# 贯叶金丝桃质控方法提升及“辨色论质”研究Δ

李喜硕[1*]，苏本正[2,3]，曲珍妮[2,3]，朱娟娟[2,3]，戴衍朋[2,3]，石典花[2,3#]（1.山东中医药大学药学院，济南 250355；2.山东省中医药研究院，济南 250014；3.国家中医药管理局中药蜜制和制炭炮制技术与原理重点研究室，济南 250014）

**中图分类号** R932；R917 **文献标志码** A **文章编号** 1001-0408（2025）06-0661-07
**DOI** 10.6039/j.issn.1001-0408.2025.06.04

**摘 要 目的** 为贯叶金丝桃的质量控制提供参考。**方法** 采用高效液相色谱法建立20批贯叶金丝桃的指纹图谱并测定其主要成分绿原酸、芦丁、金丝桃苷、异槲皮苷、萹蓄苷、槲皮苷、槲皮素的含量；采用SPSS 26.0软件进行聚类分析。采用电子眼测定贯叶金丝桃粉末的明度值（$L^*$）、红绿值（$a^*$）和黄蓝值（$b^*$），采用机器学习算法建立基于外观色度值的贯叶金丝桃上述7种成分含量的预测模型，并采用均方根误差（RMSE）评价预测模型的预测性能。**结果** 20批贯叶金丝桃指纹图谱共标定16个共有峰，指认出9个色谱峰，分别为绿原酸、芦丁、金丝桃苷、异槲皮苷、萹蓄苷、槲皮苷、槲皮素、贯叶金丝桃素和金丝桃素，20批样品与对照图谱的相似度为0.889～0.987；绿原酸、芦丁、金丝桃苷、异槲皮苷、萹蓄苷、槲皮苷、槲皮素含量分别为0.025%～0.166%、0.048%～0.339%、0.082%～0.419%、0.017%～0.209%、0.011%～0.134%、0.020%～0.135%、0.041%～0.235%；聚类分析结果显示，当欧氏距离为1.4时，18批合格贯叶金丝桃可聚为3类。20批贯叶金丝桃的$L^*$为62.814～75.668，$a^*$为1.409～3.490，$b^*$为25.249～30.759；XGBoost、LightGBM、AdaBoost 3种预测模型的RMSE为0.008～0.070，拟合效果良好。除芦丁外，XGBoost模型预测其余6种成分的含量均具有较高的预测精度。**结论** 所建指纹图谱及含量测定方法准确、重复性好、结果可靠；结合机器学习算法构建的基于外观色度值的含量预测模型可用于贯叶金丝桃的质量控制。

**关键词** 贯叶金丝桃；指纹图谱；聚类分析；含量测定；色度值；质量评价；机器学习；预测模型

## Improvement of quality control methods and “quality evaluation via color discrimination” of *Hypericum perforatum*

LI Xishuo[1], SU Benzheng[2,3], QU Zhenni[2,3], ZHU Juanjuan[2,3], DAI Yanpeng[2,3], SHI Dianhua[2,3]（1. College of Pharmacy, Shandong University of Traditional Chinese Medicine, Jinan 250355, China; 2. Shandong Academy of Traditional Chinese Medicine, Jinan 250014, China; 3. Key Laboratory for Research of Technique and Principle of Honey-processing and Carbonizing of SATCM, Jinan 250014, China）

**ABSTRACT OBJECTIVE** To provide a reference for the quality control of *Hypericum perforatum*. **METHODS** High-performance liquid chromatography (HPLC) was used to establish fingerprints for 20 batches of *H. perforatum* and determine the contents of its main components: chlorogenic acid, rutin, hyperin, isoquercitrin, avicularin, quercitrin and quercetin. Cluster analysis was conducted using SPSS 26.0 software. The chromaticity values (luminance value $L^*$, red-green value $a^*$, and yellow-blue value $b^*$) of *H. perforatum* powder were measured using electronic eye. A prediction model for the contents of seven components in *H. perforatum* based on its appearance chromaticity values was established using machine learning algorithms. The predictive performance of the models was evaluated using root-mean-square-error (RMSE). **RESULTS** A total of 16 common peaks were calibrated in the fingerprints of 20 batches of *H. perforatum*, and 9 peaks were identified, which were chlorogenic acid, rutin, hyperin, isoquercitrin, avicularin, quercitrin, quercetin, hypericin and hyperforin; the similarities of the 20 batches of samples and reference fingerprint ranged from 0.889-0.987. The contents of chlorogenic acid, rutin, hyperin, isoquercitrin, avicularin, quercitrin and quercetin were 0.025%-0.166%, 0.048%-0.339%, 0.082%-0.419%, 0.017%-0.209%, 0.011%-0.134%, 0.020%-0.135%, 0.041%-0.235%, respectively. Cluster analysis results showed that 18 batches of qualified *H. perforatum* were

Δ **基金项目** 国家中医药管理局中药炮制技术传承基地建设项目（No.国中医药办规财函〔2022〕185号）；国家中医药管理局科技司共建科技项目（No.GZY-KJS-SD-2023-53）；国家中医药管理局高水平中医药重点学科建设项目（No.ZYYZDXK-2023121）；中央引导地方科技发展专项资金项目（No.YDZX2021096）

* **第一作者** 硕士研究生。研究方向：中药炮制。E-mail：1010946092@qq.com

# **通信作者** 研究员，博士。研究方向：中药炮制。E-mail：shidianhua81@163.com

grouped into three categories, when the Euclidean distance was set to 1.4. *L** of the 20 batches of *H. perforatum* ranged from 62.814 to 75.668, *a** ranged from 1.409 to 3.490, and *b** ranged from 25.249 to 30.759. RMSE of three prediction models, namely XGBoost, LightGBM, and AdaBoost, ranged from 0.008 to 0.070, indicating good fitting performance. XGBoost model predicted the contents of the other six components with high accuracy, except for rutin. **CONCLUSIONS** The established fingerprints and content determination methods are accurate, reproducible, and reliable. The content prediction model based on appearance chromaticity values, combined with machine learning algorithms, can be used for the quality control of *H. perforatum*.

