

近红外光谱技术结合化学计量学应用于川贝母真伪与规格的快速辨识研究

谢梦迪¹, 田亮玉¹, 桂新景^{2,3,4}, 施钧瀚^{2,3,4,5}, 王君明¹, 岳佑松¹, 张璐^{2,3,4,5}, 王艳丽^{2,3,4}, 姚静^{2,3,4,5}, 李学林^{1,2,3,4,5}, 刘瑞新^{1,2,3,4,5*}

1. 河南中医药大学, 河南 郑州 450008
2. 河南中医药大学第一附属医院 药学部, 河南 郑州 450000
3. 河南省中药饮片临床应用现代化工程研究中心, 河南 郑州 450000
4. 河南中医药大学 呼吸疾病中医药防治省部共建协同创新中心, 河南 郑州 450000
5. 河南省中药临床药学中医药重点实验室(建设单位), 河南 郑州 450000

摘要: 目的 探究近红外光谱技术应用于川贝母 *Fritillariae Cirrhosae Bulbus* 快速辨识的可行性。方法 收集 80 个川贝母待测样品(含炉贝、松贝、青贝及川贝母伪品等), 为获取上述样品的真伪和规格信息, 首先进行传统人工鉴别(M₁法)与《中国药典》法鉴别(M₂法), 并以 M₁法和 M₂法相结合的辨识结果作为标杆信息(Y); 近红外光谱仪采集待测样品粉末的光谱信息(X), 结合主成分分析-判别分析(principal component analysis-discriminant analysis, PCA-DA)、偏最小二乘-判别分析(partial least squares discriminant analysis, PLS-DA)、最小二乘支持向量机(least squares support vector machines, LS-SVM) 3 种化学计量学方法建立并优化真伪及商品规格辨识模型[Y=M(X)](M₃法)。结果 80 个样品均参与真伪辨识, 经留一法交互验证, 结果发现 3 种真伪辨识模型中均未分类样品, 准确率依次为 98.75%、98.75%、97.50%; 80 个样品中有 5 个因专家对规格的判定意见难以达成一致或样品本身原因等不参与建模, 最终以 75 个样品参与规格辨识, 模型准确率分别为 100.00%(无未分类样品)、100.00%(有 4 个未分类样品)、100.00%(无未分类样品)。结论 真伪辨识均以 PCA-DA、PLS-DA 为最终辨识模型, 商品规格分类以 PCA-DA、LS-SVM 为最优辨识模型。表明近红外光谱技术在川贝母真伪和规格质量快速辨识方面具有良好应用前景。

关键词: 近红外光谱技术; 川贝母; 快速辨识; 真伪; 规格; 主成分分析-判别分析

中图分类号: R286.2 文献标志码: A 文章编号: 0253-2670(2022)08-2490-09

DOI: 10.7501/j.issn.0253-2670.2022.08.026

Fast identification of authenticity and specification of *Fritillariae Cirrhosae Bulbus* by near infrared spectroscopy combined with chemometrics

XIE Meng-di¹, TIAN Liang-yu¹, GUI Xin-jing^{2,3,4}, SHI Jun-han^{2,3,4,5}, WANG Jun-ming¹, YUE You-song¹, ZHANG Lu^{2,3,4,5}, WANG Yan-li¹, YAO Jing^{2,3,4,5}, LI Xue-lin^{1,2,3,4,5}, LIU Rui-xin^{1,2,3,4,5}

1. Henan University of Traditional Chinese Medicine, Zhengzhou 450008, China
2. Department of Pharmacy, The First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou 450000, China
3. Henan Engineering Research Center for Modernization of clinical application of Chinese Herbal Pieces, Zhengzhou 450000, China
4. Collaborative Innovation Center for Prevention and Treatment of Respiratory Diseases with Traditional Chinese Medicine, Henan University of Chinese Medicine, Zhengzhou 450000, China
5. Henan Provincial Key Laboratory of TCM Clinical Pharmacy, Zhengzhou 450000, China

