

## 基于变异系数的模糊物元模型评价不同基原甘草药材质量及快速判别研究

李瑞华<sup>1</sup>, 林夏<sup>2,3</sup>, 罗京超<sup>2,3</sup>, 陈雨竹<sup>2,3</sup>, 张永超<sup>2,3</sup>, 李执栋<sup>2,3</sup>, 闫逸伦<sup>2,3</sup>, 候化蕊<sup>2,3</sup>, 童枫<sup>2,3</sup>, 张欣<sup>2,3\*</sup>, 王振中<sup>2,3\*</sup>

1. 南京中医药大学, 江苏 南京 210023

2. 中药制药过程控制与智能制造技术全国重点实验室, 江苏 连云港 222001

3. 江苏康缘药业股份有限公司, 江苏 连云港 222001

**摘要:** 目的 综合评价乌拉尔甘草 *Glycyrrhiza uralensis* 和光果甘草 *G. glabra* 的质量, 通过化学计量学寻找质量差异标志物并建立快速判别乌拉尔甘草和光果甘草模型。方法 建立甘草药材液相指纹图谱, 确立共有峰, 共有峰数据结合模糊物元模型与偏最小二乘法判别分析 (partial least squares discriminant analysis, PLS-DA) 化学计量学方法进行质量综合评价及质量差异标志物筛选。基于近红外光谱建立不同的快速判别模型, 通过对比筛选出最佳的快速判别模型。结果 模糊物元分析表明乌拉尔甘草与光果甘草存在显著差异; 经 PLS-DA, 结果表明甘草酸铵、甘草素、甘草苷为乌拉尔甘草和光果甘草的差异标志物; 光谱经 SG 预处理, iPLS 波段筛选所建立的决策树判别模型, 精确率为 0.88、准确率为 0.88、F1 (精确率和准确率的调和平均值) 为 0.88。结论 乌拉尔甘草与光果甘草之间存在显著差异。通过近红外光谱技术建立的判别模型, 为快速区分这 2 种基原甘草提供了有效手段, 有助于提升甘草药材质量控制水平。

**关键词:** 甘草; 指纹图谱; 模糊物元模型; 化学计量学; 快速判别; 芹糖甘草苷; 甘草苷; 芹糖异甘草苷; 甘草素; 甘草酸铵

中图分类号: R286.2 文献标志码: A 文章编号: 0253-2670(2024)11-3866-08

DOI: 10.7501/j.issn.0253-2670.2024.11.028

## Evaluation of quality and spectral rapid identification of *Glycyrrhizae Radix et Rhizoma* from different origins by fuzzy matter-element model based on coefficient of variation

LI Ruihua<sup>1</sup>, LIN Xia<sup>2,3</sup>, LUO Jingchao<sup>2,3</sup>, CHEN Yuzhu<sup>2,3</sup>, ZHANG Yongchao<sup>2,3</sup>, LI Zhidong<sup>2,3</sup>, YAN Yilun<sup>2,3</sup>, HOU Huarui<sup>2,3</sup>, TONG Feng<sup>2,3</sup>, ZHANG Xin<sup>2,3</sup>, WANG Zhenzhong<sup>2,3</sup>

1. Nanjing University of Traditional Chinese Medicine, Nanjing 210023, China

2. National Key Laboratory on Technologies for Chinese Medicine Pharmaceutical Process Control and Intelligent Manufacture, Lianyungang 222001, China

3. Jiangsu Kanion Pharmaceutical Co. Ltd., Lianyungang 222001, China

**Abstract: Objective** To comprehensively evaluate the quality of *Glycyrrhiza uralensis* Fisch. and *Glycyrrhiza glabra* L. and to search for quality difference markers through chemometrics and to establish a rapid discriminant model for *G. uralensis* and *G. glabra*. **Methods** HPLC fingerprint of Gancao (*Glycyrrhizae Radix et Rhizoma*) was established, and the common peak data were combined with fuzzy matter element model and PLS-DA stoichiometric method to evaluate the quality and screen the markers of quality difference. Based on the near infrared spectrum, different rapid discrimination models were established, and the best rapid discrimination models were selected by comparison. **Results** The fuzzy matter element analysis showed that there were significant differences between *G. uralensis* and *G. glabra*. The results of PLS-DA analysis showed that ammonium glycyrrhetate, liquiritigenin and liquiritin were differential markers of *G. uralensis* and *G. glabra*. The decision tree discriminant model was established by SG pre-processing and iPLS band screening, and the accuracy rate was 0.88, the accuracy was 0.88, and the F1 was 0.88. **Conclusion** There

基金项目: 连云港市揭榜挂帅项目 (CGJBGS2101)

作者简介: 李瑞华, 硕士研究生, 研究方向为中药制药过程技术。E-mail: 3294995873@qq.com

通讯作者: 王振中, 研究员, 硕士生导师, 研究方向为中药新药研发及应用研究。E-mail: kyywzz@163.com

张欣, 研究方向为中药智造及过程分析技术。E-mail: ZXV1608@KANION.com

is a significant difference between *G. uralensis* and *G. glabra*. The discriminant model established by the near-infrared spectroscopy provided an effective method for quickly distinguishing the two kinds of *Glycyrrhizae Radix et Rhizoma*, and helped to improve the quality control of *Glycyrrhizae Radix et Rhizoma*.