**KEYWORDS** *Hypericum perforatum*; fingerprint; cluster analysis; content determination; chromaticity values; quality evaluation; machine learning; prediction model

贯叶金丝桃为藤黄科植物贯叶金丝桃*Hypericum perforatum* L.的干燥地上部分，味辛，性寒，归肝经，具有疏肝解郁、清热利湿、消肿通乳的功效[1]，主要分布在我国湖北、四川、陕西、甘肃、湖南、新疆等地[2]。贯叶金丝桃具有抗抑郁、抗炎、抗菌、抗氧化及保护神经细胞等药理作用[3]，其化学成分主要有黄酮类、间苯三酚类、萘骈二蒽酮类及多种挥发油类[4]，其中黄酮类化合物金丝桃苷、芦丁、槲皮素、槲皮苷、异槲皮苷、萹蓄苷等，间苯三酚类化合物贯叶金丝桃素，萘骈二蒽酮类化合物金丝桃素，酚酸类化合物绿原酸是其主要活性成分[5-6]。2020年版《中国药典》（一部）贯叶金丝桃标准项下仅收载了芦丁、金丝桃苷的薄层鉴别和金丝桃苷的含量测定方法，尚不能全面、准确地评价其内在质量。

中药指纹图谱可全面反映中药的内在化学成分信息，且多成分含量测定已逐渐成为中药常用的质量评价模式[7]。性状是中药质量评价的重要指标，与中药内在质量密切相关，因此可通过中药性状判断其质量优劣[8-9]。电子眼可将物体颜色进行客观数据化，在中药采收、基原鉴定、质量评价、真伪判断等方面均有应用[10]。机器学习是基于数据的计算模型，从数据中学习和提取有用的模式，可用于预测、分类、决策等领域，具有预测精度高、特征信息提取及非线性问题处理能力强等优势[11]。本研究通过建立贯叶金丝桃的高效液相色谱（HPLC）指纹图谱测定了其中7种成分（绿原酸、芦丁、金丝桃苷、异槲皮苷、萹蓄苷、槲皮苷、槲皮素）的含量，并通过机器学习构建了基于其外观色度值的成分含量的预测模型，旨在为贯叶金丝桃的质量控制提供参考。

## 1 材料

### 1.1 主要仪器

本研究所用主要仪器包括Waters 2695型HPLC仪（美国Waters公司），IRIS VA400型电子眼系统（法国Alpha Mos公司），XS-205DU型十万分之一电子天平（瑞士Mettler Toledo公司），BJ-150型高速多功能粉碎机（上海顶帅电器有限公司），KQ5200-DE型超声仪（昆山舒美超声仪器有限公司），DK-S26型电热恒温水浴锅、DHG-9146A型电热恒温鼓风干燥箱（上海精宏实验设备有限公司）。

### 1.2 药品与试剂

15批（编号S1～S15）贯叶金丝桃饮片和5批（编号S16～S20）贯叶金丝桃药材，经山东省中医药研究院林慧彬研究员鉴定均为真品（见表1）。绿原酸（批号BP0345，纯度≥98%）、金丝桃苷（批号BP0753，纯度≥98%）、异槲皮苷（批号BP0793，纯度≥98%）、槲皮素（批号BP1187，纯度≥98%）、贯叶金丝桃素（批号BP0751，纯度≥98%）对照品均购自成都普瑞法科技开发有限公司；金丝桃素对照品（批号BP0752，纯度≥98.0%）购自成都普思生物科技股份有限公司；芦丁（批号100080-202012，纯度≥91.6%）、槲皮苷对照品（批号111538-202007，纯度≥93.5%）均购自中国食品药品检定研究院；萹蓄苷对照品（批号BBP00890，纯度≥98.0%）购自云南西力生物技术股份有限公司；乙腈为色谱纯，其余试剂均为分析纯，水为纯净水。

**表1 20批贯叶金丝桃的来源信息**

| 编号 | 产地 | 来源 | 批号 |
|---|---|---|---|
| S1 | 陕西 | 北京大恒倍生药业有限公司 | 23051401 |
| S2 | 江西 | 河北上池本草药业有限公司 | 035230402 |
| S3 | 云南 | 亳州京皖中药饮片厂 | 230401 |
| S4 | 安徽 | 安徽泸谯药业有限公司 | 2109270022 |
| S5 | 河南 | 安徽道源堂中药饮片公司 | 220501 |
| S6 | 甘肃 | 太安堂（亳州）中药饮片公司 | 230301 |
| S7 | 甘肃 | A中药饮片公司[a] | 230601 |
| S8 | 安徽 | 聊城利民中药饮片公司 | 230902 |
| S9 | 安徽 | 安徽永刚中药饮片厂 | 230801 |
| S10 | 四川 | B药业有限公司[a] | 2306001 |
| S11 | 四川 | 河北嘉恒冷背药业有限公司 | C202302001 |
| S12 | 安徽 | 厦门燕来福制药有限公司 | 221214 |
| S13 | 河南 | 山东键新堂中药饮片公司 | 230901 |
| S14 | 河北 | 河北百草康神药业有限公司 | 2309021 |
| S15 | 河北 | 北京华邈药业有限公司 | 2023111401 |
| S16 | 湖北 | 杭州市中医院 | |
| S17 | 甘肃 | 杭州市中医院 | |
| S18 | 陕西 | 杭州市中医院 | |
| S19 | 湖北 | 杭州市中医院 | |
| S20 | 四川 | 杭州市中医院 | |

a：公司名称为化名。

## 2 方法与结果

### 2.1 指纹图谱的建立

#### 2.1.1 色谱条件

以 Zafex JX-$C_{18}$（4.6 mm×250 mm，5 μm）为色谱柱；以磷酸氢二钠缓冲液（6 mmol/L，pH 6.5）-0.02%磷酸溶液-乙腈为流动相进行梯度洗脱（见表2）；检测波长为354 nm（0～58 min，绿原酸、芦丁、金丝桃苷、异槲皮苷、蒴蓄苷、槲皮苷、槲皮素）、290 nm（58～60 min，贯叶金丝桃素）、590 nm（60～65 min，金丝桃素）；流速为1 mL/min；柱温为30 ℃；进样量为5 μL。