Abstract: **Objective** To explore the feasibility of the application of near infrared spectroscopy in the rapid identification of *Fritillariae Cirrhosae Bulbus*. **Methods** A total of 80 samples of *Fritillariae Cirrhosae Bulbus* were collected. In order to obtain the authenticity and specification information of the samples, traditional manual identification (M₁ method) and pharmacopoeia identification (M₂ method) were carried out, and the identification results combined with M₁ method and M₂ method were used as

收稿日期: 2021-11-09

基金项目: 国家自然科学基金面上项目(81773892); 河南省卫生健康中青年学科带头人专项(HNSWJW-2020014); 河南省中医药拔尖人才培养项目(重点项目)(2019ZYBJ07); 河南省高层次人才特殊支持“中原千人计划”——“中原青年拔尖人才”项目(ZYQR201912158); 中管局基地专项(2018JDZX087)

作者简介: 谢梦迪, 在读硕士, 从事中药饮片临床应用现代化关键技术研究。E-mail: 2082176197@qq.com

***通信作者** 刘瑞新, 博士, 主任药师, 从事中药饮片临床应用现代化关键技术研究。Tel: (0371)66233562 E-mail: liuruixin7@163.com

benchmark information (Y). Near Infrared Spectroscopy to collect spectral information of the sample powder (X), combined with principal component analysis, discriminant analysis (PCA-DA), partial least squares discriminant analysis (PLS-DA), the least squares support vector machines (LS-SVM) three chemometrics methods to establish and optimize the authenticity and commodity specification identification model [$Y=M(X)$] (M_3) **Results** All the 80 samples participated in the identification of authenticity and falsification. The results showed that there were no unclassified samples in the three identification models, and the accuracy rates were 98.75%, 98.75% and 83.75%, respectively. Among the 80 samples, five samples did not participate in the modeling due to the difficulty in reaching an agreement on the specification by experts or due to the sample itself. Finally, 75 samples were selected to participate in the specification identification, and the model accuracy rates were 100.00% (no unclassified samples), 100.00% (four unclassified samples) and 100.00% (no unclassified sample), respectively. **Conclusion** PCA-DA and PLS-DA were used as the final identification models for authenticity identification, and PCA-DA and LS-SVM were used as the optimal identification model for commodity specification classification. Near infrared spectroscopy has a good application prospect in the rapid identification of the authenticity and specification of *Fritillariae Cirrhosae Bulbus*.

Key words: near infrared spectroscopy technology; *Fritillariae Cirrhosae Bulbus*; fast identification; authenticity; specifications; principal component analysis-discriminant analysis

川贝母为百合科植物川贝母 *Fritillaria cirrhosa* D. Don、暗紫贝母 *F. unibracteata* Hsiao et K. C. Hsia、甘肃贝母 *F. przewalskii* Maxim.、梭砂贝母 *F. delavayi* Franch.、太白贝母 *F. taipaiensis* P. Y. Li 或瓦布贝母 *F. unibracteata* Hsiao et K. C. Hsiavar. Wabuensis (S. Y. Tanget S. C. Yue) Z. D. Liu, S. Wanget S. C. Chen 的干燥鳞茎, 按性状不同分别习称“松贝”“青贝”“炉贝”和“栽培品”。川贝母始载于《神农本草经》, 列为中品, 具有润肺散结、止咳化痰等功效, 可用于治疗虚劳咳嗽、吐痰咯血等症, 有“止咳圣药”之称^[1]。由于川贝母的药用价值高、疗效显著、市场需求量大, 过度的滥采滥挖使川贝母濒临灭绝^[2], 资源量急剧下降, 导致川贝母的价格日益昂贵, 市场上出现了许多仿伪品。因此, 鉴别川贝母的仿伪品、掺杂状况^[3]及商品规格分类对保证中药饮片质量确保临床疗效至关重要。传统的中药质量评价^[4]主要为经验鉴别, 即根据外观性状特征来判断真伪^[5]优劣。但传统人工辨识方法存在个体差异, 如感官灵敏度不同、主观性强、易于疲劳等问题, 对辨识结果造成影响; 随着现代技术发展, 出现了越来越多的鉴别方法, 比如薄层鉴别^[6]、理化鉴别^[7]、分子鉴定^[8]、荧光鉴别^[9-10]、高效液相鉴别^[11]等方法, 但这些方法存在耗时长, 效率低, 价格昂贵等缺点。如何快速^[12]准确鉴别川贝母规格及真伪是现阶段需解决的问题。

