

# 运用 R 型因子与神经网络两种聚类分析方法 对不同产地赤芍进行质量评价

潘卫东<sup>1\*</sup>, 夏春光<sup>2</sup>

(1. 江苏联合职业技术学院连云港中医药分院, 江苏 连云港 222006;

2. 江苏正大天晴药业有限公司, 江苏 连云港 222006)

**[摘要]** 目的:运用 R 型因子与神经网络两种聚类分析方法对收集到得不同产地来源的赤芍药材样品 HPLC 指纹图谱进行聚类分析,建立对赤芍进行质量判别的分类模型。方法:收集不同产地来源的赤芍样品 40 个,用 HPLC 法测定芍药苷含量,结合传统的经验对赤芍进行分类并获得 HPLC 图谱,运用获得 R 型因子与神经网络两种方法对不同产地赤芍进行指纹图谱的聚类分析,形成判别函数。结果:两种分析方法结果比较仅 19,27,34 号样品判别结果有所差异,相似度达 91.4%。结论:两种聚类分类方法互相验证,所建立的模型能够为药材的质量评价提供一个快捷、准确、可行的鉴别方法。

**[关键词]** 赤芍; 色谱指纹图谱; R 型因子; 神经网络

**[中图分类号]** R284.1 **[文献标识码]** A **[文章编号]** 1005-9903(2012)19-0104-04

## Study on the Method of Quality Assessment of Radix Paeoniae from Different Places by R Factor Cluster Analysis Method and ANN

PAN Wei-dong<sup>1\*</sup>, XIA Chun-guang<sup>2</sup>

(1. Jiangsu Union Vocational and Technical College, Lianyungang 222006, China;

2. Jiangsu Zhengda Tianqing Pharmaceutical Co., LTD, Lianyungang 222006, China))

**[Abstract]** **Objective:** To establish a method to classify its quality and research for fingerprints of red peony root samples. R factor cluster analysis method and ANN are adopted in the study. **Method:** Collecting 40 red peony root samples from different places and classifying according to the chromatographic fingerprints of HPLC and traditional experience, the research has received discriminant functions by systematic cluster analysis with R factor cluster analysis method and ANN. **Result:** The similarity between two methods was 91.4%, there was difference between the samples of No. 19, 27, 34. **Conclusion:** Two methods verified each other, a good method of assessing quality of the red peony root was established.

**[Key words]** red peony root; chromatographic fingerprints; R factor cluster analysis; ANN

赤芍是毛茛科植物芍药 *Paeonia lactiflora* Pall. 或川赤芍 *Paeonia veitchii* Lynch 的干燥根,收载于《中国药典》(2010 年版)一部,具有清热凉血、散瘀镇痛<sup>[1]</sup>之功效。该药有抗血栓形成作用、能显著降低重型肝炎患者血清内毒素及血清肿瘤坏死因子,改善肝脏生化指标等药理作用,芍药苷为其主要活性成分<sup>[2]</sup>,尚有不能确定的化学成分。不同产地中药其成分及含量往往有较大差别<sup>[3]</sup>,笔者通过对赤

芍的指纹图谱进行研究,在中药药效多是在有效成分不完全明确的前提下,依靠其所含的多种化学成分的总体作用,将赤芍作为一个“整体”进行质量优劣的分类研究,建立质量判别方法,提供中药材原料选择、购买的依据。

### 1 材料

**1.1 仪器** 高效液相色谱仪 2487-1525-717 及色谱工作站(Waters 公司),旋转蒸发仪(上海医械专机厂),HH-2 型数字显示恒温水浴锅(常州国华仪器有限公司),海腾牌超声波清洗器(济宁海腾超声电子设备有限公司),XS105 型电子分析天平(METTCER TOCEDO)。

**[收稿日期]** 20111213(001)

**[通讯作者]** \*潘卫东,硕士,副教授,从事药物制剂与质量分析, Tel:18705138936, E-mail:pwd1212@sohu.com

**1.2 试剂** 甲醇(色谱纯,天津市瑞明威化工有限公司),磷酸二氢钾(分析纯,天津市瑞明威化工有限公司),95%乙醇(分析纯,天津市北辰弊跃化学试剂厂),HPD-300大孔吸附树脂(安徽三星树脂科技有限公司),水(注射用水) 江苏正大天晴药业自制。

