据或某些经量化后的指标,运用计算机对这些数据进行处理,从中获取出光谱或色谱中的有用信息,然后以量化数据代替直观鉴别对药材进行分析、判断、分类及鉴别真伪^[2]。现已被广泛应用于中药材质量评价的方法主要有贝叶斯(Bayes)判别法、聚类分析法(CA)、主成分分析(PCA)、人工神经网络技术(ANN)等。

柴胡(Radix Bupleuri)为伞形科植物柴胡(*Bupluerum chinense* DC.)和狭叶柴胡(*Bupluerum scorzonerifolium* Willd.)的干燥根,前者习称北柴胡,后者习称南柴胡。柴胡具有和解表里,疏肝,升阳之功效^[3],临床上广泛用于治疗各种原因引起的外感发热,其中柴胡挥发油被认为是其解热的物质基础。本试验利用气相色谱法(GC)建立北柴胡挥发性成分的色谱指纹图谱,运用化学计量学分析方法探讨北柴胡 GC 指纹图谱的特征及规律,以期通过化学计量学的数据处理方法对图谱进行“过滤”和简化,获取专属、稳定、实用的指纹图谱特征数据。

1 仪器与试剂

1.1 仪器与试剂

Aglient 7890A GC 色谱仪配备 Agilent Chemstation 色谱工作站(美国安捷伦科技有限公司);BS224S 型电子天平(北京塞多利斯公司);HP-5 毛细管气相色谱柱(美国安捷伦科技有限公司);氢火焰离子化检测器(美国安捷伦科技有限公司);GHL-300 高纯氢气发生器(北京汇佳精密仪器公司)。

正己烷为分析纯(天津四友科技有限公司);氮气(苏净集团保护气氛有限公司)。

1.2 药材

北柴胡样品共 12 批,分别采自于全国各地,见表 1。经河南中医学院生药教研室董诚明教授鉴定为伞形科植物柴胡 *Bupleurum chinense* DC.。按 2005 版中国药典柴胡项下规定,除去茎、叶及泥沙,40℃烘箱干燥 24 h,粉碎,过 10 目筛,干燥器内室温保存。

表 1 柴胡样品来源及编号

Tab 1 Sample source and number

样品编号	产地	生长方式	生长年限
S1	陕西渭南	野生	多年
S2	四川荣县	栽培	1 年
S3	河北丰宁	野生	多年
S4	甘肃陇西	栽培	2 年
S5	河北邯郸	栽培	1 年
S6	河南辉县	栽培	2 年
S7	河南嵩县	栽培	2 年
S8	辽宁大连	栽培	2 年
S9	河南栾川	野生	多年
S10	山西万荣	栽培	1 年
S11	河南嵩县	栽培	8 个月
S12	河南嵩县	栽培	9 个月

2 实验方法

2.1 供试品溶液的制备

取柴胡粉末 50 g,精密称定,置于 1 000 mL 圆底烧瓶中,加入 20% NaCl 溶液 500 mL,回流 1 h 后以正己烷代替二甲苯作为挥发油溶剂,按中国药典 2005 年版一部附录挥发油测定法(乙法)提取挥发油 6 h。吸取正己烷层,置于 2 mL 量瓶中,以正己烷定容至刻度,摇匀,即得。

2.2 气相色谱条件

HP-5 毛细管气相色谱柱(30 m×0.25 mm,0.25 μm);柱温升温程序:40℃保持 3 min,以 5℃·min⁻¹升至 110℃保持 2 min,再以 4℃·min⁻¹升至 250℃保持 3 min;进样量:1.0 μL,分流比:5:1;进样口温度 260℃;载气:高纯度氮气;流速:1.03 mL·min⁻¹,恒流模式;检测器:氢火焰离子化检测器(FID);检测温度:280℃;空气:400 mL·min⁻¹,氢气:30 mL·min⁻¹,氮气:25 mL·min⁻¹。

