

基于主成分分析和聚类分析的栀子种质资源评价

蔡晓洋¹, 张思荻¹, 曾俊¹, 李敏^{1*}, 钟兴彬²

(1. 成都中医药大学 中药材标准化教育部重点实验室, 中药资源系统研究与开发利用国家重点实验室, 成都 611137; 2. 四川众鼎中药发展有限公司, 四川 泸州 646000)

[摘要] 目的:通过探讨栀子种质资源之间的理化品质差异,为新品种选育及定向育种提供理论支持。方法:对24个种质来源的栀子种苗的主要农艺性状进行测定,采用UPLC对栀子药材中栀子苷和西红花苷-I,西红花苷-II的含量进行测定,并采用描述性统计、主成分和聚类分析方法对栀子种质来源和品质关系进行了分析。结果:24个来源的栀子的根系长度、叶形指数、叶片数和地径4项指标未表现出差异。江西1,2,4,浙江1,3,4等6个药材质量不合格,栀子苷质量分数未达到2015年版《中国药典》的规定(1.8%)。对存在差异的11项品质指标进行了主成分分析,根据主成分解释总变量和碎石图提取了4个主成分反映原来变量85.14%的信息。第1主成分主要综合了叶长、叶宽指标的信息,即种苗叶形因子;第2主成分主要综合了栀子苷,西红花苷-I,西红花苷-II和西红花苷类指标的信息,即药材内在品质因子;第3主成分主要综合了种苗鲜重和干重的信息,即种苗质量因子;第4主成分主要综合了苗高、苗高与根长比、叶片张开程度的含量信息,即种苗苗型因子。结合主成分分析综合评分表直观地反映出栀子品质的优良顺序;聚类分析将24个种质的栀子分为3类,聚类分析与综合评分排序表对种质资源类别的判定结果较为一致。结论:本研究表明四川1,3,4以及江西7等栀子品质较好,可作为新品种选育及定向育种的优良材料。

[关键词] 栀子; 种苗; 超高效液相色谱; 种质资源; 栀子苷; 西红花苷; 主成分分析; 聚类分析

[中图分类号] R282.2 **[文献标识码]** A **[文章编号]** 1005-9903(2017)14-0030-08

[doi] 10.13422/j.cnki.syfx.2017140030

[网络出版地址] <http://kns.cnki.net/kcms/detail/11.3495.R.20170420.0943.028.html>

[网络出版时间] 2017-04-20 9:43

Evaluation of Germplasm Resources of Gardeniae Fructus Based on Principal Component and Hierarchical Cluster Analysis

CAI Xiao-yang¹, ZHANG Si-di¹, ZENG Jun¹, LI Min^{1*}, ZHONG Xing-bin²

(1. Chengdu University of Traditional Chinese Medicine (TCM), Key Laboratory of Standardization of Chinese Crude Drug under Ministry of Education, State Key Laboratory of Systematic Research for Development and Utilization of TCM Resources Co-founded by Sichuan Province and MOST, Chengdu 611137, China;
2. Sichuan Zhongding TCM Development Co. Ltd., Luzhou 646000, China)

[Abstract] **Objective:** To investigate the differences in physical and chemical characteristics among germplasm resources of *Gardenia jasminoides*, in order to provide theoretical support for directional breeding and breeding of new varieties. **Method:** Main agronomic states of cultivated *Gardeniae Fructus* collected from 24 sources were determined. Content of geniposide, crocin I and crocin II in *Gardeniae Fructus* of different origins were determined using UPLC. Interrelationships between germplasm resources and quality were investigated by descriptive statistics, principal component analysis (PCA) and hierarchical cluster analysis (HCA). **Result:** There was no significant difference in root length, leaf shape index, the number of leaves and diameter of *G. jasminoides* among 24 germplasm resources. The quality of six medicines, namely Jiangxi 1, Jiangxi 2, Jiangxi 4,

[收稿日期] 20170106(015)

[基金项目] 四川省科技计划项目(2015SZ0035)

[第一作者] 蔡晓洋, 硕士, 从事中药品种、质量及资源研究, Tel:18302823285, E-mail:18302823285@163.com

[通讯作者] *李敏, 教授, 博士生导师, 从事中药品种、质量及资源研究, Tel:13980038316, E-mail:028limin@163.com

Zhejiang 1, Zhejiang 3, Zhejiang 4, did not reach the standards set forth in the Chinese pharmacopoeia (2015). The 11 quality indexes were investigated by principal component analysis. The first four components representing 85.14% of the total variability on the basis of the total variance explained and the screen plot of principal component analysis. The first principal component was related to leaf length and leaf width, which were the seedling leaf shape factor. The second principal component was related to geniposide, crocin I, crocin II and crocins, which were the herbs' intrinsic quality factor. The third principal component was related to fresh seedling weight and dry seedling weight, which were the seedling weight factor. The fourth principal component was related to seedling height, ratio of seedling height and root length, and leaf opening degree, which were the seedling form factor. PCA visually displayed the sequence of quality of *Gardeniae Fructus* HCA classified the 24 *Gardeniae Fructus* of different germplasm resources into three main categories on the basis of the measurement parameters, which was consistent with the results of PCA. **Conclusion:** This study shows that Sichuan 1, 3, 4, and Jiangxi 7 species of *Gardeniae Fructus* has a quality, and can be included in directional breeding and breeding of new varieties.