**1.3 药品** 芍药苷对照品(批号110736-200320,购于中国药品生物制品检定所),赤芍样品40个(均由江苏正大天晴药业提供),对照药材赤芍(批号1093-200001,含量3.78%,购于中国药品生物制品检定所)

## 2 方法

**2.1 色谱条件** Xtimate C<sub>18</sub> 色谱柱(4.6 mm × 250 mm, 5 μm),流动相甲醇-0.05 mol·L<sup>-1</sup>磷酸二氢钾溶液(32:68),流速1 mL·min<sup>-1</sup>,柱温室温,检测波长230 nm,进样量20 μL。

**2.2 检测波长的选择** 分别取适量芍药苷对照品,用甲醇溶解配成溶液,绘制紫外吸收光谱,芍药苷的最大吸收波长为227.4 nm,本研究选择230 nm为检测波长。

**2.3 系统适用性试验** 理论塔板数以芍药苷计算为7 355,此色谱峰与相邻峰之间分离度均>1.5,对称因子均在0.95~1.05。

**2.4 对照品溶液的制备** 精密称取经五氧化二磷减压干燥36 h的芍药苷对照品12.8 mg,置25 mL量瓶中,用甲醇溶解并定容至刻度,摇匀,配成0.512 g·L<sup>-1</sup>的对照品贮备液。精密吸取芍药对照品贮备液0.5 mL,置5 mL量瓶中,用甲醇溶解并定容至刻度,摇匀,配成51.2 mg·L<sup>-1</sup>的对照品溶液。

**2.5 样品收集及预处理** 收集了不同产地、不同批次赤芍样品40个,均为野生赤芍。经药房专家鉴定,其中1,5,6,12,26,32,36,3号样品为优质品,11,18,20,22,28,29,40号样品为劣质品,其余为一般品。样品经粉碎机粉碎成粉末,过60目筛置广口瓶,密闭保存。

### 2.6 供试品溶液的制备

**2.6.1 样品提取方法选择** 文献报道赤芍提取方法较多,有加热回流<sup>[4]</sup>、超声提取<sup>[5]</sup>与闪式提取<sup>[6]</sup>等方法。本实验考察加热回流与超声提取方法,精密称取赤芍细粉约1 g,分别采用水、甲醇以及50%乙醇、70%乙醇、85%乙醇作为溶剂,提取液回收溶剂,残渣用甲醇溶解制成1 ml相当于原药材4 mg的溶液,取上清液经0.45 μm微孔滤膜滤过,取续滤液作为供试品溶液以芍药苷为指标成分按照以上色

谱条件进行含量测定。结果显示以70%乙醇为提取溶剂超声提取1 h的提取效率较高,且操作方便。

**2.6.2 供试品溶液的制备** 取赤芍样品,粉碎后过60目筛,取1 g,精密称定,加10倍量的70%乙醇超声提取3次,每次1 h,合并提取液,减压回收乙醇,残渣用甲醇溶解制成1 mL相当于原药材4 mg的溶液,取上清液经0.45 μm微孔滤膜滤过,续滤液作为供试品溶液。

## 3 结果与分析

**3.1 图谱测定时间的确定** 按照《中国药典》2010年版一部赤芍含量测定方法,取供试品溶液20 μL,注入高效液相色谱仪,记录色谱图,结果表明28 min后无有意义峰出现,故确定图谱记录时间为30 min(赤芍样品色谱图见图1)。

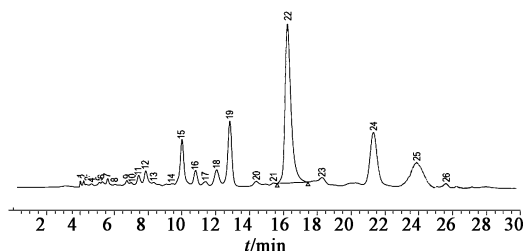


图1 赤芍样品图谱

**3.2 线性关系考察** 精密量取2.4项下芍药苷对照品贮备液0.1, 0.2, 0.4, 0.8, 1.2, 1.6, 2.0, 3.0, 4.0 mL置5 mL量瓶中,用甲醇稀释并定容至刻度,摇匀,配成系列质量浓度的溶液。经0.45 μm微孔滤膜过滤,取续滤液分别进样20 μL,在上述色谱条件下进行HPLC分析,测得峰面积,以对照品峰面积为纵坐标(Y),以质量浓度为横坐标(X)绘制标准曲线。回归方程 $Y = 106.22X + 193.35$  ( $r = 0.9993$ ),芍药苷在20.48~307.2 mg·L<sup>-1</sup>呈良好的线性关系。