2.3 方法学考察

2.3.1 仪器精密度试验 按“2.2”项下色谱条件对同一供试品溶液重复进样测定 5 次,记录色谱图。以 p3 为内参照峰,计算各共有峰的相对保留时间和相对峰面积,结果显示:各共有峰的相对保留时间 RSD<3%,相对峰面积 RSD<5%。认为该仪器精密度良好。

2.3.2 稳定性试验 取样品编号为 S12 的柴胡粉末适量,按“2.1”项下方法制备,并在制备后 0, 6, 10, 14, 18, 24 h 进样测定,记录色谱图。以 p3 为内参照峰,计算各共有峰的相对保留时间和相对峰面积,结果显示:各共有峰的相对保留时间 RSD<3%,相对峰面积 RSD<5%。认为供试品溶液

在 24 h 内稳定。

2.3.3 重复性试验 取编号为 S12 的样品 5 份，按“2.1”项下方法平行制备，按“2.2”项下色谱条件进样测定，记录色谱图。以 p3 为内参照峰，计算各共有峰的相对保留时间和相对峰面积，结果显示：各共有峰的相对保留时间 RSD<3%，相对峰面积 RSD<5%。认为该方法重复性良好。

3 结果与分析

3.1 指纹图谱共有模式的建立

利用 CHROMAP 色谱指纹图谱系统解决方案软件(1.51 版，珠海科曼中药研究有限公司)，以样品 S12 的指纹图谱为保留时间参考标准，选取峰面积较大且较稳定，分离度好的 3 号峰(p3)作为内参照峰。对 12 批北柴胡样品的色谱指纹峰进行匹配及校正，获得 28 个共有峰，峰值算法以中位数

法计算，建立了北柴胡挥发性成分的指纹图谱共有模式(CP)。轮廓图见图 1A，共有峰的保留时间及相对峰面积见表 4。运用高斯曲线对共有模式进行色谱漂移校正，获得了共有模式的高斯曲线图，见图 1B。高斯曲线图更好地反映了色谱指纹图谱的共有信息，消除了基线漂移的干扰，简化了色谱图。

3.2 样品测定

取 12 批柴胡样品粉末，按“2.1”项下方法制备供试液，并在“2.2”项下色谱条件进样测定，记录色谱图。柴胡样品气相指纹图谱叠加图如图 2 所示。

3.3 GC 指纹图谱相似度计算

以共有模式为参照，采用夹角余弦法对 12 批北柴胡 GC 指纹图谱进行相似度计算，结果见表 2。

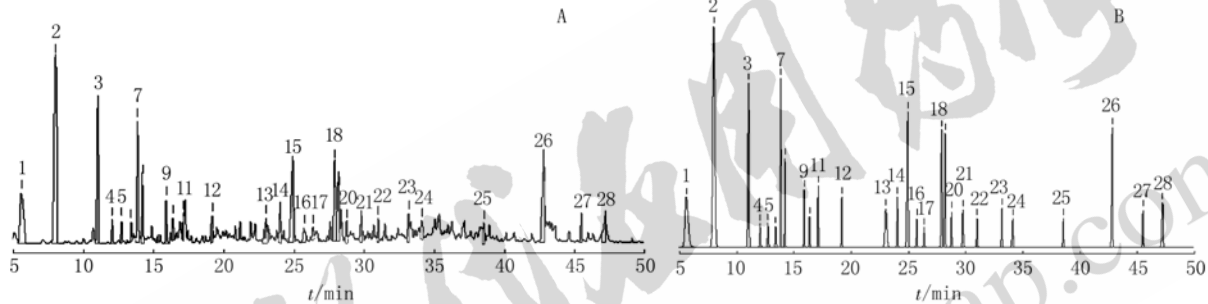


图 1 北柴胡 GC 指纹图谱共有模式
A-共有模式轮廓图；B-共有模式高斯曲线图
Fig 1 The common pattern of GC fingerprint
A-the profile curve of common pattern; B-the Gaussian curve of common pattern