**表2 梯度洗脱条件**

| t/min | 磷酸氢二钠缓冲液/% | 0.02%磷酸溶液/% | 乙腈/% |
|---|---|---|---|
| 0 | 0 | 88 | 12 |
| 20 | 0 | 78 | 22 |
| 40 | 0 | 65 | 35 |
| 50 | 0 | 21 | 79 |
| 51 | 21 | 0 | 79 |
| 65 | 21 | 0 | 79 |

#### 2.1.2 供试品溶液的制备

取贯叶金丝桃粉末（过三号筛）约2 g，精密称定，置于具塞锥形瓶中，精密加入60%乙醇50 mL，称重，加热回流1 h，放冷，再次称重，用60%乙醇补足减失的质量，摇匀，经0.22 μm微孔滤膜滤过，取续滤液，即得供试品溶液。

#### 2.1.3 对照品溶液的制备

取绿原酸、芦丁、金丝桃苷、异槲皮苷、蒴蓄苷、槲皮苷、槲皮素、贯叶金丝桃素对照品各适量，精密称定，分别置于5 mL容量瓶中，加甲醇至刻度，即得各单一对照品储备液。取金丝桃素对照品适量，精密称定，置于5 mL容量瓶中，加二甲基亚砜（DMSO）1 mL溶解，再加50%甲醇至刻度，即得金丝桃素对照品储备液。精密量取上述各单一对照品储备液，置于5 mL容量瓶中，加甲醇定容，即得绿原酸、芦丁、金丝桃苷、异槲皮苷、蒴蓄苷、槲皮苷、槲皮素、贯叶金丝桃素、金丝桃素质量浓度分别为37.08、41.92、47.28、51.66、46.16、49.80、39.84、24.40、39.60 μg/mL的混合对照品溶液。

#### 2.1.4 精密度试验

取“2.1.2”项下供试品溶液（编号S11），按“2.1.1”项下色谱条件连续进样6次，以金丝桃苷为参照峰，得各共有峰相对保留时间的RSD均小于0.10%（$n$=6），相对峰面积的RSD均小于2.96%（$n$=6），表明仪器的精密度良好。

#### 2.1.5 稳定性试验

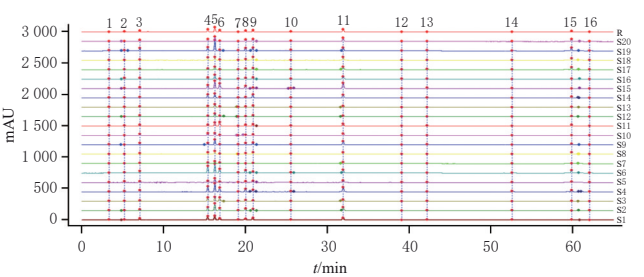
取“2.1.2”项下供试品溶液（编号S11），分别在制备后于室温下放置0、2、4、8、12、24 h时按“2.1.1”项下色谱条件进样测定，以金丝桃苷为参照峰，得各共有峰相对保留时间的RSD均小于0.09%（$n$=6），相对峰面积的RSD均小于2.63%（$n$=6），表明供试品溶液在室温下放置24 h内稳定性良好。

#### 2.1.6 重复性试验

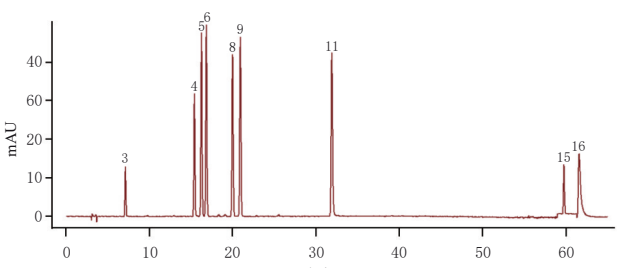
取样品（编号S11），共6份，按“2.1.2”项下方法制备供试品溶液，再按“2.1.1”项下色谱条件进样测定，以金丝桃苷为参照峰，得各共有峰相对保留时间的RSD均小于0.10%（$n$=6），相对峰面积的RSD均小于2.13%（$n$=6），表明方法的重复性良好。

#### 2.1.7 指纹图谱建立与相似度评价

取20批样品，按“2.1.2”项下方法制备供试品溶液，再按“2.1.1”项下色谱条件进样测定。将所得数据导入《中药色谱指纹图谱相似度评价系统（2012版）》，以S11样品为参照图谱（各峰信号强、分离度好），设置时间窗为0.1 min，以平均数法生成对照图谱，经多点校正和Marker峰匹配后，得到样品的叠加指纹图谱、对照指纹图谱（R）及混合对照品图谱（图1）。最终确定16个共有峰，共指认出9个峰，分别为绿原酸、芦丁、金丝桃苷、异槲皮苷、蒴蓄苷、槲皮苷、槲皮素、贯叶金丝桃素和金丝桃素。对20批样品与对照图谱进行相似度评价，结果显示，各批样品与对照图谱的相似度为0.889～0.987，均大于0.850，表明贯叶金丝桃中各色谱峰组成相似，所建指纹图谱可用于初步控制贯叶金丝桃的整体质量。