**Key words:** *Glycyrrhizae Radix et Rhizoma*; fingerprinting; fuzzy matter-element model; chemometrics; rapid discriminant; liquiritin apioside; liquiritin; isoliquiritin apioside; liquiritigenin; ammonium glycyrrhetate

甘草 *Glycyrrhizae Radix et Rhizoma* 是一种应用广泛的中药材, 其来源于乌拉尔甘草 *Glycyrrhizae uralensis* Fisch.、胀果甘草 *G. inflata* Bat.或光果甘草 *G. glabra* L.的干燥根和根茎<sup>[1]</sup>。科学研究表明, 其具有抗菌, 抗炎<sup>[2-3]</sup>、抗病毒<sup>[4]</sup>、抗肿瘤<sup>[5-6]</sup>、抗氧化<sup>[7]</sup>、降血糖<sup>[8-9]</sup>、调血脂<sup>[10]</sup>、抗动脉粥样硬化<sup>[11-12]</sup>等作用。然而, 由于甘草种质遗传和地域环境等条件的影响, 不同基原甘草在某些化学成分上存在含量上的差异, 并具有种属特异性成分<sup>[13]</sup>, 因此药效也会有所不同。

为了客观评价不同基原甘草药材品质, 本实验以乌拉尔甘草和光果甘草为实验对象, 基于物元分析结合欧氏贴近度的方法, 构建了变异系数权重的模糊物元模型, 以综合评价不同基原甘草药材品质。

此外, 近红外光谱技术在中药鉴别和多基原中药识别方面具有快速、无损、简便、绿色、环保等优势。例如, 董继晶等<sup>[14]</sup>通过近、中红外光谱数据融合建立的主成分分析法 (principal components analysis, PCA) 与支持向量机 (support vector machines, SVM) 模型, 对不同基原“金银花”类药材识别准确率为 100%; 钟玉兰等<sup>[15]</sup>采用独立软模式簇类判别分析法建立近红外定性鉴别模型, 对于三七粉和掺杂比例高于 10%的伪品识别成功率为 100%。本实验利用近红外光谱技术, 结合判别模型, 建立了决策树判别分析、随机森林判别分析、支持向量机判别分析、朴素贝叶斯分类等快速判别模型。通过对比分析, 确定了最佳模型, 为近红外光谱在甘草等多基原药材快速判别方面提供了有力支持。

## 1 仪器与材料

### 1.1 仪器

Waters e2695 型高效液相色谱仪, 美国 Waters 有限公司; Ms204 型电子分析天平, 梅特勒托利多仪器 (上海) 有限公司; H1650-W 型台式高速离心机, 湖南湘仪实验室仪器开发有限公司; KQ-500DB 型数控超声波清洗器, 昆山市超声仪器有限公司; Antaris II型傅里叶近红外变换光谱仪, 配积分球漫反射检测器, 美国

Thermo Fisher Scientific 有限公司; 超纯水为实验室自制。

### 1.2 材料

甲醇 (分析纯), 南京化学试剂股份有限公司; 磷酸 (分析纯), 南京化学试剂股份有限公司; 甲醇 (色谱纯), 北京迈瑞达科技有限公司; 乙腈 (色谱纯), 北京迈瑞达科技有限公司。

甘草素 (批号 578869, 质量分数 98%); 甘草苷 (批号 551155, 质量分数 93.7%); 芹糖甘草苷 (批号 74639-14-8, 质量分数 98%); 甘草酸铵 (批号 53956-04-0, 质量分数 93%); 芹糖异甘草苷 (批号 120926-46-7, 质量分数 99%) 均购自中国食品药品检定研究院。22 批甘草药材样品信息如表 1 所示, 由江苏省食品药品检验研究院胡浩斌主任药师鉴定为乌拉尔甘草 *G. uralensis* Fisch.、光果甘草 *G. glabra* L.的干燥根和根茎。