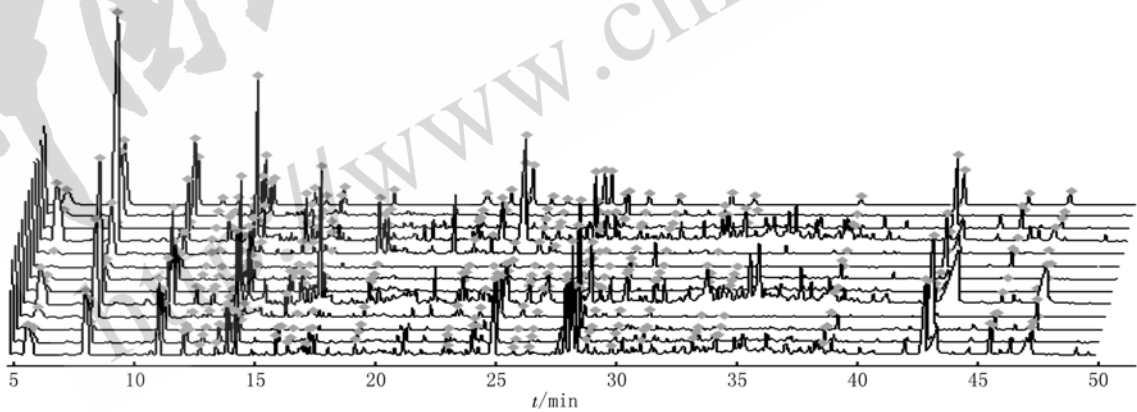


图 2 不同北柴胡样品 GC 指纹图谱叠加图(n=12, 由内至外依次为 CP, S1, S2, S3, ..., S12)
Fig 2 GC fingerprints of different samples (n=12, From interior to exterior, there are CP, S1, S2, S3, ..., S12)

表 2 相似度评价结果

Tab 2 The results of similarity of *B. chinense* DC. among different habitats

样品编号	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
相似度	0.963	0.375	0.910	0.864	0.454	0.908	0.956	0.908	0.954	0.327	0.806	0.821
空间夹角/°	16	68	24	30	63	25	17	25	17	71	36	35

相似度结果表明：编号为 S1, S3, S6, S7, S8, S9 的样品相似度最高。表明这 6 个北柴胡样品中的挥发性成分在组分构成方面具有较好的相似性，相对含量较其他样品更加稳定。值

得注意的是，样品 S2, S5, S10 的相似度极低(<0.500)，表明这 3 个柴胡样品中所含的挥发性成分与共有模式存在较大的差异。共有模式与样品 S10 的指纹图谱镜像对比图见图 3。

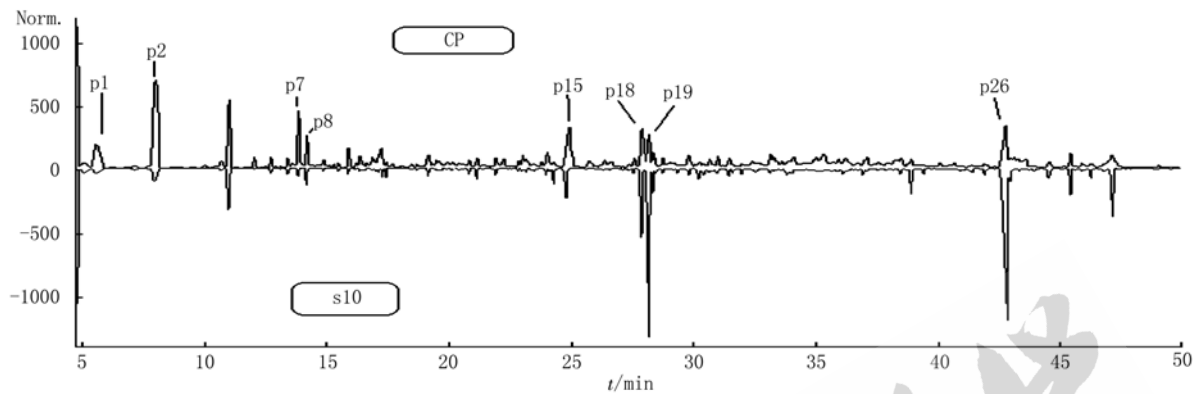


图 3 共有模式(上半部分)与样品 S10(下半部分)的指纹图谱镜像对比图
Fig 3 Comparison map between common pattern (superior) and the fingerprint of S10 (inferior).