[Key words] *Gardeniae Fructus*; seeding; UPLC; germplasm resource; geniposide; crocin; principal component analysis; cluster analysis

梔子为茜草科植物梔子的干燥成熟果实,具有泻火除烦、清热利尿、凉血解毒的功效^[1]。我国药用梔子资源主要分布于江西、湖南、四川、浙江等地^[2]。在梔子主产区有不少的乡镇将其作为支柱产业,以期为农户增收致富。在梔子的长期栽培过程中,产生了许多变异,以致于表型丰富多样^[3-5]。钟传文等^[6]、曹岚等^[7]对江西产梔子主要性状和成分含量等进行了比较;周昌华等^[8-9]通过植物生物性状之间相关性的研究对梔子的形状能否稳定遗传进行了探讨。长期以来对梔子种质资源的系统研究始终较为薄弱,在梔子优势植株筛选、定向品种选育及其相关开发利用方面较薄弱且相关研究鲜有报道。从评价技术角度来看,中药材理化评价技术及评价体系尚不完善,可借鉴其他果蔬原料的研究评价方法,公丽艳等^[10]、刘振兴等^[11]、扶定等^[12]均采用主成分分析及聚类分析等方法分别对苹果、小豆、稻米的多项品质指标进行了品种评价。

有学者是通过从不同产地采集样品进行分析,得出种质资源差异的结论^[2];而本研究是课题组在梔子种质资源调查和收集的基础上,将不同产地的梔子同地栽培进行观察和分析,对 24 份梔子的 15 项品质进行测定分析,以苗高、根系长度、叶长、叶宽、果实化学成分等指标参数为依据,研究梔子种源间各指标的关系,旨在分析不同种源梔子理化品质的差异性,探讨梔子种质资源评价方法,以期为梔子种质资源评价和优良种质选育提供参考依据。

1 材料

梔子药材和种质采自四川、江西、湖南、浙江等

地(表 1)共 24 份种质资源及对应药材,经成都中医药大学李敏教授鉴定为茜草科植物梔子 *Gardenia jasminoides* 的植株及干燥成熟果实,所有种质均于 2015 年春季播种于四川省泸县梔子种质资源圃,本实验不同种质资源梔子栽培所采用的立地生长条件、气候相同,常规管理条件一致。

ACQUITY UPLC H-Class System(包括四元高压泵系统,柱温箱,自动进样器,TUV 检测器,Empower 软件色谱工作站,美国 Waters 公司),BP-121s 型 1/10 万精密电子天平(浙江精密仪器有限公司),BS-200S-WEI 型 1/1 000 电子天平(北京赛多利斯仪器系统有限公司);对照品梔子苷(四川省维克奇生物科技有限公司,批号 130428);西红花苷-I,西红花苷-II 对照品(中国食品药品检定研究院,批号分别为 111588-201202,111589-201304);甲醇(分析纯),超纯水(四川优普超纯科技有限公司)。

2 方法和结果

2.1 梔子种苗性状指标测定 测定梔子种苗具有代表性的农艺性状特征数据,包括苗高、根系长度、叶长、叶宽、叶片数、地径、叶片张开程度、种苗鲜重和干重。在种苗生长旺盛时期,测定梔子种苗各性状指标。见表 2。

种苗苗高、根系长度、叶长及叶宽等采用直尺进行测量;叶形指数为叶长与叶宽的比例,叶形一般有圆叶形、椭圆叶形、长椭圆叶形及披针形,划分依据为^[13],圆叶形为叶形指数 < 2,椭圆叶形为叶形指数在 2 ~ 2.5,长椭圆叶形为叶形指数在 2.5 ~ 3,披针形为叶形指数 > 3。

表 1 梔子资源类型与来源

Table 1 Sources and types of Gardeniae Fructus germplasm resources

| 编号 | 产地 | 种质类型 | 编号 | 产地 | 种质类型 |
|-------|----------------------|------|------|-----------|------|
| 江西 1 | 江西省抚州市金溪县合市镇富山村-冠丰一号 | 栽培 | 湖南 3 | 湖南省益阳市 | 野生 |
| 江西 2 | 江西省南昌市进贤县张公镇白李村-黄梔子 | 栽培 | 湖南 4 | 湖南衡阳市衡山县 | 栽培 |
| 江西 3 | 江西省樟树市 | 栽培 | 湖南 5 | 湖南省浏阳市 | 野生 |
| 江西 4 | 江西省樟树市 | 栽培 | 湖南 6 | 湖南省湘潭市湘潭县 | 栽培 |
| 江西 5 | 江西省抚州市金溪县 | 栽培 | 浙江 1 | 浙江省金华市磐安县 | 野生 |
| 江西 6 | 江西省吉安市永丰县 | 栽培 | 浙江 2 | 浙江省杭州市淳安县 | 栽培 |
| 江西 7 | 江西省吉安市峡江县 | 栽培 | 浙江 3 | 浙江省衢州市 | 栽培 |
| 江西 8 | 江西省吉安市泰和县 | 栽培 | 浙江 4 | 浙江省温州市苍南县 | 栽培 |
| 江西 9 | 江西省吉安市新干县 | 栽培 | 四川 1 | 四川省巴中市巴州区 | 栽培 |
| 江西 10 | 江西省宜春市上高县 | 栽培 | 四川 2 | 四川省宜宾市南溪县 | 栽培 |
| 湖南 1 | 湖南省岳阳市平江县 | 野生 | 四川 3 | 四川省泸州市泸县 | 栽培 |
| 湖南 2 | 湖南省长沙市宁乡县 | 野生 | 四川 4 | 四川省泸州市纳溪区 | 栽培 |

叶片开张程度为叶片与植株的夹角,采用量角器进行测定,一般有开张型、较开张型和闭合型 3 种形式;划分依据,开张型为叶片张角 $>45^\circ$,较开张型为叶片张角在 $30^\circ \sim 45^\circ$,闭合型为叶片张角在 30° 以下。