近红外光谱技术 (near infrared spectroscopy, NIRS) ^[13]系通过测定物质在近红外光谱区 (按波数计为 12 800~4 000 cm^{-1}) 的特征光谱并利用化学计量学方法^[14]提取相关信息, 对物质进行定性、定量

分析^[15]的一种光谱分析技术, 具有快速、准确、对样品无破坏的检测特性。该技术^[16]广泛应用于食品微生物检测^[17]、肉类新鲜度评估^[18]、不同年份大米的产地溯源检测^[19], 烟草中总糖、还原糖、总烟碱的含量测定^[20], 另外在中药及复杂样品定量分析^[21-24]、定性鉴别^[25-29]等方面也有较多的应用。已有学者基于便携式声光可调滤光器 (AOTF)-近红外光谱技术通过主成分分析法 (principal component analysis, PCA) 分别对贝母类药材、川贝母类药材及商品“松贝”与伪品进行鉴别^[30], 该文献采用近红外技术结合化学计量学^[31]对川贝母进行鉴别, 具有一定的学术价值及鉴别意义。然而该文献用于分析的部分样品数相对较少, 同时仅运用简单无指导的 PCA, 并且未对川贝母中青贝、炉贝规格进行分类。因此, 为了建立更加准确、完善的模型, 本研究收集 80 个川贝母待测样品, 拟采用 NIRS 结合多种有指导的化学计量学方法进行分析, 从样品的整体信息进行判别, 建立并优化“川贝母”的 NIRS 真伪及商品规格辨识模型, 以实现川贝母样品的快速、准确鉴别分析, 为中药饮片的快速鉴别提供参考。

1 仪器与材料

Nicolet 6700 型傅里叶变换近红外光谱仪, 配 InGaAs 检测器, Opus5.5 分析软件, TQ analys 软件; 药典筛 (四号筛, 0.25 μm 孔径, 浙江上虞市道墟五四仪器厂); LC-20A 型高效液相色谱仪 (日本岛津公司); 2000ES 型蒸发光散射检测器 (美国奥泰公司); DM1000 型电子显微镜 (Leica 公司); 电子游标卡尺 (精密度 0.01 mm, GB/T21389, 成都量具工具集团有限公司); BCA2248-CW 型 1/1

万电子天平 (Sartorius 公司), CP225D 型 1/10 万电子分析天平 (Sartorius 公司); 高速万能粉碎机 (FW-100, 60~180 目, 北京科伟永兴仪器有限公司); 电热鼓风干燥箱 (GZX-9146MBE, 上海博迅实业有限公司医疗设备厂); HK250 型超声波清洗机 (上海科导超声仪器有限公司)。

对照品贝母素甲 (质量分数 96.2%, 批号 110750-201612, 中国食品药品检定研究院), 贝母素乙 (质量分数≥98%, 批号 B20081, 上海源叶生物科技有限公司), 平贝母对照药材 (批号 Y29A9H69015, 上海源叶生物有限公司), 西贝母碱 (质量分数 96%, 批