表 1 甘草样品信息

Table 1 Sample information of *Glycyrrhizae Radix et Rhizoma*

编号	基原	产地
S1	乌拉尔甘草	内蒙古磴口县
S2	乌拉尔甘草	内蒙古乌拉特前旗
S3	乌拉尔甘草	内蒙古鄂托克前旗
S4	乌拉尔甘草	内蒙古乌赤峰
S5	乌拉尔甘草	内蒙古乌拉特前旗
S6	乌拉尔甘草	内蒙古杭锦旗
S7	乌拉尔甘草	内蒙古武川县
S8	乌拉尔甘草	内蒙古杭锦旗
S9	乌拉尔甘草	内蒙古鄂托克前旗
S10	乌拉尔甘草	内蒙古赤峰
S11	乌拉尔甘草	新疆阿勒泰
S12	乌拉尔甘草	甘肃省民勤县
S13	乌拉尔甘草	内蒙古赤峰市
S14	乌拉尔甘草	内蒙古武川县
S15	光果甘草	新疆喀什巴楚县
S16	光果甘草	新疆阿克苏
S17	光果甘草	新疆喀什巴楚县
S18	光果甘草	新疆喀什巴楚县
S19	光果甘草	新疆喀什巴楚县
S20	光果甘草	新疆喀什巴楚县
S21	光果甘草	新疆喀什巴楚县
S22	光果甘草	新疆喀什巴楚县

## 2 方法

### 2.1 色谱条件

色谱柱：FXB-Y-328 XSelect CSH C<sub>18</sub> (150 mm×4.6mm, 3.5 μm)，柱温 30 °C，体积流量 1 mL/min，检测波长 237 nm，进样量 5 μL。以乙腈 (A) -0.1%磷酸 (B) 缓冲溶液为流动相，洗脱梯度为 0~3 min, 19%~22% A; 3~6 min, 22%~30% A; 6~9 min, 30%~36% A; 9~20 min, 36%~51% A; 20~25 min, 51%~63% A; 25~35 min, 80% A。

### 2.2 溶液的制备

**2.2.1 混合对照品溶液的制备** 精密称取甘草素、甘草苷、芹糖甘草苷、甘草酸铵、芹糖异甘草苷对照品适量，置 100 mL 量瓶中，加甲醇适量，超声使溶解，放冷，加甲醇定容至刻度，制得甘草素 200 μg/mL、甘草苷 20 μg/mL、芹糖甘草苷 99.78 μg/mL、甘草酸铵 99.06 μg/mL、芹糖异甘草苷 50 μg/mL 的混合对照品溶液。

**2.2.2 供试品溶液的制备** 甘草制成甘草粉末，取约 0.2 g，精密称定，置锥形瓶中，精密加入 50 mL 甲醇，称定质量，超声提取 40 min，放冷，称定质量，加 70% 甲醇补足缺失的质量，滤过，取续滤液作为供试品溶液。

### 2.3 方法学考察

**2.3.1 参照峰选取** 图谱中 6 号峰 (甘草酸铵) 峰面积较大，分离度较好，出峰时间居中，故以其作为参照峰。

**2.3.2 重复性试验** 取甘草药材粉末，按“2.2.2”项下方法平行制备 6 份供试品溶液，分别进样进行测定。计算 7 个共有峰相对峰面积的 RSD 分别为 3.04%、0.48%、0.64%、0.96%、0.85%、1.49%、1.08%。

**2.3.3 精密度试验** 精密吸取同一供试品 (S<sub>2</sub>) 溶液 5 μL，连续进样 6 次，测定。计算 7 个共有峰相对峰面积的 RSD 分别为 3.56%、0.56%、0.01%、1.26%、2.09%、0.67%、1.73%。

**2.3.4 稳定性试验** 取“2.2.2”项下同一供试品 (S<sub>2</sub>) 溶液分别于 0、4、8、12、16、20、24 h 进行测定。计算 7 个共有峰相对峰面积的 RSD 分别为 3.90%、0.72%、0.01%、2.58%、2.71%、1.99%、1.95%。

### 2.4 光谱信息采集

**2.4.1 近红外光谱采集** 取 3 g 左右样品粉末装入石英样品杯中，平铺均匀平整，采用积分球漫反射方法按以下条件进行图谱采集：分辨率 8 cm<sup>-1</sup>，扫

描波数范围 10 000~4 000 cm<sup>-1</sup>，信号累计扫描 64 次，根据甘草的不同性状，分别从 22 批甘草中随机抽取 3 个样本，共计 66 个样本，每个样本重复扫描 3 次，取其平均值作为最后的样品谱图，共计 66 条近红外光谱，如图 1 所示。

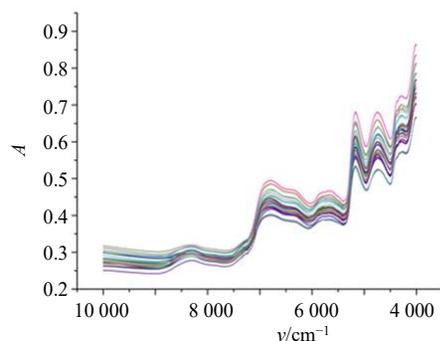


图 1 66 条甘草近红外光谱图

Fig. 1 Near-infrared spectra of 66 batches of *Glycyrrhizae Radix et Rhizoma*

**2.4.2 光谱预处理** 光谱采集时会受到环境温度、湿度、仪器状态、杂散光和随机噪声等因素的影响，红外光谱会出现基线漂移和噪声干扰等问题，建模前对光谱进行预处理，可以有效地减少噪声，净化无用信息，提高模型的精度和预测效果<sup>[16-17]</sup>，本研究采用 SG 平滑法 (Savitzky-Golay) 对 NIRS 光谱进行预处理。