种苗鲜重与干重:采用电子天平称量正常生长的种苗的鲜重和干重。

结果显示,不同来源种质梔子叶形以长椭圆形(如江西 1,2,10,湖南 1~6,浙江 3,4,四川 1,2,4)和披针形(如江西 3~9,浙江 1,2,四川 3)为主;不同来源种质梔子叶片开张程度以开张型(如江西 3,9,湖南 3,4,6)和较开张型(如江西 2,4~8,10,湖南 1,2,5,浙江 1~4,四川 1~4)为主,少有闭合型(如江西 1)。不同来源种质梔子苗期苗高、根系高度、叶片数、地径、干物质积累量等存在不同程度的差异,苗高与根系比在 $0.89 \sim 1.18$,浙江种源的种苗苗高与根系比平均值最高,为 1.18 ,其次江西、四川、湖南的最低,仅为 0.89 ;叶片数在 $12.12 \sim 13.81$,四川种源的种苗叶片数平均值最高,为 13.81 ,其次湖南、浙江、江西的最低,为 12.12 ;地径在 $4.39 \sim 5.25$ mm,四川种源的种苗地径平均值最高,为 5.25 mm,其次江西、湖南、浙江的最低,为 4.39 mm;种苗单株鲜重在 $4.51 \sim 7.50$ g,四川种源的种苗干单株鲜重平均值最高,为 7.50 g,其次江西、湖南、浙江的最低,为 4.51 g;种苗单株干重在 $1.68 \sim 2.50$ g,四川种源的种苗干单株干重平均值最高,为 2.50 g,其次江西、湖南、浙江的最低,为 1.68 g。

2.2 梔子中化学成分的测定

2.2.1 色谱条件 Waters ACQUITY UPLC BEH C₁₈

色谱柱($2.1 \text{ mm} \times 50 \text{ mm}, 1.7 \mu\text{m}$);流动相为乙腈(A)-水(B)(梯度洗脱 $0 \sim 7 \text{ min}, 3\% \sim 23\% \text{ A}; 7 \sim 13 \text{ min}, 23\% \sim 50\% \text{ A}; 13 \sim 14 \text{ min}, 50\% \sim 3\% \text{ A}; 14 \sim 15 \text{ min}, 3\% \text{ A}$);检测波长 $240 \text{ nm}(0 \sim 7 \text{ min}), 440 \text{ nm}(7 \sim 15 \text{ min})$;流速 $0.3 \text{ mL} \cdot \text{min}^{-1}$;柱温 $30 \text{ }^\circ\text{C}$;进样量 $2 \mu\text{L}$ 。

2.2.2 对照品溶液制备 分别精密称定梔子苷,西红花苷-I,西红花苷-II适量,加甲醇溶解,制成混合对照品溶液,其质量浓度分别为 $205.2, 84.4, 22.2 \text{ mg} \cdot \text{L}^{-1}$ 。

2.2.3 供试品溶液的制备 分别称取各样品粉末(过四号筛)约 0.1 g ,精密称定,分别置具塞锥形瓶中,精密加入 50% 甲醇 50 mL ,称定质量,超声处理 30 min ,静置放冷,再称定质量,用 50% 甲醇补足减失的质量,摇匀, $1 \text{ 万} \text{ r} \cdot \text{min}^{-1}$ 离心 5 min ,取上清液,用 $0.22 \mu\text{m}$ 微孔滤膜滤过,即得。

2.2.4 线性关系考察 分别精密吸取上述混合对照品溶液 $0.625, 1.250, 2.500, 3.250, 5.000 \text{ mL}$ 于 5 个 10 mL 量瓶中,制成不同质量浓度的系列溶液,按 2.2.1 项下色谱条件测定,分别以对照品浓度(X)为横坐标,峰面积(Y)为纵坐标,得到相应的回归方程,结果见表 3,表明梔子苷,西红花苷-I,西红花苷-II在线性范围内线性良好。

2.2.5 精密度试验 取相同供试品溶液,按 2.2.1 项下色谱条件连续进样 6 次,计算得到梔子苷、西红花苷-I和西红花苷-II的色谱峰面积的 RSD 分别为 $0.26\%, 0.38\%, 0.59\%$,表明仪器精密度良好。

2.2.6 稳定性试验 取同批次梔子样品,按 2.2.3

表 2 不同来源梔子的主要农艺性状 (n = 20)

Table 2 Main agronomic traits of different sources of *Gardeniae Fructus* (n = 20)