**3.3 精密度试验** 精密吸取芍药苷对照品贮备液0.5 mL,置5 mL量瓶中,用甲醇稀释并定容至刻度,摇匀,配成51.2 mg·L<sup>-1</sup>的对照品溶液。在上述色谱条件下进行HPLC分析,连续进样5 d,每天进样5次,得一日内及日间精密度,日间RSD 0.45%。

**3.4 重复性试验** 称取同一样品(No. 10)6份,每份1 g,精密称定,按2.6项下操作,在上述液相条件下进行分析测定,测定其芍药苷的含量,RSD 1.4%,方法精密度良好。

**3.5 回收率试验** 称取已知含量(2.564%)的赤芍10号样品(No. 40)9份,精密称定,计算芍药苷底物含量,分别精密加入一定量的对照品溶液,按2.6项下

操作,在上述液相条件下进行 HPLC 分析,平均回收率为 99.49%,RSD 1.2%,见表 1。

表 1 芍药苷含量测定的加样回收率

| 样品中含量<br>/mg·L <sup>-1</sup> | 加入量<br>/mg·L <sup>-1</sup> | 测得量<br>/mg·L <sup>-1</sup> | 回收率<br>/% | 平均值<br>/% | RSD<br>/% |
|------------------------------|----------------------------|----------------------------|-----------|-----------|-----------|
| 24.98                        | 25.6                       | 50.46                      | 99.76     | 99.49     | 1.2       |
| 25.05                        | 25.6                       | 49.34                      | 97.41     |           |           |
| 24.57                        | 25.6                       | 49.72                      | 99.10     |           |           |
| 50.92                        | 51.2                       | 100.75                     | 98.66     |           |           |
| 51.58                        | 51.2                       | 102.01                     | 99.25     |           |           |
| 50.85                        | 51.2                       | 103.90                     | 101.81    |           |           |
| 76.45                        | 76.8                       | 153.08                     | 99.89     |           |           |
| 76.69                        | 76.8                       | 152.20                     | 99.16     |           |           |
| 76.76                        | 76.8                       | 154.05                     | 100.32    |           |           |

**3.6 样品溶液稳定性试验** 取 10 号样品溶液放置 1,2,4,8,16,24,48 h 后,在上述色谱条件下测定。结果表明,芍药苷在 48 h 内稳定性良好,RSD 1.7%。

**3.7 样品含量测定** 取赤芍样品溶液,在上述色谱条件下进行分析测定,计算赤芍中芍药苷的含量,结果见表 2。

表 2 赤芍样品中芍药苷含量测定结果 %

| No. | 芍药苷   | No. | 芍药苷   | No. | 芍药苷   |
|-----|-------|-----|-------|-----|-------|
| 1   | 3.711 | 15  | 1.802 | 28  | 2.066 |
| 2   | 3.120 | 16  | 1.989 | 29  | 1.930 |
| 3   | 2.176 | 17  | 1.726 | 30  | 1.853 |
| 4   | 2.015 | 18  | 1.998 | 31  | 2.151 |
| 5   | 3.349 | 19  | 1.760 | 32  | 3.368 |
| 6   | 3.129 | 20  | 1.734 | 33  | 3.677 |
| 7   | 2.907 | 21  | 2.057 | 34  | 2.244 |
| 8   | 2.270 | 22  | 1.717 | 35  | 2.040 |
| 9   | 2.465 | 23  | 2.822 | 36  | 3.277 |
| 10  | 2.564 | 24  | 2.618 | 37  | 3.294 |
| 11  | 1.683 | 25  | 2.346 | 38  | 2.018 |
| 12  | 2.933 | 26  | 3.584 | 39  | 2.644 |
| 13  | 2.516 | 27  | 2.108 | 40  | 1.709 |
| 14  | 3.035 |     |       |     |       |