**2.4.3 光谱特征波段筛选** 间隔区间偏最小二乘法 (iPLS) 波段筛选是由 Norgaard 等提出的一种波长区间筛选方法<sup>[18]</sup>。将全波段平均划分成设定值的子区间，每个子区间以不同潜变量数建立不同的 PLS 模型，依据 RMSECV 数值，筛选出最佳建模子区间。本研究利用 iPLS 波段筛选将全光谱均分为 20 个子区间，每个子区间以不同潜变量数建立不同的 PLS 模型，依据 RMSECV 数值，区间 8 对应的变量区间 5 804~6 101 cm<sup>-1</sup> 为最佳建模波段。

**2.4.4 特征提取及样本的划分** 利用 PCA 对近红外光谱数据降维<sup>[19]</sup>，剔除冗余信息，按照主成分大于 1%，累积贡献率大于 75% 的原则提取特征成分，以降维后的主成分 1 和主成分 2 对应的数据建模；以 70% 作为训练集，30% 作为测试集对样本进行划分。

## 3 结果与分析

### 3.1 指纹图谱建立及相似度分析

通过对照品比对，指认了 5 个特征峰，分别为峰 2 (芹糖甘草苷)、峰 3 (甘草苷)、峰 4 (芹糖异甘草苷)、峰 5 (甘草素) 和峰 6 (甘草酸铵)。

将所得的样品数据全谱导入国家药典委员会《中药色谱指纹图谱相似度评价系统(2012版)》软件中,采用中位数法,时间窗口设定为0.1 min,以样品S1的指纹图谱为参照图谱,选取7个共有峰进行Mark峰匹配。相似度值为0.895~0.997,表明建立的指纹图谱具有代表性,可用于甘草药材的质量控制。22批甘草的指纹图谱见图2,对照指纹图谱见图3,相似度结果如表2所示。

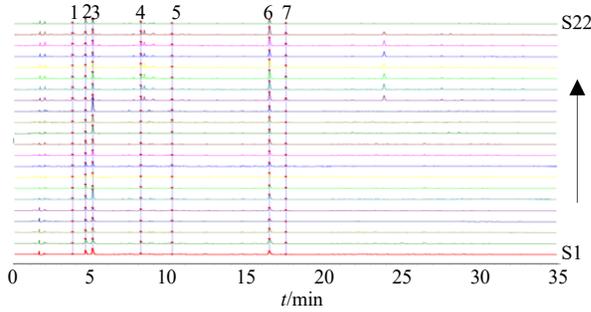
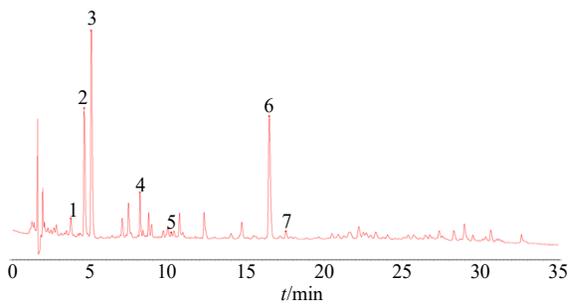


图2 22批甘草药材HPLC指纹图谱

Fig. 2 HPLC fingerprints of 22 batches of *Glycyrrhizae Radix et Rhizoma*



2-芹糖甘草苷; 3-甘草苷; 4-芹糖异甘草苷; 5-甘草素; 6-甘草酸铵。  
2-celery sugar glycyrrhizin; 3-glycyrrhizin; 4-celery sugar isoglycyrrhizin;  
5-glycyrrhizin; 6-ammonium glycyrrhetic acid.

图3 对照指纹图谱

Fig. 3 Reference peak fingerprint

表2 22批甘草相似度评价

Table 2 Evaluation of similarity of 22 batches of *Glycyrrhizae Radix et Rhizoma*

批号	相似度	批号	相似度
S1	0.959	S12	0.932
S2	0.979	S13	0.973
S3	0.996	S14	0.895
S4	0.972	S15	0.975
S5	0.983	S16	0.969
S6	0.947	S17	0.985
S7	0.922	S18	0.981
S8	0.997	S19	0.971
S9	0.986	S20	0.965
S10	0.997	S21	0.982
S11	0.988	S22	0.905