| 样品 | 苗高 /cm | 根系长度 /cm | 苗高与根系比 | 叶长 /cm | 叶宽 /cm | 叶形指数 | 叶片数 /片 | 地径 /mm | 叶片张开程度/° | 鲜重 /g/株 | 干重 /g/株 |
|-------|--------|----------|--------|--------|--------|------|--------|--------|----------|---------|---------|
| 江西 1 | 21.41 | 24.02 | 1.14 | 8.16 | 2.77 | 2.97 | 13.50 | 5.14 | 28.35 | 6.36 | 2.11 |
| 江西 2 | 21.74 | 24.97 | 1.16 | 8.76 | 3.19 | 2.75 | 10.50 | 4.85 | 33.85 | 4.69 | 1.60 |
| 江西 3 | 31.20 | 26.66 | 0.86 | 8.62 | 2.66 | 3.28 | 12.50 | 4.88 | 49.40 | 4.58 | 1.81 |
| 江西 4 | 32.01 | 26.97 | 0.85 | 8.78 | 2.84 | 3.14 | 13.40 | 5.18 | 43.00 | 5.25 | 2.33 |
| 江西 5 | 32.02 | 28.05 | 0.90 | 9.04 | 2.95 | 3.08 | 11.35 | 4.91 | 43.90 | 8.49 | 2.88 |
| 江西 6 | 29.01 | 24.88 | 0.87 | 8.82 | 2.66 | 3.36 | 11.45 | 4.82 | 37.65 | 7.70 | 2.75 |
| 江西 7 | 29.34 | 31.02 | 1.08 | 8.48 | 2.79 | 3.09 | 11.80 | 4.95 | 32.00 | 6.82 | 2.46 |
| 江西 8 | 29.37 | 24.41 | 0.87 | 9.42 | 3.09 | 3.07 | 12.85 | 4.91 | 42.00 | 7.13 | 2.51 |
| 江西 9 | 25.62 | 27.69 | 1.11 | 7.44 | 2.46 | 3.09 | 10.95 | 4.80 | 49.60 | 7.89 | 2.73 |
| 江西 10 | 31.08 | 29.71 | 0.98 | 8.59 | 2.99 | 2.89 | 12.90 | 4.86 | 42.95 | 3.81 | 1.72 |
| 湖南 1 | 31.20 | 27.18 | 0.87 | 10.80 | 3.99 | 2.76 | 12.30 | 5.45 | 39.65 | 6.27 | 2.40 |
| 湖南 2 | 30.02 | 28.51 | 0.95 | 10.14 | 3.42 | 2.99 | 12.45 | 4.91 | 39.05 | 8.58 | 2.59 |
| 湖南 3 | 30.45 | 25.20 | 0.84 | 9.22 | 3.25 | 2.89 | 13.40 | 4.90 | 46.60 | 5.63 | 2.19 |
| 湖南 4 | 27.32 | 28.17 | 1.04 | 9.40 | 3.54 | 2.68 | 12.40 | 4.35 | 50.60 | 4.36 | 1.63 |
| 湖南 5 | 30.01 | 23.67 | 0.80 | 10.19 | 3.64 | 2.82 | 13.15 | 4.57 | 43.30 | 6.40 | 2.39 |
| 湖南 6 | 30.39 | 24.35 | 0.81 | 9.32 | 3.17 | 2.96 | 13.40 | 4.98 | 50.25 | 6.01 | 2.03 |
| 浙江 1 | 27.41 | 31.07 | 1.13 | 8.34 | 2.80 | 3.03 | 12.70 | 4.11 | 42.40 | 4.99 | 1.81 |
| 浙江 2 | 27.66 | 27.05 | 0.99 | 8.82 | 2.90 | 3.05 | 12.20 | 4.64 | 41.95 | 4.58 | 1.76 |
| 浙江 3 | 20.03 | 24.13 | 1.22 | 6.99 | 2.46 | 2.85 | 11.70 | 4.22 | 43.50 | 3.89 | 1.47 |
| 浙江 4 | 17.27 | 23.14 | 1.36 | 7.03 | 2.46 | 2.89 | 13.30 | 4.58 | 36.25 | 4.58 | 1.68 |
| 四川 1 | 30.05 | 33.63 | 1.12 | 8.78 | 3.39 | 2.61 | 12.90 | 5.35 | 35.30 | 10.08 | 3.37 |
| 四川 2 | 29.50 | 27.80 | 0.94 | 9.63 | 3.83 | 2.53 | 11.90 | 5.27 | 41.80 | 6.69 | 2.28 |
| 四川 3 | 33.08 | 27.35 | 0.84 | 9.46 | 3.16 | 3.03 | 14.30 | 5.10 | 32.15 | 7.64 | 2.53 |
| 四川 4 | 29.91 | 27.24 | 0.92 | 9.25 | 3.70 | 2.52 | 13.30 | 5.28 | 34.25 | 5.57 | 1.81 |

表 3 梔子苷、西红花苷-I和西红花苷-II线性关系

Table 3 Calibration curves and linear ranges of geniposide, crocin I and crocin II

| 成分 | 标准曲线 | R ² | 线性范围/mg·L ⁻¹ |
|---------|----------------------|----------------|-------------------------|
| 梔子苷 | Y = 8 680X + 18 000 | 0.999 8 | 12.825 0 ~ 205.2 |
| 西红花苷-I | Y = 30 600X + 61 100 | 0.999 6 | 5.275 0 ~ 84.4 |
| 西红花苷-II | Y = 37 079X + 10 190 | 0.999 8 | 0.277 5 ~ 22.2 |

项下方法制备供试品溶液,按 2.2.1 项下色谱条件分别于 0,2,4,8,16,24 h 时测定,计算梔子苷、西红花苷-I和西红花苷-II的色谱峰面积的 RSD 分别为 0.3%,0.6%,1.0%,表明供试品溶液在 24 h 内稳定。

2.2.7 重复性试验 取同批次的梔子样品 6 份,按 2.2.3 项下方法制备供试品溶液,并按 2.2.1 项下色谱条件测定,计算得到梔子样品中梔子苷、西红花苷-

I和西红花苷-II的色谱峰面积的 RSD 分别为 0.7%,0.7%,1.1%,表明该方法重复性良好。

2.2.8 加样回收试验 取已完成测定的梔子样品粉末 6 份(每份约 0.05 g),精密称定,每份梔子样品中加入等量的梔子苷、西红花苷-I和西红花苷-II对照品,按 2.2.3 项下方法制备供试品溶液,按 2.2.1 项下色谱条件测定,计算回收率,结果梔子苷,西红花苷-I,西红花苷-II的平均回收率分别为 102.34%,101.63%,100.77%,RSD 分别为 1.7%,1.6%,1.3%,表明该方法准确可靠。

2.2.9 样品含量测定 取 24 批样品,分别按 2.2.3 项下方法制备供试品溶液,每个样品 3 次重复,按 2.2.1 项下色谱条件进行含量测定,结果见表 4,图 1。