#### 4 利用赤芍指纹图谱进行质量评价

**4.1 相对保留值和归一化面积值的计算** 测定中国生物制品检定所购得的对照药材赤芍,共获得 26 个色谱峰。其中 22 号峰为芍药苷,为赤芍主要指标成分,将其作为内参比峰。各色谱峰对内参比峰的相对保留时间定性,按照公式 4-1 计算各色谱峰相对于其内参比峰的峰面积比 X。

$$X = A_{ij} / A_{is} \quad \text{式(1)}$$

$A_{ij}$  为第  $i$  个样品第 1 个组分的峰面积,  $A_{is}$  为第  $i$  个样品的参比峰面积。若某一指纹图谱在某相对保留时间处无峰,相对峰面积为零,仍给相应的编号,以保证各色谱指纹图谱都有相同的色谱峰数。

#### 4.2 R 型因子聚类分析

**4.2.1 样品聚类分析** 将各色谱峰相对于内参比色谱峰的峰面积量化,选取 40 个赤芍样品中经过性状与理化鉴定可以确定为优质品的 36、37 号样品,一般品 38、39 号样品及劣质品 40 号作为验证样品,进行指纹图谱研究,将每个峰面积的相对峰面积  $X_i$  作为变量,得到  $40 \times 25$  阶原始数据矩阵。其中 1 ~ 35 号样品为“训练集”,建立判别模型,36 ~ 40 号样品为“学习集”,用于检验模型建立的合理性。

数据不经转换,运用“DPS 软件”采用 R 型因子聚类分析<sup>[7]</sup>,相关系数作为样品的测度。聚类分析将 35 个赤芍样品分为 3 类,结合形态学及专家鉴定结果,判定第 1 类为优质品,第 2 类为一般品,第 3 类劣质品,聚类结果见表 3。

表 3 聚类分析

| 样品级别 | 样品编号  |
|------|---|
| 1    | 1,33,12,5,6,32,26   |
| 2    | 2,3,7,10,9,4,21,13,14,15,8,16,25,19,24,31,30,27,23,34,17,35 |
| 3    | 11,18,20,22,28,29   |

**4.2.2 用已知样品对聚类方法进行检验** 将经过性状、理化鉴定分别为优质品的 36、37 号样品,一般品的 38、39 号样品及劣质品的 40 号样品作为检验样品,将色谱数据带入聚类分析数组。结果显示,判别分类结果与预测结果相符,聚类分析方法在这 5 个样品的检测中成功率为 100%。

**4.2.3 获得判别函数** 将各色谱峰相对于内参比色谱峰的峰面积量化后得到的  $40 \times 25$  阶原始数据矩阵,进行聚类分析后,获得了 3 种类别的样本,对判别结果 3 种类别中产生影响较大的变量做出选择,判别准则为贝叶斯判别函数,以马氏距离为测度,获得判别系数,并形成判别函数,3 类样本模型的判别函数分别为:

$$Y_1 = -34.0112 + 17.51 \times X_1 + 72.3921 \times X_7 - 146.0773 \times X_9 + 147.0177 \times X_{11} + 25.8711 \times X_{14}$$

$$Y_2 = -3.4159 - 0.5246 \times X_1 + 18.7094 \times X_7 + 18.6650 \times X_9 + 2.4383 \times X_{11} - 12.4320 \times X_{14}$$

$$Y_3 = -262.6948 + 69.6192 \times X_1 + 315.5969 \times X_7 - 365.6156 \times X_9 + 587.2904 \times X_{11} + 531.7006 \times X_{14}$$

将 1 ~ 35 号样品的相应变量值代入判别函数

中,求的其判别值,并根据判别值计算每个样品各属于类别的后验概率。逐步判别分析所获得的判别结果与聚类结果一致,判对概率为100%。将已知分类的36~40号样品的相应变量值代入判别函数中,判定结果为36,37号为第1类,38,39号为第2类,40号样品为第3类样品(检验样本的判别结果见表4)。

表4 检验样本的判别

| 样品<br>编号 | 实际<br>类别 | 判别<br>类别 | 后验<br>概率 | 样品<br>编号 | 实际<br>类别 | 判别<br>类别 | 后验<br>概率 |
|----------|----------|----------|----------|----------|----------|----------|----------|
| 36       | 1        | 1        | 1.000    | 39       | 2        | 2        | 1.000    |
| 37       | 1        | 1        | 1.000    | 40       | 3        | 3        | 1.000    |
| 38       | 2        | 2        | 1.000    |          |          |          |          |