A. 20批贯叶金丝桃的叠加指纹图谱及对照指纹图谱




B. 混合对照品溶液的HPLC图

3：绿原酸；4：芦丁；5：金丝桃苷；6：异槲皮苷；8：蒴蓄苷；9：槲皮苷；11：槲皮素；15：贯叶金丝桃素；16：金丝桃素。

**图1 贯叶金丝桃的叠加指纹图谱及混合对照品的HPLC图**

### 2.2 多成分含量测定

#### 2.2.1 混合对照品溶液的制备

分别精密称取“2.1.3”项下绿原酸、芦丁、金丝桃苷、异槲皮苷、蒴蓄苷、槲皮苷、槲皮素对照品储备液适量，置于5 mL容量瓶中，加甲醇定容，即得上述各成分质量浓度分别为29.90、86.52、132.30、30.82、19.60、20.14、

67.52 μg/mL的混合对照品溶液。

2.2.2　线性关系考察

取“2.2.1”项下混合对照品溶液，分别取1、2、4、8、10、12、14 μL，按“2.1.1”项下色谱条件进样测定，以各成分进样量为横坐标（$X$）、峰面积为纵坐标（$Y$）进行线性回归。结果见表3。

**表3　各成分的回归方程与线性范围**

| 成分 | 线性范围/μg | 回归方程 | $R^2$ |
|---|---|---|---|
| 绿原酸 | 0.029 9～0.418 6 | $Y$=1 095 849.26$X$+2 180.00 | 0.999 8 |
| 芦丁 | 0.086 5～1.211 3 | $Y$=1 667 891.97$X$−5 518.18 | 0.999 9 |
| 金丝桃苷 | 0.132 3～1.852 2 | $Y$=2 150 250.03$X$−7 338.79 | 0.999 9 |
| 异槲皮苷 | 0.030 8～0.431 5 | $Y$=2 258 416.25$X$−3 051.32 | 0.999 9 |
| 萹蓄苷 | 0.019 6～0.274 4 | $Y$=1 864 218.40$X$−1 334.04 | 0.999 9 |
| 槲皮苷 | 0.020 1～0.282 0 | $Y$=2 298 336.04$X$−3 584.50 | 0.999 9 |
| 槲皮素 | 0.067 5～0.945 3 | $Y$=2 531 572.42$X$−12 178.32 | 0.999 8 |

2.2.3　精密度试验

取“2.1.2”项下供试品溶液（编号S11），按“2.1.1”项下色谱条件连续进样6次，记录峰面积。结果显示，绿原酸、芦丁、金丝桃苷、异槲皮苷、扁蓄苷、槲皮苷、槲皮素峰面积的RSD分别为0.39%、0.50%、0.41%、0.52%、0.53%、0.71%、0.94%（$n$=6），表明仪器的精密度良好。

2.2.4　稳定性试验

取“2.1.2”项下供试品溶液（编号S11），分别于制备后室温下放置0、2、4、8、12、24 h时按“2.1.1”项下色谱条件进样测定，记录峰面积。结果显示，绿原酸、芦丁、金丝桃苷、异槲皮苷、扁蓄苷、槲皮苷、槲皮素峰面积的RSD分别为0.75%、1.35%、0.89%、1.00%、0.93%、1.07%、1.27%（$n$=6），表明供试品溶液在室温下放置24 h内稳定性良好。

2.2.5　重复性试验

取样品（编号S11），共6份，按“2.1.2”项下方法制备供试品溶液，再按“2.1.1”项下色谱条件进样测定，记录峰面积并根据外标法计算各成分的含量。结果显示，绿原酸、芦丁、金丝桃苷、异槲皮苷、萹蓄苷、槲皮苷、槲皮素含量的RSD分别为1.39%、1.77%、0.64%、0.55%、1.32%、0.93%、1.41%（$n$=6），表明该方法的重复性良好。

2.2.6　加样回收率试验

精密量取已知含量的贯叶金丝桃粉末（编号S11），共6份，每份1 g，按各成分含量质量比约1∶1加入适量的各单一对照品溶液，精密加入60%乙醇至50 mL，然后按“2.1.2”项下方法制备供试品溶液，再按“2.1.1”项下色谱条件进样测定，记录峰面积并计算加样回收率。结果显示，绿原酸、芦丁、金丝桃苷、异槲皮苷、萹蓄苷、槲皮苷、槲皮素的平均加样回收率分别为93.85%、95.85%、94.53%、95.68%、93.86%、100.29%、96.48%，RSD分别为2.81%、2.68%、2.71%、2.74%、2.31%、2.57%、2.50%（$n$=6），表明方法的准确度良好。

2.2.7　样品含量测定

取20批样品粉末各2 g，精密称定，按“2.1.2”项下方法制备供试品溶液，再按“2.1.1”项下色谱条件进样测定，按外标法计算样品含量，每批样品平行测定2次。结果显示，20批样品中绿原酸、芦丁、金丝桃苷、异槲皮苷、萹蓄苷、槲皮苷、槲皮素的含量分别为0.025%～0.166%、0.048%～0.339%、0.082%～0.419%、0.017%～0.209%、0.011%～0.134%、0.020%～0.135%、0.041%～0.235%。表明不同批次贯叶金丝桃中各成分含量存在较大差异。进一步分析发现，S7和S10样品中金丝桃苷含量低于《中国药典》规定不得少于0.10%标准，故为不合格品，其余18批样品为合格品。结果见表4。