不同来源梔子化学成分测定结果显示:江西 1,

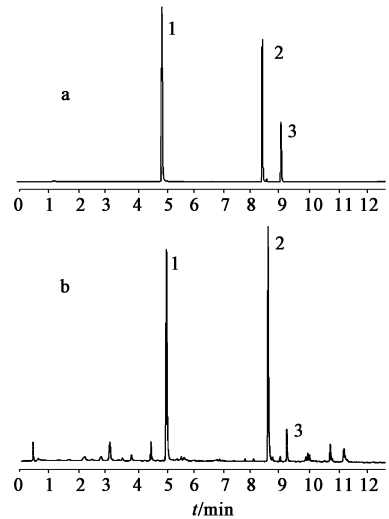
表 4 不同来源栀子药材中有效成分质量分数测定

Table 4 Determination of medicinal components in Gardeniae Fructus from different sources

| 样品 | 栀子苷 | 西红花苷-I | 西红花苷-II | 西红花苷-I, II 总量 |
|-------|------|--------|---------|---------------|
| 江西 1 | 1.23 | 1.40 | 0.22 | 1.62 |
| 江西 2 | 0.79 | 0.51 | 0.12 | 0.63 |
| 江西 3 | 2.14 | 0.93 | 0.08 | 1.01 |
| 江西 4 | 1.52 | 0.67 | 0.12 | 0.79 |
| 江西 5 | 2.33 | 0.60 | 0.08 | 0.68 |
| 江西 6 | 2.42 | 0.75 | 0.16 | 0.91 |
| 江西 7 | 3.03 | 1.08 | 0.17 | 1.25 |
| 江西 8 | 4.59 | 1.28 | 0.11 | 1.39 |
| 江西 9 | 4.24 | 1.33 | 0.24 | 1.57 |
| 江西 10 | 3.59 | 0.94 | 0.18 | 1.12 |
| 湖南 1 | 4.81 | 1.50 | 0.08 | 1.58 |
| 湖南 2 | 2.64 | 0.74 | 0.14 | 0.88 |
| 湖南 3 | 2.00 | 0.46 | 0.09 | 0.55 |
| 湖南 4 | 3.09 | 1.13 | 0.05 | 1.18 |
| 湖南 5 | 4.65 | 1.34 | 0.10 | 1.44 |
| 湖南 6 | 4.49 | 1.37 | 0.12 | 1.49 |
| 浙江 1 | 1.06 | 0.20 | 0.01 | 0.21 |
| 浙江 2 | 2.70 | 0.20 | 0.14 | 0.34 |
| 浙江 3 | 1.39 | 0.37 | 0.12 | 0.49 |
| 浙江 4 | 1.36 | 0.43 | 0.07 | 0.50 |
| 四川 1 | 5.74 | 0.72 | 0.10 | 0.82 |
| 四川 2 | 3.18 | 1.04 | 0.29 | 1.33 |
| 四川 3 | 4.56 | 1.29 | 0.15 | 1.44 |
| 四川 4 | 6.43 | 1.42 | 0.17 | 1.59 |

2,4,浙江 1,3,4 等属于不合格样品,样品中的栀子苷含量未达到 2015 年版一部《中国药典》含量测定规定的标准(1.8%)。

各省栀子样品中栀子苷,西红花苷-I,西红花苷-II 以及西红花苷-I, II 总量存在差异,栀子苷质量分数在 1.63% ~ 4.98%,四川的栀子苷平均质量分数最高,为 4.98%,其次为湖南、江西,浙江的最低,为 1.63%;西红花苷-I 质量分数在 0.30% ~ 1.12%,其中四川的西红花苷-I 平均质量分数最高,为 1.12%,其次为湖南、江西,浙江的最低,为 0.30%;西红花苷-II 质量分数在 0.09% ~ 0.18%,四川的西红花苷-II 平均质量分数最高,为 0.18%,其次为江西、湖南,浙江的最低,为 0.06%;西红花苷-I, II 质量分数在 0.39% ~ 1.30%,四川的西红花苷-I, II 总量的平均质量分数最高,为 1.30%,其



a. 对照品; b. 栀子样品; 1. 栀子苷; 2. 西红花苷-I; 3. 西红花苷-II
图 1 栀子样品 UPLC 色谱

Fig. 1 UPLC chromatograms of sample of Gardeniae Fructus

次为湖南、江西,浙江的最低,为 0.39%。

3 结果与分析

3.1 栀子种苗与药材成分描述性分析 对 24 个不同来源栀子种苗农艺性状指标和药材化学成分数据差异分析结果见表 5。其中,根系长度、叶形指数、叶片数和地径等 4 项指标的变异系数分别为 9.75%, 7.34%, 7.31% 和 6.96% 均 < 10%, 离散程度较小,上述几项指标变异系数小,取值分布较为接近,未表现出显著差异性,这可能与 24 类栀子种苗生长条件、气候、水肥管理等相同有关,其他 11 项指标的变异系数均较大,说明不同种苗多数外观品质与内在品质指标差异较大,栀子苷含量数据变异系数最大为 50.72%,数据离散程度大,各药材间指标测定值差异大,这一结果与试验所用样品产地来源等有关。比较均值与中位数,其各项指标的中位数均接近其平均数,说明这些数据的离群点较少。以上结果表明,试验所选的各项指标均在可接受范围内,离群点较少,各品质特性差异显著,具有一定的代表性。本研究所选的栀子种苗等的相关栽培管理条件一致,因此各样品品质性状的多样性来源于产地间的差异性。

3.2 主成分分析^[14] 根据前文结果,剔除根系长度、叶形指数、叶片数和地径 4 项指标,24 份栀子药材及对应栀子苗的 11 项理化品质指标的主成分分析结果见表 6。为了能够满足数据降维目的的同时包括更多的信息,常用累计方差贡献率不低于某一阈值来确定所选择的主成分数;目前研究显示,第一主成分能够最大限度地反映样本间的差异,是概括所

表 5 栀子种苗和药材理化品质描述性统计 (n = 24)