号 110767-201710, 中国食品药品检定研究院), 二乙胺 (批号 20121006, 天津市永大化学试剂开发中心产品), 二氯甲烷 (批号 20170812, 天津市恒兴化学试剂制造有限公司), 甲醇 (批号 20180301, 烟台市双双化工有限公司), 氨水 (批号 20161101, 烟台市双双化工有限公司), 醋酸乙酯 (批号 20150324, 天津市永大化学试剂有限公司), 稀甘油 (国药准字 H20073910, 南昌白云药业有限公司), 娃哈哈纯净水, 色谱级乙腈, 三氯甲烷、浓氨水等均为分析纯。实验用 80 批川贝母待测样品分别购自于郑州中医院、药材市场、医药公司等, 产地及批号信息见表 1。

表 1 川贝母样品购买信息 (含 20 批平贝母样品)

Table 1 Purchase information of *Fritillariae Cirrhosae Bulbus* samples (contains 20 batches of *Fritillariae Ussuriensis Bulbus* samples)

编号	产地	批号	购买单位	编号	产地	批号	购买单位
1	四川	1712073	河南中医药大学第一附属医院	41	四川	—	三七西洋参专卖
2	四川	1712143	河南中医药大学第一附属医院	42	四川	171101	郑州市中医院 (小包装)
3	四川	1801201	河南中医药大学第一附属医院	43	四川	180201	郑州市中医院 (散装)
4	四川	1801211	河南中医药大学第一附属医院	44	四川	1712043	河中一医药经营有限公司
5	四川	1803261	河南中医药大学第一附属医院	45	四川	1801141	河中一医药经营有限公司
6	四川	180408	河中一医药经营有限公司	46	四川	1710273	河中一医药经营有限公司
7	四川	180409	河中一医药经营有限公司	47	四川	180110	河中一医药经营有限公司
8	四川	180410	河中一医药经营有限公司	48	四川	180111	河中一医药经营有限公司
9	四川	180411	河中一医药经营有限公司	49	四川	180112	河中一医药经营有限公司
10	四川	180412	河中一医药经营有限公司	50	四川	180113	河中一医药经营有限公司
11	四川	180413	河中一医药经营有限公司	51	四川	180114	河中一医药经营有限公司
12	四川	180414	河中一医药经营有限公司	52	四川	180115	河中一医药经营有限公司
13	四川	180415	河中一医药经营有限公司	53	四川	180116	河中一医药经营有限公司
14	四川	180416	河中一医药经营有限公司	54	四川	180117	河中一医药经营有限公司
15	四川	180417	河中一医药经营有限公司	55	四川	180118	河中一医药经营有限公司
16	四川	180418	河中一医药经营有限公司	56	四川	180119	河中一医药经营有限公司
17	四川	180419	河中一医药经营有限公司	57	四川	180120	河中一医药经营有限公司
18	四川	180420	河中一医药经营有限公司	58	四川	180121	河中一医药经营有限公司
19	四川	180421	河中一医药经营有限公司	59	四川	180122	河中一医药经营有限公司
20	四川	180422	河中一医药经营有限公司	60	四川	180123	河中一医药经营有限公司
21	四川	—	三七参茸堂	61	浙江	—	三七参茸行专卖店
22	四川	—	三七西洋参专卖	62	四川	—	三七参茸堂
23	四川	—	郑州市鲲鹏药材行	63	贵州	—	三七西洋参专卖
24	四川	—	郑州市鲲鹏药材行	64	四川	—	昌盛药材行
25	四川	1701073	河中一医药经营有限公司	65	四川	—	昌盛药材行
26	四川	1707233	河中一医药经营有限公司	66	黑龙江	180301	河中一医药经营有限公司
27	四川	1711263	河中一医药经营有限公司	67	黑龙江	180302	河中一医药经营有限公司
28	四川	1803051	河中一医药经营有限公司	68	黑龙江	180303	河中一医药经营有限公司
29	四川	1712271	河中一医药经营有限公司	69	黑龙江	180304	河中一医药经营有限公司
30	四川	180125	河中一医药经营有限公司	70	黑龙江	180305	河中一医药经营有限公司
31	四川	180126	河中一医药经营有限公司	71	黑龙江	180306	河中一医药经营有限公司
32	四川	180201	河中一医药经营有限公司	72	黑龙江	180307	河中一医药经营有限公司
33	四川	180202	河中一医药经营有限公司	73	黑龙江	180308	河中一医药经营有限公司
34	四川	180203	河中一医药经营有限公司	74	黑龙江	180309	河中一医药经营有限公司
35	四川	180204	河中一医药经营有限公司	75	黑龙江	180310	河中一医药经营有限公司
36	四川	180205	河中一医药经营有限公司	76	黑龙江	180311	河中一医药经营有限公司
37	四川	180206	河中一医药经营有限公司	77	吉林	1710241	河中一医药经营有限公司
38	四川	180207	河中一医药经营有限公司	78	吉林	1711261	河中一医药经营有限公司
39	四川	180208	河中一医药经营有限公司	79	黑龙江	180313	河中一医药经营有限公司
40	四川	180209	河中一医药经营有限公司	80	黑龙江	180312	河中一医药经营有限公司