**表4　贯叶金丝桃中7种成分含量测定结果（$n$=2，%）**

| 编号 | 绿原酸 | 芦丁 | 金丝桃苷 | 异槲皮苷 | 萹蓄苷 | 槲皮苷 | 槲皮素 |
|---|---|---|---|---|---|---|---|
| S1 | 0.111 | 0.175 | 0.172 | 0.066 | 0.070 | 0.090 | 0.094 |
| S2 | 0.077 | 0.167 | 0.286 | 0.099 | 0.105 | 0.071 | 0.140 |
| S3 | 0.048 | 0.189 | 0.172 | 0.152 | 0.059 | 0.023 | 0.062 |
| S4 | 0.088 | 0.232 | 0.290 | 0.137 | 0.117 | 0.063 | 0.162 |
| S5 | 0.071 | 0.066 | 0.151 | 0.018 | 0.047 | 0.058 | 0.087 |
| S6 | 0.111 | 0.339 | 0.360 | 0.164 | 0.134 | 0.061 | 0.138 |
| S7 | 0.025 | 0.056 | 0.082 | 0.040 | 0.011 | 0.020 | 0.064 |
| S8 | 0.078 | 0.086 | 0.125 | 0.018 | 0.050 | 0.082 | 0.091 |
| S9 | 0.103 | 0.084 | 0.211 | 0.021 | 0.079 | 0.091 | 0.153 |
| S10 | 0.073 | 0.117 | 0.089 | 0.020 | 0.025 | 0.040 | 0.067 |
| S11 | 0.072 | 0.054 | 0.141 | 0.162 | 0.058 | 0.074 | 0.100 |
| S12 | 0.081 | 0.048 | 0.115 | 0.038 | 0.022 | 0.047 | 0.041 |
| S13 | 0.054 | 0.059 | 0.122 | 0.027 | 0.037 | 0.066 | 0.056 |
| S14 | 0.060 | 0.168 | 0.165 | 0.035 | 0.039 | 0.135 | 0.130 |
| S15 | 0.081 | 0.162 | 0.294 | 0.209 | 0.123 | 0.031 | 0.235 |
| S16 | 0.077 | 0.069 | 0.162 | 0.019 | 0.051 | 0.059 | 0.101 |
| S17 | 0.037 | 0.054 | 0.120 | 0.017 | 0.046 | 0.063 | 0.125 |
| S18 | 0.048 | 0.137 | 0.123 | 0.025 | 0.046 | 0.128 | 0.134 |
| S19 | 0.166 | 0.160 | 0.419 | 0.039 | 0.121 | 0.126 | 0.137 |
| S20 | 0.100 | 0.112 | 0.199 | 0.029 | 0.083 | 0.093 | 0.149 |

2.2.8　贯叶金丝桃中7种成分含量结果的聚类分析

以18批合格贯叶金丝桃中绿原酸等7种成分的含量为变量，采用SPSS 26.0软件进行聚类分析（图2）。结果显示，当欧氏距离为1.4时，18批样品聚为3类：S2、S4、S6、S15、S19聚为第1类，S1、S9、S14、S18、S20聚为第2类，S3、S5、S8、S11～S13、S16、S17聚为第3类。当欧氏距离为2.0时，第2类和第3类可聚为一类。结合含量测定结果发现，第1类样品中所含的7种成分的总含量明显高于第2类和第3类样品，而大部分第2类样品中所含的7种成分的总含量高于第3类样品，但存在S3和S11样品按总含量应归第2类，实际却归为第3类的特殊情况。对此结果进一步分析发现，原因可能与S3样品中槲皮苷和槲皮素含量、S11样品中芦丁和槲皮苷含量均低于第2类样品，与第3类样品类似有关。综上，基于聚类分析结合含量测定结果，可将18批样品中的第1类划为优品，第2类划为良品，第3类划为合格品。

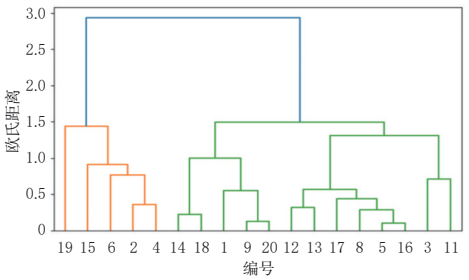



**图2　18批样品聚类分析的树状图**

### 2.3　贯叶金丝桃的颜色测定

颜色空间基于人眼对颜色的感知，能够精确捕捉和量化中药颜色的细微变化，从而真实反映其颜色特性。明度值（*L**）范围为0～100，表示颜色从黑到白的变化；红绿值（*a**）表示颜色从红到绿的变化，其中正值代表红色，负值代表绿色；黄蓝值（*b**）表示颜色从黄到蓝的变化，正值代表黄色，负值代表蓝色[12]。

取20批样品，粉碎，过三号筛，均匀放置于表面皿中，使用电子眼系统采集图像，平行采集3次后取平均值。结果显示，20批贯叶金丝桃的*L**为62.814～75.668，*a**为1.409～3.490，*b**为25.249～30.759，表明贯叶金丝桃粉末整体偏白、偏红、偏黄。结果见表5。

**表5　20批贯叶金丝桃的色度值测定结果（*n*=3）**

| 样品编号 | *L** | *a** | *b** | 样品编号 | *L** | *a** | *b** |
|---|---|---|---|---|---|---|---|
| S1 | 70.006 | 1.911 | 27.304 | S11 | 63.617 | 3.038 | 27.299 |
| S2 | 66.716 | 3.149 | 27.348 | S12 | 75.668 | 2.062 | 29.700 |
| S3 | 74.926 | 1.763 | 28.295 | S13 | 68.271 | 3.090 | 26.515 |
| S4 | 64.314 | 2.568 | 26.891 | S14 | 70.406 | 2.455 | 29.368 |
| S5 | 70.100 | 2.304 | 26.520 | S15 | 69.188 | 2.006 | 30.759 |
| S6 | 73.893 | 1.409 | 28.370 | S16 | 68.856 | 2.450 | 27.120 |
| S7 | 73.668 | 2.890 | 29.447 | S17 | 62.814 | 3.353 | 25.249 |
| S8 | 71.808 | 2.314 | 26.895 | S18 | 63.099 | 2.855 | 27.435 |
| S9 | 66.202 | 3.490 | 26.832 | S19 | 64.413 | 3.131 | 27.842 |
| S10 | 70.134 | 1.892 | 26.602 | S20 | 64.822 | 2.805 | 26.526 |