Table 5 Descriptive statistics of Gardeniae Fructus seedings and physico-chemical characterization (n = 24)

| 指标 | 极小值 | 极大值 | 中位数 | 均值 | 标准差 | 变异系数 |
|--------------|-------|-------|-------|-------|------|-------|
| 苗高/cm | 17.27 | 33.08 | 29.71 | 28.21 | 4.11 | 14.58 |
| 根系长度/cm | 23.14 | 33.63 | 27.12 | 26.95 | 2.63 | 9.75 |
| 苗高与根长比 | 0.80 | 1.36 | 0.95 | 0.99 | 0.15 | 15.15 |
| 叶长/cm | 6.99 | 10.80 | 8.82 | 8.90 | 0.91 | 10.26 |
| 叶宽/cm | 2.46 | 3.99 | 3.04 | 3.09 | 0.44 | 14.24 |
| 叶形指数 | 2.52 | 3.36 | 2.97 | 2.93 | 0.21 | 7.34 |
| 叶片数/片 | 10.50 | 14.30 | 12.60 | 12.53 | 0.92 | 7.31 |
| 地径/mm | 4.11 | 5.45 | 4.91 | 4.88 | 0.34 | 6.96 |
| 叶片张开程度/° | 28.35 | 50.60 | 41.98 | 40.82 | 6.16 | 15.09 |
| 栀子种苗鲜重/g/株 | 3.81 | 10.08 | 6.14 | 6.17 | 1.65 | 26.71 |
| 栀子种苗干重/g/株 | 1.47 | 3.37 | 2.24 | 2.20 | 0.48 | 21.89 |
| 栀子苷含量/% | 0.79 | 6.43 | 2.87 | 3.08 | 1.56 | 50.72 |
| 西红花苷-I 含量/% | 0.20 | 1.50 | 0.94 | 0.90 | 0.41 | 45.74 |
| 西红花苷-II 含量/% | 0.01 | 0.29 | 0.12 | 0.13 | 0.06 | 48.03 |
| 西红花苷类含量/% | 0.21 | 1.62 | 1.07 | 1.03 | 0.44 | 42.67 |

选指标差异信息的最佳线性函数^[15]。本研究通过主成分分析得到的第一主成分累计方差贡献率(43.31%)较低,不适合该方法确定主成分数目。

表 6 主成分分析解释的总变量

Table 6 Total variance explained of principal component analysis

| 主成分数 | 特征值 | 方差贡献率/% | 累积方差贡献率/% |
|------|------|---------|-----------|
| 1 | 4.76 | 43.31 | 43.31 |
| 2 | 1.99 | 18.11 | 61.42 |
| 3 | 1.60 | 14.57 | 75.99 |
| 4 | 1.01 | 9.15 | 85.14 |
| 5 | 0.68 | 6.14 | 91.28 |
| 6 | 0.44 | 4.02 | 95.30 |
| 7 | 0.35 | 3.21 | 98.51 |
| 8 | 0.10 | 0.90 | 99.41 |
| 9 | 0.04 | 0.35 | 99.75 |
| 10 | 0.03 | 0.25 | 100.00 |

碎石图可以用来辅助确定最佳主成分数目(图 2),碎石图中横坐标表示主成分数目,纵坐标为特征值。本研究通过考察特征值 $\lambda > 1$ 并综合考虑碎石图(主成分特征值的连续陡峭部分即为所选取的主成分数)和累积方差贡献率来确定最佳主成分数。由图 2 和表 6 可知,前 4 个主成分的特征值($\lambda > 1$)较大且连线较陡峭,即前 4 个主成分对解释变量的贡献最大,提取 4 个主成分最合适,累积方差

贡献率 85.14%,综合了栀子药材和种苗的大部分信息。

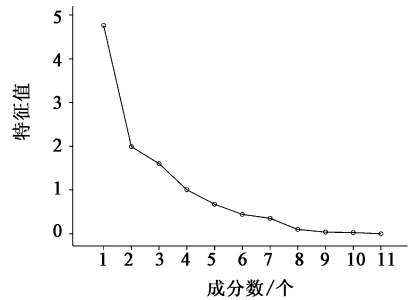


图 2 主成分分析碎石

Fig. 2 Screen plot of principal component analysis

主成分的载荷矩阵旋转之后载荷系数更接近 1,这样得到的主成分能够更好的解释和命名变量。由表 7 可知,第一主成分 PC1 主要综合了叶长、叶宽指标的信息,该类指标在第一主成分上呈正向分布,即 PC1 越大,叶长、叶宽指标的值越大;PC1 可命名为种苗叶形因子。第二主成分 PC2 主要综合了栀子苷、西红花苷-I,西红花苷-II 和西红花苷类指标的信息,栀子苷、西红花苷-I,西红花苷-II 和西红花苷类在第二主成分上呈正向分布,即 PC2 越大,栀子苷、西红花苷-I,西红花苷-II 和西红花苷类的值越大;PC2 可命名为药材内在品质因子。第三主成分 PC3 主要综合了栀子种苗鲜重和干重的信息,该类指标在第三主成分上呈正向分布,即 PC3 越大,栀子种苗鲜重和干重的值越大;PC3 可命名为

表 7 主成分分析旋转后的成分载荷矩阵

Table 7 Rotated component matrix of principal component analysis

| 指标 | 主成分 | | | |
|------------|--------|--------|--------|--------|
| | PC1 | PC2 | PC3 | PC4 |
| 苗高 | 0.605 | 0.081 | 0.439 | 0.530 |
| 苗高与根长比 | -0.617 | -0.181 | -0.238 | -0.599 |
| 叶长 | 0.936 | 0.101 | 0.173 | 0.133 |
| 叶宽 | 0.912 | 0.189 | -0.005 | -0.109 |
| 叶片张开程度 | -0.045 | -0.098 | -0.201 | 0.882 |
| 栀子种苗鲜重 | 0.128 | 0.166 | 0.942 | -0.124 |
| 栀子种苗干重 | 0.124 | 0.130 | 0.966 | 0.020 |
| 栀子苷 | 0.491 | 0.560 | 0.305 | 0.102 |
| 西红花苷-I 含量 | 0.344 | 0.885 | 0.059 | 0.064 |
| 西红花苷-II 含量 | -0.237 | 0.709 | 0.166 | -0.164 |
| 西红花苷类 | 0.289 | 0.930 | 0.079 | 0.037 |