从图 3 可看出，共有模式与样品 S10 的差异主要存在于色谱峰 p1, p2, p7, p8, p15, p18, p19, p26 之间。这 8 个色谱峰相对峰面积比例的失衡是致使其相似度较低的主要原因。

3.4 北柴胡 GC 指纹图谱的聚类分析

本试验对 12 批柴胡样品的 GC 指纹图谱及其共有模式(编号 0)进行聚类分析。其中数据源为峰面积，所有的数据经标准化变换预处理(用于消除变量的变异相差过大或变量缺失值过多对聚类结果的影响)，聚类方法为最邻近法，样品距离度量采用欧几里德(Euclidean)距离，即样品几何空间的分布。得到样品的聚类分析树状图，见图 4。

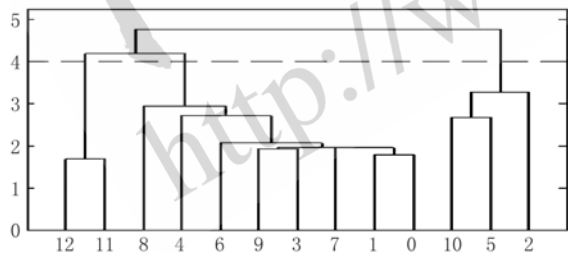


图 4 样品聚类分析树状图
Fig 4 Dendrogram for samples

聚类分析图反应的是各样品的相似程度，样品间的分类距离值越小则样品间的差异越小，反之则越大。当分类距离取 4 时，12 批柴胡样品被分为 3 类，见表 3。

表 3 样品聚类分析及主成分分类
Tab 3 The classifications of clustering analysis and principal component analysis for samples

样品聚类分析		主成分分类	
类别	样品编号	类别	样品编号
第 I ^a 类	S1, S3, S4, S6, S7, S8, S9	第 I ^b 类	S1, S3, S4, S6, S7, S8, S9
第 II ^a 类	S11, S12	第 II ^b 类	S11, S12
第 III ^a 类	S2, S5, S10	第 III ^b 类	S2, S5, S10

3.5 主成分分析

利用 CHROMAP 色谱指纹图谱系统解决方案软件，对 12 批北柴胡样品的 GC 指纹图谱进行主成分分析，以峰面积为数据源，所有的数据经标准化变换处理，主成分(PC)个数为 10。得到主成分分析结果见表 4。

从表 4 中可以看出：用 10 个 PC 最后得到的累积均方差(accumulate mean square)为 99.70%，表示引入 10 个主成分后可以概括平均 99.70%的原变量信息。第 1 主成分(PC1)和 PC1 + PC2 的累积均方差分别为 61.21%，74.39%。通过比较表 4 中因子得分系数矩阵的绝对值可看出：p2, p26 号峰对 PC1 的贡献最大；对 PC2 有贡献的色谱峰主要有 p1, p15, p18, p19。由于前 2 个主成分已经可以概括大部分原指标的信息(累积均方差>70%)，将 12 批样品(数据源为共有峰峰面积)在空间内的坐标投影于主成分 PC1~PC2 所确定的平面内，得到了样品在平面 PC1-PC2 内的分布散点图，见图 5。

表 4 北柴胡 GC 指纹图谱共有峰及主成分分析结果

Tab 4 The co-possessing peaks of GC fingerprint and the result of principal component analysis