### 2.4　贯叶金丝桃外观色度预测含量研究

为探寻贯叶金丝桃外观色泽和内在质量的相关性，探索建立“辨色论质”的方法，本研究利用机器学习方法构建了基于外观色度预测贯叶金丝桃主要成分含量的模型（见图3）。

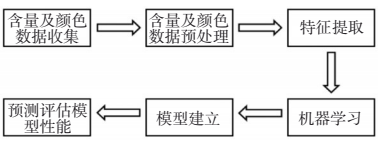



**图3　基于外观色度预测贯叶金丝桃主要成分含量的模型路径图**

#### 2.4.1　机器学习模型的建立

以20批贯叶金丝桃中绿原酸等7种成分含量、*L**、*a**、*b**值为样本特征，采用Min-Max归一化方法进行标准化或归一化处理；再通过Python绘制核密度估计图（kernel density estimation，KDE）以识别异常值并进行填充；使用Python中的“train_test_split”命令将数据集按8∶2的比例随机划分为训练集和测试集。为减少过拟合风险，本研究结合网格搜索和三折交叉验证提升模型精度，结果见图4。由图4可见，训练集与测试集中各成分分布基本一致，表明模型能很好地应用于测试集。

#### 2.4.2　预测模型的筛选

以XGBoost、LightGBM、AdaBoost等机器学习算法作为预测模型，使用均方根误差（root-mean-square-error，RMSE）评价机器学习模型效果，RMSE=$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$，式中，$y_i$代表真实的测量值，$\hat{y}_i$代表机器学习模型的预测值，$n$代表样本量。RMSE值越小，表示模型可以识别数据的能力越强，模型的预测越准确[13]。模型评估结果见表6。由表6可知，XGBoost、

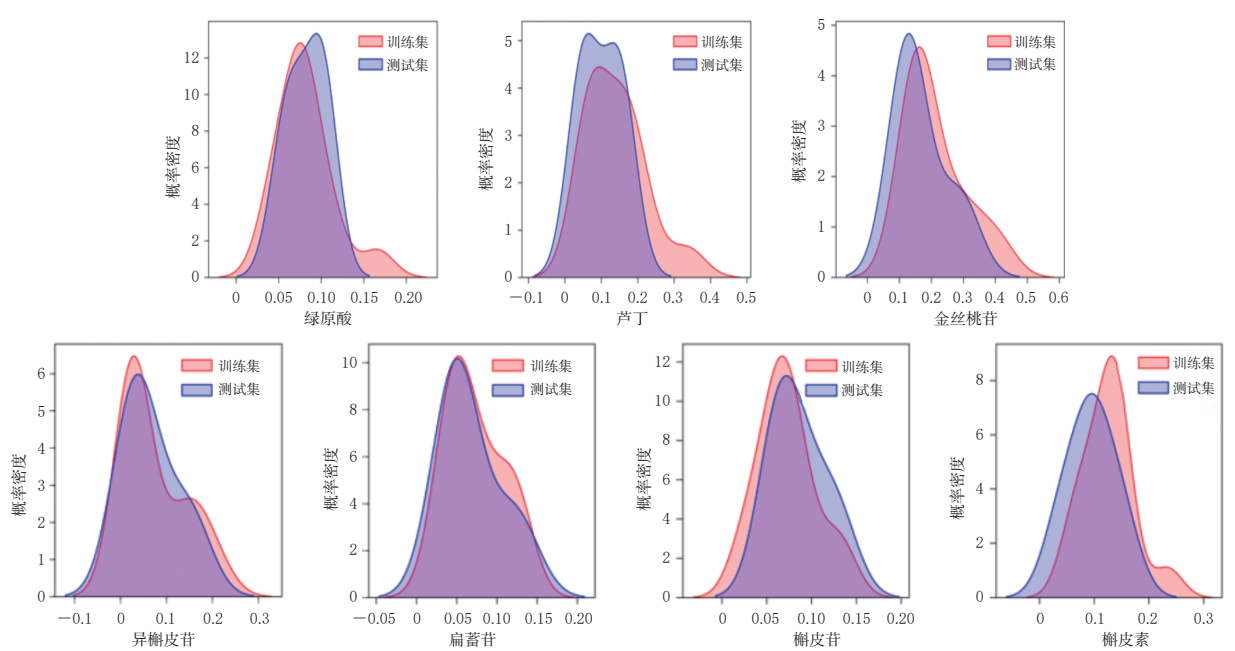



**图4　贯叶金丝桃中7种成分含量训练集与测试集的KDE分布图**

LightGBM、AdaBoost 3种机器学习算法预测模型的RMSE为0.008～0.070，均较小，未见明显差异。但是鉴于XGBoost模型能以正则化方法规避过拟风险，保持模型稳定性，精准评估特征和筛选关键项，可实现含量的精准预测，因此，本研究选择XGBoost模型用于预测贯叶金丝桃中绿原酸等7种成分的含量。

**表6　不同预测模型效果评估的RMSE值**

| 模型 | 绿原酸 | 芦丁 | 金丝桃苷 | 异槲皮苷 | 扁蓄苷 | 槲皮苷 | 槲皮素 |
|---|---|---|---|---|---|---|---|
| XGBoost | 0.031 | 0.038 | 0.020 | 0.043 | 0.019 | 0.043 | 0.014 |
| LightGBM | 0.022 | 0.044 | 0.031 | 0.037 | 0.023 | 0.033 | 0.021 |
| AdaBoost | 0.025 | 0.064 | 0.042 | 0.070 | 0.021 | 0.029 | 0.008 |