注:旋转在 5 次迭代后收敛;PC1 ~ PC4 分标表示第一至第四主成分。

种苗质量因子。第四主成分 PC4 主要综合了苗高、苗高与根长比、叶片张开程度的含量信息,苗高、叶片张开程度在第四主成分上呈正向分布,苗高与根长比在第四主成分上呈负向分布,即 PC4 越大,苗高、叶片张开程度的值越大,苗高与根长比的值越小;PC4 可命名为种苗苗型因子。

3.3 不同种质栀子的综合评价 以各主成分对应的方差贡献率作为权重,对主成分得分和对应的权重进行线性加权,构建不同产地栀子评价函数 $Y = 43.31PC1 + 18.11PC2 + 14.57PC3 + 9.15PC4$,计算各种质栀子的综合评分值,分值越高表示该品种越好^[16]。由表 8 可知,24 个不同种质的栀子中,“四川 3”综合评价最好,其次为“湖南 1”和“四川 4”,综合评价最差的为“浙江 4”。

3.4 聚类分析 聚类分析是将样品按照品质特性相似程度逐渐聚合在一起,相似度最大的优先聚合在一起,最终按照类别的综合性质多个品种聚合,从而完成聚类分析的过程^[14]。本研究根据 24 份不同种质栀子种苗和药材的理化品质采用 Ward 法进行系统聚类分析。由聚类分析谱系图可知,在类间距离为 10 时,24 份样品分为 3 类。

第一类聚集了 16 个样品即江西 3,4,5,6,8,9,10,湖南 1,2,3,4,5,6,浙江 1,2,四川 2,这一类聚集了株高较高,苗高与根系比值居中,叶片较宽,叶片较开张,种苗较重,栀子苷、西红花苷-I 和西红花苷类成分含量较高的栀子类型,这一结果与主成分分

表 8 不同产地栀子的主成分得分及优良排序

Table 8 Principal component scores and fine sorting of different varieties of Gardeniae Fructus

| 样品 | PC1 | PC2 | PC3 | PC4 | 综合得分 | 综合排名 |
|-------|-------|-------|------|-------|--------|------|
| 江西 1 | 0.53 | -0.32 | 4.46 | 23.21 | 294.33 | 21 |
| 江西 2 | 0.11 | -0.91 | 3.21 | 26.97 | 282.06 | 22 |
| 江西 3 | -2.10 | -0.25 | 3.83 | 40.76 | 333.13 | 16 |
| 江西 4 | -0.78 | -0.99 | 4.92 | 36.44 | 353.46 | 13 |
| 江西 5 | -0.99 | -1.13 | 6.66 | 36.72 | 369.57 | 9 |
| 江西 6 | -0.26 | -0.93 | 6.18 | 31.74 | 352.09 | 14 |
| 江西 7 | 0.89 | -0.78 | 5.98 | 27.96 | 367.57 | 10 |
| 江西 8 | -0.39 | -0.08 | 5.46 | 34.75 | 379.25 | 8 |
| 江西 9 | -3.51 | 0.88 | 4.99 | 39.19 | 295.29 | 20 |
| 江西 10 | -0.58 | -0.21 | 3.81 | 36.22 | 358.08 | 12 |
| 湖南 1 | 1.32 | -0.33 | 5.20 | 33.33 | 431.90 | 2 |
| 湖南 2 | 0.43 | -1.18 | 6.41 | 32.51 | 388.14 | 6 |
| 湖南 3 | -1.22 | -0.90 | 4.53 | 38.26 | 346.73 | 15 |
| 湖南 4 | -1.92 | 0.24 | 2.76 | 40.10 | 328.37 | 17 |
| 湖南 5 | 0.07 | -0.10 | 4.95 | 35.67 | 399.45 | 5 |
| 湖南 6 | -1.77 | 0.41 | 4.30 | 40.84 | 366.94 | 11 |
| 浙江 1 | -1.29 | -1.12 | 3.96 | 34.63 | 298.73 | 19 |
| 浙江 2 | -0.78 | -0.83 | 3.83 | 34.42 | 322.02 | 18 |
| 浙江 3 | -2.87 | 0.00 | 2.12 | 33.60 | 213.98 | 23 |
| 浙江 4 | -1.81 | -0.21 | 2.56 | 27.69 | 208.65 | 24 |
| 四川 1 | 0.67 | -0.85 | 7.93 | 29.97 | 403.32 | 4 |
| 四川 2 | -0.02 | -0.42 | 5.03 | 34.39 | 379.33 | 7 |
| 四川 3 | 1.91 | -0.83 | 6.91 | 28.97 | 433.22 | 1 |
| 四川 4 | 1.69 | 0.06 | 4.92 | 29.53 | 416.04 | 3 |