**4.3 应用神经网络对样本进行分析,并进行识别**  
采用自组织竞争型神经网络<sup>[8]</sup>,完成1~35号样本的质量识别,然后用36~40号样本对该网络进行测试,看网络是否能够成功地对它进行识别。

**4.3.1 网络创建** 利用函数创建一个自组织竞争型神经网络。由于需要区分的类别数目为3,因此,神经元的数目也为3。将学习速率设置为0.1。

**4.3.2 训练网络** 网络创建后,对网络进行训练并仿真。由于训练步数的大小能影响网络的聚类性能,故分别设置训练步数为100,150,200,分别观察其分类性能。

结果发现,当训练步数为100时,网络对样本进行了初步的分类,但是分成了5类,不够精确。当训练步数为200时,所有样本一共被归为两类,其中大部分样本被分成一类,仅18,19等少部分样本归为一类,此种结果不符合要求。当训练步数为150时,样本恰好被分为三类。这与实际情况是吻合的。此时如果再提高训练步数,已经没有实际意义了。因此,确定训练步数为150次。当达到最大训练次数(150次)时,训练停止。

**4.3.3 网络测试** 利用仿真函数检验网络对上述样品模式的分类。分类结果见表5。

由此可见,网络成功地对样本进行了分类,且分类结果与DPS聚类结果基本相同,仅19,27,34号有所差异,相似度达91.4%,可以认为二者结论相同。

将应用此网络对36~40号样本进行测试,结果说明,36,37号样本属于第一类,38,39号样本属于第二类,40号样本属于第三类,与DPS聚类分析完全吻合,说明该神经网络模型可以应用于赤芍质量识别且能够取得较好的结果。

表5 自组织竞争型神经网络1~35号样品分析

| 样品<br>级别 | 样本<br>个数 | 样品编号   |
|----------|----------|--|
| 1        | 10       | 1,5,6,12,19,26,27,32,33,34                             |
| 2        | 19       | 2,3,4,7,8,9,10,13,14,15,16,17,21,23,24,<br>25,30,31,35 |
| 3        | 6        | 11,18,20,22,28,29                                      |

## 5 讨论

收集了不同产地的赤芍样品40个,并对其进行了性状与理化鉴定。对于赤芍原材料来说,其每一个化学成分及含量都是构成质量高低好坏的因素。《中国药典》中规定含量测定是以芍药苷为检测指标,但仅这赤芍苷这一种化学成分含量不足以评价赤芍原药材的质量。本研究用系统聚类分析和自组织竞争型神经网络对赤芍样品进行了HPLC指纹图谱研究,实现了对赤芍中25个化学成分色谱峰的分离,获取了反映该药材整体化学数据,将所获得的化学数据进行化学模式识别研究,将样品分为三类:优质品、一般品和劣质品,并建立了判别函数,用已知样品进行模式识别的方法验证结果符合预定的目标。用该法分析未知样品时,只需获得该样品的色谱指纹图谱,将相应变量值代入判别函数中,求的其判别值,并根据判别值计算样品所属类别,即可对样品进行鉴定、分类和评价。

## [参考文献]

- [1] 中国药典.一部[S].2005:109.
- [2] 李文,殷小杰.六种产地赤芍对大鼠抗凝血及抗血小板聚集作用的比较[J].中国实验方剂学杂志,2001,7(6):30.
- [3] 陈随清,崔璨,裴莉昕,等.不同产地和来源冬凌草药材的质量评价[J].中国实验方剂学杂志,2011,17(17):122.
- [4] 徐先祥,周丽,马燕,等.正交实验优选赤芍提取工艺研究[J].安徽中医学院学报,2008,27(5):41.
- [5] 王程程,刘玉峰,刘宇,等.正交试验法优化赤芍总苷闪式提取工艺[J].中国实验方剂学杂志,2012,18(3):12.
- [6] 高秀强,李奉勤,范文成.正交设计法研究赤芍超声提取工艺条件[J].中国现代中药,2008,10(4):24.
- [7] Vapnik V N.统计学习理论的本质[M].张学工,译.北京:清华大学出版,2000:35.
- [8] 施彦,韩力群,廉小亲.神经网络设计方法与实例分析[M].北京:北京邮电大学出版社,2009:12.

[责任编辑 顾雪竹]