2.4.3　外观色度值预测的贯叶金丝桃中7种成分含量与实测结果分析

将样本的$L^*$、$a^*$、$b^*$特征值导入XGBoost模型作为输入项，贯叶金丝桃中7种成分的含量结果作为输出项，再将7种基于色度值预测的成分含量与实测结果进行比较，结果见图5。由图5可知，芦丁成分的预测含量与实测结果偏差较大；其余6种成分的24组预测结果中，有21组的预测值与实际值基本一致，仅3组（绿原酸中的A组，异槲皮苷和扁蓄苷中的D组）的预测值与实际值存在一定偏差；模型对金丝桃苷成分含量预测的准确率为100%。可见，建立的模型基本能够实现根据外观颜色预测贯叶金丝桃内在成分的含量。

## 3　讨论

### 3.1　色谱条件的选择

贯叶金丝桃中的间苯三酚类（贯叶金丝桃素）、萘骈二蒽酮类（金丝桃素）、黄酮类（金丝桃苷、芦丁、槲皮苷、槲皮素等）及有机酸类（绿原酸）等为其主要活性成分，这些成分复杂多样，其中黄酮类成分和有机酸类成分在354 nm波长处响应值较高，贯叶金丝桃素和金丝桃素的最大吸收波长分别为290、590 nm。为最大化获取指纹图谱信息，实现多种成分含量的同时测定，本研究在综合考察峰形、峰数、基线、吸收差异等多项干扰因素后，最终选取290、354、590 nm作为检测波长。此外，本研究对提取溶剂（不同比例甲醇溶液和乙醇溶液）和提取方式（超声提取、加热回流提取）进行了筛选，结果发现，以60%乙醇为提取溶剂，加热回流提取1 h时，各成分提取效果最佳，含量最高。

### 3.2　20批贯叶金丝桃的含量测定结果分析

本研究结果显示，绿原酸、芦丁、金丝桃苷、异槲皮苷、扁蓄苷、槲皮苷、槲皮素含量分别为0.025%～0.166%、0.048%～0.339%、0.082%～0.419%、0.017%～0.209%、0.011%～0.134%、0.020%～0.135%、0.041%～0.235%。聚类分析结果显示，18批合格的样品聚为3类，其中第1类样品中7种成分的总含量均相对较高，为优品；第2类样品为良品；第3类样品的总含量相对较低，为合格品。这提示不同产地贯叶金丝桃的成分含量存在差异，可能与生长环境不同有关；同一产地不同批次贯叶金丝桃的成分含量也存在差异，可能与药材的采收时间、农户种植水平等因素有关。

### 3.3　基于机器学习的贯叶金丝桃的含量预测研究

颜色是中药外观质量评价的重要指标，其与内在有效成分具有一定相关性[14]。张一凡等[12]基于机器学习算法将颜色特征值与测得的姜炭的5种成分含量进行相关性分析，成功建立了颜色-成分的定量预测模型，为姜炭炮制程度控制和质量评价提供了参考依据。本研究采用机器学习中的XGBoost、LightGBM和AdaBoost算法，利用$L^*$、$a^*$、$b^*$特征值与有效成分含量测定结果建立了贯叶金丝桃主要成分的含量预测模型。结果发现，绿原酸、金丝桃苷、异槲皮苷、扁蓄苷、槲皮苷、槲皮素6种成分的含量预测模型均具有较高的预测精度。其中，2020年版《中国药典》（一部）规定的指标成分金丝桃苷含量的预测准确率达到100%。上述结果提示可通过测定外观色度值来辨识贯叶金丝桃的质量。

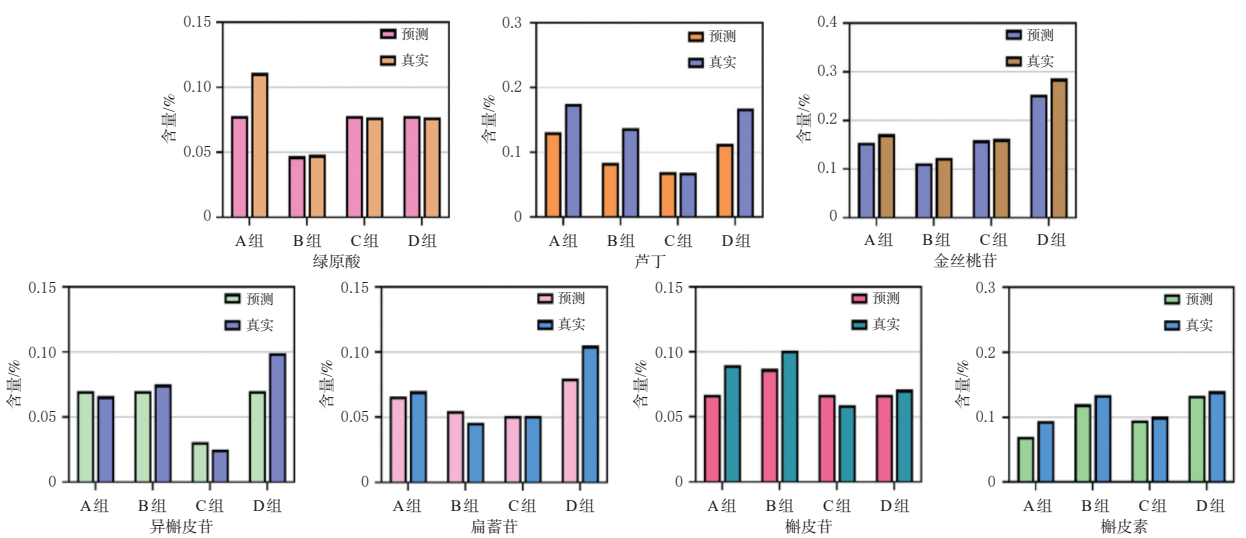



**图5　基于外观色度值预测的贯叶金丝桃中7种成分含量与实测结果的柱形图**

综上所述，所建指纹图谱及含量测定方法准确、重复性好、结果可靠；结合机器学习算法构建的基于外观色度值的含量预测模型可用于贯叶金丝桃的质量控制。

## 参考文献


[1] 国家药典委员会.中华人民共和国药典：一部[M]. 2020年版.北京：中国医药出版社，2020：242.
National Pharmacopoeia Commission. Chinese pharmacopoeia：part Ⅰ [M]. 2020 edition. Beijing：China Medical Science Press，2020：242.