析综合评分排序在前面的样品结果基本一致。第二类聚集了江西 7,四川 1,3,4,这一类聚集了株高最高,苗高与根系比值较大,叶片较宽,叶片较闭合,种苗最重,栀子苷,西红花苷-I,西红花苷-II 和西红花苷类成分含量高的栀子类型,这一结果与主成分分析综合评分排序在前面的样品结果基本一致,该类栀子树型较大且化学成分含量较高,可作为向市场提供提取栀子苷、西红花苷类等色素的优质原料及优质栀子药材的原植物材料,作为优良品种选育的首选材料。第三类聚集了江西 1,2,浙江 3,4,这一类聚集了株高最矮,苗高与根系比值大,叶片窄,叶片较闭合,种苗最轻,栀子苷,西红花苷-I,西红花苷-II 和西红花苷类成分含量低的栀子类型,这一结果与主成分分析综合评分排序在后面的样品结果基本一致,该类栀子树型较小且化学成分含量较低,不适当当前育种的目标,更不适合大面积生产利用,但具有良好的植株树型性状,具有单独保存的价值,作为优异资源保留,后期可作为盆栽等观花品种选育的首选材料。见图 3。

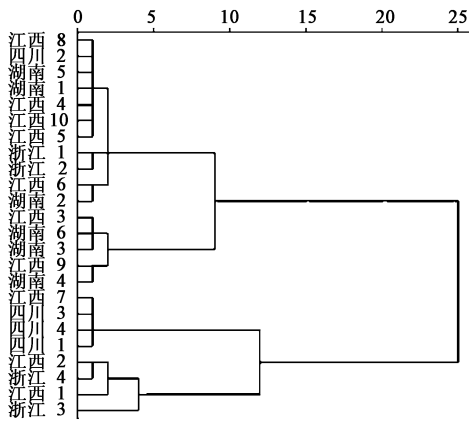


图 3 11 个品质评价因子的聚类分析谱系
Fig. 3 Cluster dendrogram of eleven indicators of quality evaluation

4 结论与讨论

主成分分析通过剔除不重要的信息,保留重要信息,以少量综合指标代表原来多个指标的大部分信息的分析方法;聚类分析是在不对信息进行删减和重要性划分的前提下,将关系更接近的研究对象合并为一类,区分类别之间的组成及界限。本研究对 15 项理化指标进行主成分分析,提取重要信息,并通过主成分综合评分分析各个主成分、理化品质指标与栀子种质资源之间的关系,为栀子的选育提供可靠直观的依据。

对江西、湖南、四川、浙江等地栀子种质资源的 15 项品质指标进行了测定并进行了结果分析,描述性分析结果表明根系长度、叶形指数、叶片数和地径等 4 项指标指标的变异系数分别为 9.75%, 7.34%, 7.31% 和 6.96% 均 < 10%, 不同种质之间以上 4 项指标的差异性不明显,这可能与 24 类栀子种苗立地生长条件、气候、水肥管理等相同有关。对变异系数较大,存在明显差异的 11 项指标进行主成分分析,提取了 4 个主成分反映原有指标的 85.14% 的信息,通过对各主成分得分和相应的权重进行线性加权,通过综合评分排序表能直观地反映出各种质栀子的优良,24 个不同种质的栀子中,“四川 3”综合评价最好,其次为“湖南 1”和“四川 4”,综合评价最差的为“浙江 4”。

按照栀子理化品质特性,采用聚类分析将 24 份种质资源分为 3 类,初步判断四川 1,3,4 以及江西 7 等种质适合成为进行品种选育的优良材料。根据综合评分结果与聚类分析结果比较,聚类分析与综合评分排序表对种质资源类别的判定结果较为

一致,所以四川 1,3,4 以及江西 7 等种质品质较好,可作为新品种选育及定向育种的优良材料。

在药材检测过程中所出现的 6 个不合格样品,可能与样品采收时间、产地、种质类型等有关。

[参考文献]

[1] 国家药典委员会. 中华人民共和国药典. 一部[M]. 北京:中国医药科技出版社,2015:284.

[2] 付小梅,赖学文,葛菲,等. 中药栀子类药材资源调查和商品药材鉴定[J]. 中国野生植物资源,2002(5): 23-25.

[3] 王碧娟. 湖北省孝南区不同类型栀子的优选及综合评价[D]. 武汉:华中农业大学,2011.

[4] 刘若楠. 栀子农家品种主要性状及有效成分比较研究[D]. 北京:中国林业科学研究院,2011.

[5] 邓绍勇,曹泉,余林,等. 栀子野生居群叶片和果实性状的表型多样性[J]. 林业科学研究,2015,28(2): 289-296.

[6] 钟传文,温达志. 黄栀子栽培技术与利用[J]. 林业科技通讯,2001(2):9-12.

[7] 曹岚,梁芳,戴泽霞. 江西新干产栀子种质资源研究初报[J]. 时珍国医国药,2008,19(4):794-795.

[8] 周昌华,张兴翠,罗健,等. 山栀子品种资源的综合评价[J]. 中国中药杂志,1998,23(3):13-60.

[9] 周昌华,张兴翠,罗健,等. 山栀生物性状之间的相关性研究[J]. 中国中药杂志,1997,22(4):16-18.

[10] 公丽艳,孟宪军,刘乃侨,等. 基于主成分与聚类分析的苹果加工品质评价[J]. 农业工程学报,2014,30(13):276-285.

[11] 刘振兴,周桂梅,陈健,等. 基于主成分与聚类分析的小豆品种综合评价[J]. 农学学报,2015,5(9): 57-63.

[12] 扶定,王青林,赵万兵,等. 基于主成分分析的稻米品质评价及聚类分析[J]. 湖北农业科学,2013,52(15):3488-3491.

[13] 詹亚华. 全国高等医药院校中医药系列教材药植植物学[M]. 3版. 北京:中国医药科技出版社,2016: 70-84.

[14] 高惠璇. 应用多元统计分析[M]. 北京:北京大学出版社,2005:118-120.

[15] 张文彤. SPSS 统计分析高级教程[M]. 北京:高等教育出版社,2004:213-233.

[16] 宋江峰,刘春泉,姜晓青,等. 基于主成分与聚类分析的菜用大豆品质综合评价[J]. 食品科学,2015,36(13):12-17.

[责任编辑 邹晓翠]