### 3.2 模糊物元模型综合评价

基于变异系数的模糊物元分析,能够避免专家赋权的主观偏好,客观反映各共有峰的相对重要程度,将22批甘草7个共有峰峰面积数据(表3)代入参考文献<sup>[20]</sup>所对应的模糊物元公式中,其中以峰面积越大越优型的优隶属度原则建立从优隶属模糊物元。相关计算公式如下所示。

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i \quad (1)$$

$$D_i = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_i)^2} \quad (2)$$

$$\delta_i = D_i \sqrt{X_i} \quad (3)$$

$$W_i = \frac{\delta_i}{\sum_{i=1}^n \delta_i} \quad (4)$$

$$\rho H_j = 1 - \sqrt{\sum_{n=1}^n W_i \Delta_{ji}} \quad (5)$$

$\bar{x}$ 为第*i*项评价指标的均值, $D_i$ 为第*i*项评价指标的均方差, $\delta_i$ 为第*i*项评价指标的变异系数, $W_i$ 为第*i*项评价指标的权重, $\rho H_j$ 为贴近度。

不同基原甘草贴近度值箱线图如图4所示,以7个共有峰峰面积为评价指标建立基于变异系数的模糊物元模型,结果表明,光果甘草和乌拉尔甘草存在质量差异。采用SPSS Statistics 26中单因素方差分析两者之间的品质差异,结果表明,乌拉尔甘草和光果甘草存在显著性差异, $P$ 值小于0.05。

### 3.3 偏最小二乘法判别分析 (partial least squares discriminant analysis, PLS-DA) 筛选差异标志物

为了进一步区分与比较乌拉尔甘草、光果甘草药材,寻找不同基原之间的差异标志物,采用有监督的PLS-DA。按照基原将甘草分为2组,以甘草指纹图谱的7个共有峰峰面积为自变量,建立PLS-DA模型,如图5所示。模型拟合参数为 $R^2_X=0.692$ , $R^2_Y=0.875$ , $Q^2=0.755$ ,表明模型稳定,预测能力良好。变量重要性投影(variable importance in projection scores, VIP)在PLS-DA中用于评价,变量对分类的贡献程度,将 $VIP>1$ 的峰作为差异标志物。甘草药材VIP见图6,共筛选出3个差异性标志物,是峰6(VIP=1.28)、峰5(VIP=1.17)、峰3(VIP=1.04),分别对应着甘草酸铵、甘草素、甘草苷,表明这3种成分在2种基原甘草之间存在显著差异。

为防止模型出现过拟合,通过置换检验200次对模型内部进行验证,结果如图7所示, $R^2$ 、 $Q^2$ 回

表 3 22 批甘草共有峰峰面积  
Table 3 Common peak area of 22 batches of *Glycyrrhizae Radix et Rhizoma*

批号	相对峰面积						
	峰 1	芹糖甘草苷	峰 7	甘草苷	芹糖异甘草苷	甘草素	甘草酸铵
S1	29 349	187 421	312 633	48 718	17 114	222 932	12 672
S2	41 071	261 458	282 657	90 781	29 425	338 841	46 090
S3	36 906	130 654	217 691	43 099	14 982	280 380	39 876
S4	22 252	106 292	209 819	42 312	29 135	157 775	12 771
S5	28 198	146 901	160 096	53 716	35 040	212 249	30 783
S6	54 112	169 791	448 756	57 556	35 129	269 502	23 145
S7	34 851	180 939	113 677	63 989	12 090	166 793	26 899
S8	44 450	133 532	199 845	42 510	21 694	243 242	40 141
S9	51 596	165 906	202 292	60 085	23 008	224 874	39 545
S10	44 143	105 944	226 036	36 248	20 623	241 227	36 296
S11	51 002	176 792	218 611	64 067	19 852	243 986	41 741
S12	41 094	86 520	384 385	26 198	39 198	231 362	11 214
S13	55 301	123 748	329 192	48 511	21 720	291 399	26 118
S14	38 407	70 024	588 324	38 255	50 246	305 414	12 874
S15	53 652	351 458	176 960	110 956	28 801	420 241	88 460
S16	110 538	120 706	128 518	301 967	60 852	498 305	131 262
S17	106 619	103 426	297 995	131 872	38 571	516 368	113 146
S18	91 070	138 850	325 486	124 859	39 976	490 192	121 134
S19	92 064	144 664	298 698	127 630	42 460	481 400	129 978
S20	85 621	97 675	204 989	88 333	41 780	377 501	85 481
S21	87 086	111 932	246 875	127 061	56 946	474 839	115 048
S22	78 170	129 947	287 912	121 667	31 444	460 821	106 918