峰号	共有模式		主成分分析				
	保留时间/min	相对峰面积	PC1	PC2	PC3		PC10
P1	5.59	0.412 5	-0.153 6	0.400 2	-0.0063	...	0.237 0
P2	7.99	1.622 7	-0.570 6	0.060 0	0.231 9	...	0.047 7
P3 ^R	11.05	1.000 0	-0.223 0	0.226 5	0.608 4	...	-0.042 5
P4	12.05	0.057 3	-0.011 3	0.104 0	-0.049 2	...	-0.253 6
P5	12.72	0.077 8	-0.019 7	0.033 3	-0.012 0	...	0.063 8
P6	13.41	0.070 2	-0.003 0	0.040 5	-0.004 8	...	-0.063 8
P7	13.86	0.420 1	-0.172 6	0.078 2	0.159 1	...	-0.007 3
P8	14.21	0.299 9	-0.102 7	0.123 8	-0.176 5	...	0.246 8
P9	15.89	0.154 8	-0.048 7	0.033 2	-0.013 4	...	0.148 2
P10	16.36	0.076 3	-0.021 9	0.019 9	-0.017 8	...	0.077 0
P11	17.12	0.114 7	-0.050 3	0.114 8	0.049 8	...	-0.167 6
P12	19.18	0.132 7	-0.072 2	-0.135 5	0.136 3	...	-0.048 5
P13	23.02	0.248 1	-0.074 7	0.023 6	-0.028 6	...	0.405 9
P14	24.04	0.165 7	0.017 5	0.066 8	0.169 8	...	-0.224 8
P15	24.96	0.513 4	0.114 0	0.387 4	0.441 0	...	-0.147 5
P16	25.74	0.123 0	-0.039 7	-0.029 3	0.001 5	...	-0.134 1
P17	26.37	0.065 5	-0.008 3	0.007 1	-0.026 0	...	-0.043 5
P18	27.91	0.298 9	0.071 0	0.429 4	0.152 0	...	0.212 1
P19	28.21	0.116 2	0.160 1	0.438 1	-0.019 0	...	-0.053 2
P20	28.77	0.110 1	-0.019 2	0.005 2	-0.007 1	...	-0.341 8
P21	29.77	0.115 6	-0.082 1	0.091 0	-0.365 7	...	-0.341 9
P22	31.02	0.090 2	-0.018 4	0.085 7	-0.065 9	...	0.194 9
P23	33.17	0.115 6	-0.014 0	0.168 8	0.032 4	...	0.102 1
P24	34.13	0.087 7	-0.024 5	0.092 8	-0.071 1	...	0.154 6
P25	38.55	0.080 1	-0.081 7	0.229 8	-0.226 2	...	-0.149 6
P26	42.83	0.240 5	0.688 8	0.076 2	-0.149 0	...	0.125 5
P27	45.51	0.103 5	0.093 6	0.019 4	-0.020 3	...	-0.065 3
P28	47.22	0.106 2	0.019 9	0.269 9	0.090 8	...	-0.307 9
累积均方差	-	-	61.21%	74.39%	83.25%	...	99.70%

注：主成分分析结果为因子得分系数矩阵以及累积均方差

Note: The result of PCA is consisted of component score coefficient matrix and the accumulate mean square

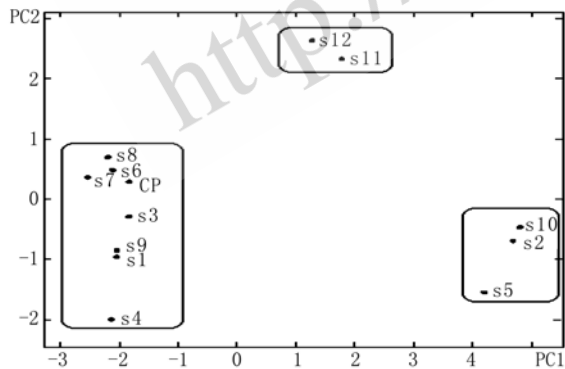


图 5 样品在平面 PC1-PC2 内的分布散点图

Fig 5 Scatter plot of PCA on two principal components for 12 batches of samples

从图 5 可以看出：样品 S3 与 S10 的差异主要体现在 PC1 轴上，而 S4 与 S12 的差别主要体现在 PC2 轴上。图 5 能明显的将 12 批北柴胡样品分为 3 类，分类结果见表 3。结合图 5、表 3 和表 4 可得出：色谱峰 p1, p2, p15, p18, p19, p26 的差异体现了 12 批北柴胡样品的主要差异(累积均方差为 74.39%)。