2 方法与结果

2.1 标杆信息的确定

M₁ 法为传统人工鉴别方法, 聘请在中药鉴定研究领域具有丰富经验的专家 8 名, 参照《中国药典》2020 年版^[6]一部质量标准外观性状规定, 将样品编号后随机放置, 由专家依据气味、颜色、性状、味道、质地等外观指标的差异进行真伪鉴别及规格辨识。M₂ 法为经典的药典方法, 参照《中国药典》2020 年版一部^[6], 其中含量测定方法是基于文献方法^[32]改进的高效液相色谱蒸发光散射检测法 (HPLC-ELSD),

选择 3 位操作人员进行实验, 结果取平均值, 色谱条件为 Waters XBridge shield RP₁₈ 色谱柱, 流动相乙腈 (A) -0.1% 二乙胺水溶液 (B) 梯度洗脱 (0~15 min, 50%~23% A; 15~16 min, 23%~50% A; 16~28 min, 50% A), 体积流量 1.0 mL/min, 柱温 25 °C。ELSD 参数为漂移管温度 108 °C, 载气体积流量 2.8 L/min, 进样量 10 μL。当 M₁ 法与 M₂ 法结果出现差别时, 将对样品进行复检确定最终结果。若结果仍不统一时, 考虑到 M₂ 法比 M₁ 法更具有客观性和系统性, 考虑以药典结果作为最终结果。结果见表 2。