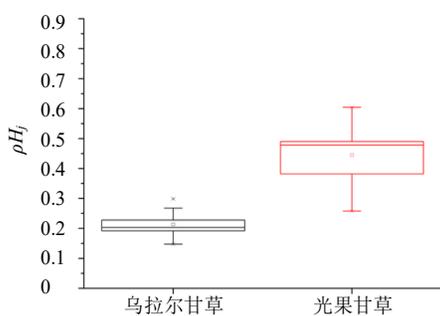


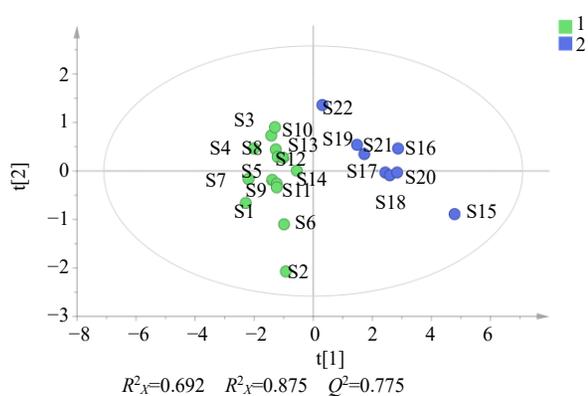
图 4 不同基原甘草贴近度值  $\rho H_j$

Fig. 4 Euclid approach degree  $\rho H_j$  for *Glycyrrhizae Radix et Rhizoma* from different origins

归线的截距分别为 0.397、-0.267, 均小于原始值<sup>[21]</sup>, 证明模型稳定可靠。

### 3.4 近红外快速判别模型

模糊物元模型结果表明, 光果甘草和乌拉尔甘草存在质量差异, PLS-DA 结果也表明了甘草酸铵、甘草素、甘草苷在两者之间存在显著差异。因此, 为保障药用资源合理开发利用及提升产业经济效



1-乌拉尔甘草; 2-光果甘草。  
1-*G. uralensis*; 2-*G. glabra*。

图 5 不同基原甘草 PLS-DA 得分图

Fig. 5 PLS-DA scores of *Glycyrrhizae Radix et Rhizoma* from different origins

益, 亟需一种新方法可以实现 2 种基原甘草的快速准确区分。目前, 近红外光谱结合判别模型应用于中药产地, 基原等的快速鉴别, 应用广泛的判别模

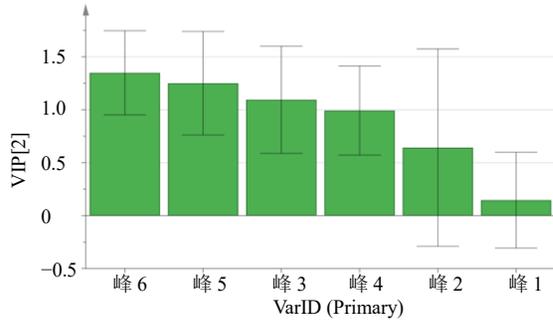


图 6 不同基原甘草 VIP 值

Fig. 6 VIP values of *Glycyrrhizae Radix et Rhizoma* from different origins

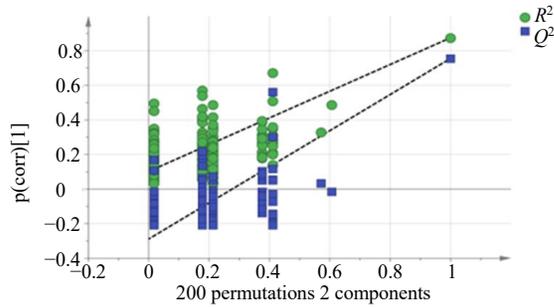


图 7 PLS-DA 模型置换检验

Fig. 7 PLS-DA model displacement test

型有支持向量机 (support vector machine, SVM), 决策树法 (decision tree, DT), 朴素贝叶斯法 (bayesian model, NBM), 随机森林算法 (random forset, RF) 等。

SVM 是一类按监督学习方法对数据进行二元分类的广义线性分类器, 其决策边界是对学习样本求解的最大边距超平面<sup>[22-24]</sup>; DT 是一种树状结构,

每棵决策树由 3 类节点构成, 即根节点、内部节点和叶节点, 根节点是判定的起始节点, 对棵决策树的决策效果起到重要的作用, 相比于其他特征的判定条件有更高的重要性, 叶节点就是在多级判定条件下的类别决策结果<sup>[21]</sup>; NBM 是贝叶斯分类学中使用较为广泛的算法。该算法本身来源于贝叶斯定理。在确定目标时, 认为各部分的属性特征相互独立, 每个对象的特征矢量的维度也都相互独立, 互不相关<sup>[25-28]</sup>; RF 法是一种著名的集成学习方法, 其核心思想是用随机的方式建立一个森林, 该森林由许多决策树组成, 以决策树为基分类器构成大型多分类器<sup>[29]</sup>。本实验中, 4 种模型的超参数设置见表 4。