[2] 闫翠起，赵曼茜，王炜，等. HPLC测定不同产地和不同部位贯叶金丝桃中的金丝桃素[J]. 华西药学杂志，2017，32(2)：211-213.
YAN C Q，ZHAO M X，WANG W，et al. Determination of hypericin in the different parts of *Hypericum perforatum* from different habitats by HPLC[J]. West China J Pharm Sci，2017，32(2)：211-213.

[3] 盛可心，王昊冉，徐凌川，等.贯叶连翘药理作用及其机制研究新进展[J]. 中药药理与临床，2023，39(3)：117-122.
SHENG K X，WANG H R，XU L C，et al. Advances in the pharmacological effects and mechanisms of *Hypericum perforatum* L.[J]. Pharmacol Clin Chin Mater Med，2023，39(3)：117-122.

[4] ALAHMAD A，ALGHORAIBI I，ZEIN R，et al. Identification of major constituents of *Hypericum perforatum* L. extracts in Syria by development of a rapid，simple，and reproducible HPLC-ESI-Q-TOF MS analysis and their antioxidant activities[J]. ACS Omega，2022，7(16)：13475-13493.

[5] 周海琴，何雨晴，陈盛君.基于UPLC指纹图谱和多指标成分定量的野生与栽培贯叶金丝桃的质量评价[J]. 中南药学，2023，21(12)：3274-3280.
ZHOU H Q，HE Y Q，CHEN S J. Quality analysis of wild and home hyperici perforati herba by UPLC fingerprint and multi-component quantitation[J]. Cent South Pharm，2023，21(12)：3274-3280.

[6] LI X X，YAN Y，ZHANG J，et al. Hyperforin：a natural lead compound with multiple pharmacological activities [J]. Phytochemistry，2023，206：113526.

[7] 李洋，陈健，张越，等.基于指纹图谱结合化学模式识别及多成分含量测定的白芍药材质量评价研究[J]. 中草药，2022，53(1)：231-237.
LI Y，CHEN J，ZHANG Y，et al. Determination of multi-components of *Paeoniae Alba Radix* based on fingerprints and chemical pattern recognition[J]. Chin Tradit Herb Drugs，2022，53(1)：231-237.

[8] 郭换，刘飞，梅国荣，等.色度分析花椒黄酮类成分含量与颜色值的相关性[J]. 中国实验方剂学杂志，2017，23(6)：91-97.
GUO H，LIU F，MEI G R，et al. Correlation between flavonoids and color values of Zanthoxyli Pericarpium based on chromatometry[J]. Chin J Exp Tradit Med Formulae，2017，23(6)：91-97.

[9] 薛淑娟，车奇涛，陈随清.基于“成分-颜色”关联分析地黄药材的质量及体外活性评价[J]. 时珍国医国药，2023，34(11)：2791-2796.
XUE S J，CHE Q T，CHEN S Q. Analysis of quality and activity evaluation *in vitro* of Radix Rehmanniae based on component color correlation[J]. Lishizhen Med Mater Med Res，2023，34(11)：2791-2796.

[10] 刘振东，兰金旭，陈随清.HPLC结合电子眼技术分析不同采收期的枸橘药材质量[J]. 中国中药杂志，2021，46(20)：5253-5259.
LIU Z D，LAN J X，CHEN S Q. Quality analysis of Citri Trifoliatae Fructus in different harvest time by HPLC combined with electronic eye technology[J]. China J Chin Mater Med，2021，46(20)：5253-5259.

[11] 周纯祎，王思雪，袁越，等.轻度认知障碍向阿尔兹海默病进展风险预测模型的系统评价[J]. 护理学报，2024，31(12)：47-53.
ZHOU C Y，WANG S X，YUAN Y，et al. Systematic evaluation of the risk prediction model of the progression of mild cognitive impairment to Alzheimer's disease[J]. J Nurs，2024，31(12)：47-53.

[12] 张一凡，周苏娟，孟江，等.基于机器视觉系统的姜炭炮制程度判别及颜色-成分相关性分析[J]. 中国药房，2022，33(22)：2712-2718.
ZHANG Y F，ZHOU S J，MENG J，et al. Discrimination of processing degree of *Zingiber officinale* charcoal and analysis of the correlation between color and component based on machine vision system[J]. China Pharm，2022，33(22)：2712-2718.

[13] 李成举，刘寅笃，秦天元，等.基于机器学习的马铃薯叶片叶绿素含量估算[J]. 光谱学与光谱分析，2024，44(4)：1117-1127.
LI C J，LIU Y D，QIN T Y，et al. Estimation of chlorophyll content in potato leaves based on machine learning [J]. Spectrosc Spectr Anal，2024，44(4)：1117-1127.

[14] 尚超凡，周巧，朱娟娟，等.电子感官技术在中药质量辨识及炮制的应用概况[J]. 中草药，2024，55(24)：8654-8663.
SHANG C F，ZHOU Q，ZHU J J，et al. An overview of the application of electronic sensory technology in the identification and processing of traditional Chinese medicine quality[J]. China Tradit Herb Drugs，2024，55(24)：8654-8663.


（收稿日期：2024-09-04　修回日期：2025-02-21）
（编辑：李　劲）