表 2 川贝母样品信息及 2 种辨识方法的结果

Table 2 Sample of *Fritillariae Cirrhosae Bulbus* and two identification results

编号	真伪辨识			商品规格辨识			编号	真伪辨识			商品规格辨识		
	M ₁	M ₂	标杆信息 Y	M ₁	M ₂	标杆信息 Y		M ₁	M ₂	标杆信息 Y	M ₁	M ₂	标杆信息 Y
1	1	1	1	1	1	1	41	—	1	1	1	3	3
2	1	1	1	1	1	1	42	171101	0	1	1	0	0
3	1	1	1	1	1	1	43	180201	0	2	2	0	0
4	1	1	1	1	1	1	44	1712043	1	1	1	3	3
5	1	1	1	1	1	1	45	1801141	1	1	1	3	3
6	1	1	1	1	1	1	46	1710273	1	1	1	3	3
7	1	1	1	1	1	1	47	180110	1	1	1	3	3
8	1	1	1	1	1	1	48	180111	1	1	1	3	3
9	1	1	1	1	1	1	49	180112	1	1	1	3	3
10	1	1	1	1	1	1	50	180113	1	1	1	3	3
11	1	1	1	1	1	1	51	180114	1	1	1	3	3
12	1	1	1	1	1	1	52	180115	1	1	1	3	3
13	1	1	1	1	1	1	53	180116	1	1	1	3	3
14	1	1	1	1	1	1	54	180117	1	1	1	3	3
15	1	1	1	1	1	1	55	180118	1	1	1	3	3
16	1	1	1	1	1	1	56	180119	1	1	1	3	3
17	1	1	1	1	1	1	57	180120	1	1	1	3	3
18	1	1	1	1	1	1	58	180121	1	1	1	3	3
19	1	1	1	1	1	1	59	180122	1	1	1	3	3
20	1	1	1	1	1	1	60	180123	1	1	1	3	3
21	1	1	1	2	2	2	61	—	2	2	2	4	4
22	1	1	1	1	2	2	62	—	2	2	2	4	4
23	2	2	2	0	0	0	63	—	2	2	2	4	4
24	2	2	2	0	0	0	64	—	2	2	2	4	4
25	1	1	1	2	2	2	65	—	2	2	2	4	4
26	1	1	1	2	2	2	66	180301	2	2	2	4	4
27	1	1	1	2	2	2	67	180302	2	2	2	4	4
28	1	1	1	2	2	2	68	180303	2	2	2	4	4
29	1	1	1	2	2	2	69	180304	2	2	2	4	4
30	1	1	1	2	2	2	70	180305	2	2	2	4	4
31	1	1	1	2	2	2	71	180306	2	2	2	4	4
32	1	1	1	2	2	2	72	180307	2	2	2	4	4
33	1	1	1	2	2	2	73	180308	2	2	2	4	4
34	1	1	1	2	2	2	74	180309	2	2	2	4	4
35	1	1	1	2	2	2	75	180310	2	2	2	4	4
36	1	1	1	2	2	2	76	180311	2	2	2	4	4
37	180206	1	1	1	2	2	77	1710241	2	2	2	4	4
38	180207	1	1	1	2	2	78	1711261	2	2	2	4	4
39	180208	1	1	1	2	2	79	180313	2	2	2	0	0
40	180209	1	1	1	2	2	80	180312	2	2	2	4	4

真伪鉴别: 0-未分类 1-真品 2-伪品; 商品规格: 0-未分类 1-炉贝 2-松贝 3-青贝 4-平贝 —不可得

Identification of authenticity: 0-unclassified 1-authentic 2-counterfeit; specifications: 0-Unclassified 1-furnace Fritillary 2-pine Fritillary 3-green fritillary 4-flat fritillary —not available

2.2 傅里叶变换近红外光谱信息的采集

将 1~80 号贝母样品放置于 60 °C 烘箱中烘干 6 h, 粉碎, 过筛 (四号药典筛), 粉末置于提前标记好的自封袋内, 置干燥器中, 待测。每个样品重复

装样置石英样品池中, 每次扫描前摇荡石英杯, 使样品均匀平整, 平行扫描 3 次, 见图 1。取平均值作为其近红外光谱基础数据, 每个样品导出 CSV 文件, 每个 CSV 文件包含 2 列共 2076 行数据。

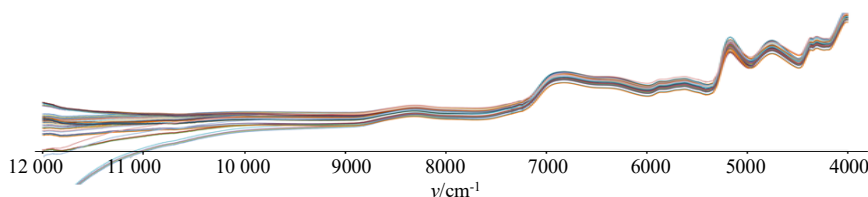


图 1 近红外采集 80 个贝母样品光谱图

Fig. 1 Spectrogram of 80 fritillaria samples collected by NIRS

2.3 真伪模型的建立及优化

结合 M_1 法和 M_2 法所得最终结果作为标杆信息 (Y), 如表 2 所示。利用 NIRS 对 80 批川贝母待测样品进行检测, 提取其光学数据并借助化学计量学方法建立适宜的辨识模型, 模型以近红外采集的 2075 个色度信息值为自变量矩阵 X (80×2075), 建立 X 和 Y 之间的关系 $Y=M(X)$ 模型。利用 MATLAB 2016b 软件建模, 分别建立 80 个样品的 PCA-DA、PLS-DA、LS-SVM 真伪辨识模型 (M_3 法) [33-34]。