本研究基于近红外光谱, 将通过预处理, 特征波段筛选建立不同的判别模型, 以准确率、精确率、F1 3 个量化指标评价判别模型性能, 通过对比筛选出最佳判别模型。其中, 准确率为预测正确样本占总样本的比例, 准确率越大越好。精确率为预测出来为正样本的结果中, 实际为正样本的比例, 精确率越大越好。F1 为精确率和准确率的调和平均。基于近红外光谱技术, 构建了一种快速判别模型, 具体结果如表 5 所示。

经过模型性能参数对比分析, 训练集结果显示 DT 模型和 RF 模型效果最优, 精确率为 1、准确率为 1, F1 为 1。测试集结果显示, DT 模型精确率为 0.88、准确率为 0.88, F1 为 0.88, 模型效果最优, RF 模型效果次之, SVM 模型效果最差。SVM 模型效果最差, 这是由于该模型需要的训练集样本量大, 本研究样本量较少。SG 具有滤除噪音的作用, iPLS 波段筛选去除无关和干扰变量, 提高了模型的稳健

表 4 4 种模型超参数设置

Table 4 Hyperparameter setting table of four models

模型	模型超参数
CART	树的最大深度为 10, 叶子节点的最大数量为 50, 叶子节点的最小样本数为 1, 内部节点分裂的最小样本数为
RF	预测变量数设置为 11, 生成树的棵树为 200, 缺失值惩罚为 1
SVM	核函数为 linear, 核函数系数为 scale, 缺失值惩罚为 1, 最大迭代次数为 1 000
NBM	先验分布为高斯分布, $\alpha$ 为 1

表 5 判别模型比较

Table 5 Comparison of rapid discriminant models

模型名称	训练集			测试集		
	准确率	F1	精确率	准确率	精确率	F1
DT	1.00	1.00	1.00	0.88	0.88	0.92
RF	1.00	1.00	1.00	0.63	0.63	0.83
SVM	0.53	0.53	0.37	0.50	0.50	0.33
NBM	0.94	0.94	0.95	0.50	0.50	0.43

性,此外,DT模型需要的训练集样本量少,能够处理多路输出问题,避免数据泛化性,因此,DT模型相较另外3种模型能够较好的实现2种基原的快速分离。

#### 4 讨论

本研究首先对供试品溶液制备方法进行了探讨,考察不同提取溶剂、提取方式、提取时间及提取溶剂用量,最终确定以70%甲醇为提取溶剂,溶剂用量为50 mL,超声处理样品,提取40 min为最佳样品前处理方法。

研究基于模糊物元分析,采用变异系数评价甘草共有峰的权重系数,避免专家赋权的主观偏好,客观反映各共有峰的相对重要程度,结合欧氏贴近度,对不同基原甘草质量进行综合评价,建立了一种适用于不同基原药材品质的综合评价方法,该方法可以有效的评价乌拉尔甘草和光果甘草药材质量,并通过建立PLS-DA模型可以准确识别出2种基原之间成分差异。

鉴于乌拉尔甘草和光果甘草存在较为明显的质量差异,为保证临床用药的准确性,需要对这2种基原的甘草实现区分,因而本研究采用近红外光谱技术实现2种基原的定性分析。研究采用SG预处理和iPLS特征波段筛选后,建立的DT模型效果最佳,模型精确率、准确率和F1均达到了0.88。

本研究表明模糊模型可以用于不同基原药材质量评价,PLS-DA可以识别不同基原药材成分差异,为不同基原药材综合评价提供一种新的思路。然而,鉴于中药成分复杂且多靶点作用,仅用7个共有峰可能无法完全代表甘草药材质量。因此,未来通过方法改进,分离出更多共有峰,并基于变异系数权重的模糊物元模型综合评价甘草药材质量。决策树判别模型较为准确实现两种基原甘草的定性判别,但是准确率尚显不足,无法完全实现对两种甘草基原的识别,为提高模型的准确率,可以通过进一步扩大建模样本量,优化模型超参数,以及引进新的检测技术和算法等方式,建立更加科学、精确的甘草基原药材快速检测手段。

**利益冲突** 所有作者均声明不存在利益冲突

#### 参考文献

[1] 中国药典 [S]. 四部. 2020: 2501.  
[2] 钟承志, 吴德松, 颜宏, 等. 丹参和甘草提取物抗菌、抗炎作用研究 [J]. 云南中医中药杂志, 2023, 44(7): 68-

71.