4 讨论

本试验过程中共选择了 4 个升温程序：
(A)50 ℃保持 3 min 以 4 ℃·min⁻¹ 升至 270 ℃后保持 5 min；
(B)45 ℃保持 5 min 以 3 ℃·min⁻¹ 升至 170 ℃保持 0 min 再以 5 ℃·min⁻¹ 升至 250 ℃保持

5 min;(C)40℃保持5 min以5℃·min⁻¹升至110℃保持0 min再以4℃·min⁻¹升至250℃保持5 min;(D)40℃保持3 min以5℃·min⁻¹升至110℃保持2 min再以4℃·min⁻¹升至250℃保持3 min。以上升温程序中,以(D)升温程序的色谱图信息量最为丰富,色谱峰分布均匀,特征性强。因此选D升温程序为本次试验的分析条件。

本试验以保留时间在4.80~50.00 min区间内的色谱峰为分析对象。主要是因为通过阴性溶液对照,笔者发现保留时间4.78 min为溶剂的色谱信号峰,而且保留时间50.00 min以后已无色谱峰。鉴于此,只对4.80~50.00 min时间区间内的谱图进行分析。

本试验采用夹角余弦法对气相指纹图谱进行相似度分析。通过指纹图谱的相似度计算结果,可得出:①不同产地的北柴胡样品所含的挥发性成分相似;②不同生长周期的北柴胡样品所含的挥发性成分存在差异;③如果以北柴胡中的挥发性成分为衡量依据,则生长周期为2年的人工栽培北柴胡样品与多年生的野生北柴胡样品之间具有较好的相似性。提示:对于目前现有的人工栽培柴胡的地区而言,生长周期是影响药材成分均一性及稳定性的主要因素之一。因此,笔者建议在柴胡药材的种植过程中应总结药材在不同生长周期内化学成分的动态积累规律,从中找出最适宜的种植年限及最适采收期。另外,当聚类分类距离取4时,可将指纹图谱相似度较高(相似度

≥0.864)的样品聚在一类群内。表明聚类分析在用于指纹图谱相关数据的处理时,亦具有较好的适用性,可作为指纹图谱数据分析的补充。

由于中草药中的活性成分大多属于植物的二次代谢产物,考虑到植物体内同一结构类型的化学物质的生物合成途径是相互联系的,使得植物体内同一结构类型的化学物质之间也存在着一定的关联性,所以笔者将主成分分析方法应用于北柴胡GC指纹图谱的研究中。本次试验笔者以样品指纹图谱中的共有峰为分析对象进行主成分分析,从中找出北柴胡挥发性成分指纹图谱色谱峰的差异可由色谱峰p1, p2, p15, p18, p19, p26的差异体现出来(累积均方差为74.39%)。如果只考虑色谱峰峰面积对相似度的影响,可以认为色谱峰p1, p2, p15, p18, p19, p26是影响北柴胡样品指纹图谱相似度评价结果的主要因素。但由于这些色谱峰所代表的成分尚不明确,尚需通过GC-MS的分析以确定其化学成分结构,此部分内容尚需进一步研究。

REFERENCES

- [1] LIANG Y Z, GONG F, YU R Q. Chemometrics applied to researches on Chinese traditional medicine [J]. Prog Chem(化学进展), 1999, 2 (11): 208-212.
- [2] QIN J P, WANG ZH M, LI M L. Application of chemometrics to the identification and quality control of Chinese traditional medicine [J]. J Math Med(数据医药学杂志), 2004, 17(4): 351-354.
- [3] Ch.P(2005)Vol II(中国药典2005年版.二部)[S]. 2005: 198.

收稿日期: 2010-02-01