- [3] 高嘉豪. 甘草酸对肉仔鸡免疫、抗炎及抗氧化作用的影响 [D]. 石河子: 石河子大学, 2022.  
[4] 王晓敏, 张芳红, 郭红云, 等. 甘草查尔酮 B 抗肿瘤、抗炎及抗病毒作用研究进展 [J]. 甘肃医药, 2023, 42(10): 865-867.  
[5] 徐卓, 陈颖翀, 林龙飞, 等. 异甘草素抗肿瘤新型给药系统的研究进展 [J]. 中草药, 2023, 54(17): 5806-5815.  
[6] 张尚龙, 闫潇, 杨秀娟, 等. 基于网络药理学和细胞验证探讨甘草抗肺癌机制 [J]. 中国中医药信息杂志, 2023, 30(10): 36-40.  
[7] 董巍, 王文豹, 王晓丽, 等. 超声法辅助提取内蒙古甘草多糖工艺及其抗氧化活性研究 [J]. 现代食品, 2023, 29(16): 91-95.  
[8] 王晴, 洪叶, 柳振宇, 等. 甘草多糖的抗氧化及降血糖作用研究 [J]. 食品工业科技, 2023, 44(1): 398-404.  
[9] 刘杰. 甘草次酸提取、氨基葡萄糖修饰及其降血糖活性研究 [D]. 天津: 天津科技大学, 2022.  
[10] 刘晓丹, 刘俊, 李云, 等. 甘草提取物及有效成分抗代谢相关脂肪性肝病的药理作用及机制研究进展 [J]. 药物评价研究, 2024, 47(3): 642-649.  
[11] 王花, 姚元滨, 何伍健. 异甘草素通过调节自噬作用减轻动脉粥样硬化大鼠斑块形成 [J]. 中国老年学杂志, 2023, 43(19): 4795-4799.  
[12] 周强, 杜芬. 甘草抗动脉粥样硬化机制的网络药理学研究 [J]. 生物资源, 2020, 42(2): 194-204.  
[13] 李娜, 张晨, 钟赣生, 等. 不同品种甘草化学成分、药理作用的研究进展及质量标志物(Q-Marker)预测分析 [J]. 中草药, 2021, 52(24): 7680-7692.  
[14] 董继晶, 齐路明, 王科, 等. 基于灰色关联-TOPSIS 法的“金银花”类药材品质评价及其多元光谱基原识别研究 [J]. 中国中药杂志, 2023, 48(10): 2713-2724.  
[15] 钟玉兰, 乐智勇. 利用近红外光谱技术快速鉴别三七粉及其伪品 [J]. 江西中医药大学学报, 2018, 30(3): 70-73.  
[16] 童枫, 徐芳芳, 张欣, 等. 基于近、中红外光谱的热毒宁注射液制剂过程投料和二次热处理工序快速检测方法研究 [J]. 中草药, 2022, 53(21): 6706-6715.  
[17] 第五鹏瑶, 卞希慧, 王姿方, 等. 光谱预处理方法选择研究 [J]. 光谱学与光谱分析, 2019, 39(9): 2800-2806.  
[18] 周孟然, 孙磊, 卞凯, 等. iPLS 波段筛选方法在食用油品上快速检测研究 [J]. 激光杂志, 2020, 41(7): 13-17.  
[19] 陈露萍, 徐芳芳, 张欣, 等. 基于偏最小二乘法建立大株红景天片素片硬度近红外光谱预测模型 [J]. 中草药, 2023, 54(8): 2446-2452.  
[20] 朱田田, 王引权, 郭俊霞. 基于变异系数权重的中药质量综合评价模糊物元模型 [J]. 中药材, 2008, 31(1): 139-142.  
[21] 邵淑贤, 徐梦婷, 林燕萍, 等. 基于电子鼻与 HS-

- SPME-GC-MS 技术对不同产地黄观音乌龙茶香气差异分析 [J]. 食品科学, 2023, 44(4): 232-239.
- [22] Hearst M A, Dumais S T, Osuna E, *et al.* Support vector machines [J]. *IEEE Intell Syst Appl*, 1998, 13(4): 18-28.
- [23] 马永杰, 马芸婷, 陈佳辉. 结合卷积神经网络多层特征和支持向量机的车辆识别 [J]. 激光与光电子学进展, 2019, 56(14): 55-61.
- [24] Noble W S. What is a support vector machine? [J]. *Nat Biotechnol*, 2006, 24(12): 1565-1567.
- [25] 栾勇鹏, 陆春方, 乔永康, 等. 基于决策树的城市早期地下工程更新改造模式规划评价方法 [J]. 隧道建设, 2023, 43(9): 1473-1484.
- [26] 束美艳, 魏家玺, 周也莹, 等. 基于朴素贝叶斯分类的柑橘叶片溃疡病诊断 [J]. 浙江大学学报: 农业与生命科学版, 2021, 47(4): 429-438.
- [27] 丛树林. 基于贝叶斯分类器的工业过程故障诊断研究 [D]. 大连: 大连交通大学, 2021.
- [28] 胡晶. 基于朴素贝叶斯的新闻分类问题算法改进问题的研究 [J]. 电脑与信息技术, 2023, 31(2): 5-8.
- [29] 戢静红, 张振宇, 邓平. 一种基于随机森林的 LOS/NLOS 基站识别方法 [J]. 电讯技术, 2023, 63(10): 1596-1602.

[责任编辑 时圣明]