2.3.1 PCA-DA 辨识模型及交互验证结果 参数设置: 变量选择 7 个, 识别模式选择线性, 模型最优正判率为 98.75%。前 3 个主成分之和达 75% 以上, 可解释原变量大部分信息。PCA-DA 辨识结果示: 23 个伪品中 79 号被错分为正品, 不存在未分类样品。见图 2。

2.3.2 PLS-DA 辨识模型及交互验证结果 参数设置: 变量选择 6、7、8、9、10 时正判率最高, 识别模式选择线性; 模型最优正判率为 98.75%。前 4 个主成分之和达 75% 以上, 可解释原变量大部分信息。PLS-DA 辨识结果示, 23 个伪品中 79 号被错分为正品, 不存在未分类样品。见图 3。

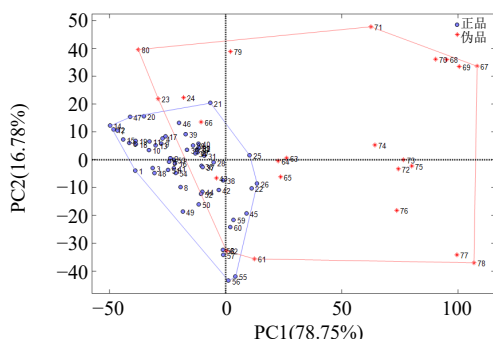


图 2 PCA-DA 模型主成分得分图

Fig. 2 Principal component score of PCA-DA model

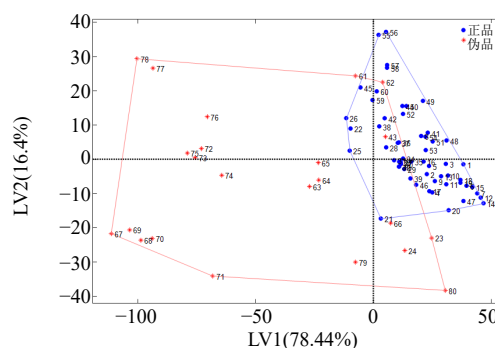


图 3 PLS-DA 模型主成分得分图

Fig. 3 Principal component scores of PLS-DA model

2.3.3 LS-SVM 辨识模型及交互验证结果 通过 LS-SVM R2009b-R2011a 载入矩阵 X (80×2075)、矩阵 Y (80×1), 采用 3 种不同函数代入程序进行运算, 结果发现: 线性核函数 (lin-kernel) 正判率为 97.50%, 径向基核函数 (RBF-kernel) 正判率为 71.25%, 多核函数 (poly-kernel) 正判率为 86.25%。因此, 以正判率最高的线性核函数为最终运行函数, 运行结果: 23 个伪品有 2 个被错分为正品, 不存在未分类样品, 模型整体正判率为 97.50%。

2.4 商品规格模型的建立及优化

结合 M_1 法和 M_2 法所得最终结果作为标杆信息 (Y), 如表 2 所示。利用 NIRS 对 75 批 (剔除 5 个未分类样品) 川贝母待测样品进行检测, 提取其光学数据并借助化学计量学方法建立适宜的辨识模型, 模型以近红外采集的 2075 个色度信息值为自变量矩阵 X (75×2075), 标杆信息为因变量矩阵 Y (80×1) 建立 X 和 Y 之间的关系 $Y=M(X)$ 模型。利用 MATLAB 2016b 软件, 分别建立 75 个样品的 PCA-DA、PLS-DA、LS-SVM 商品规格辨识模型 (M_3 法), 方法同上。

2.4.1 PCA-DA 辨识模型及交互验证结果 参数设置：变量选择 16、17、18、19、20 个时正判率最高，识别模式选择线性；模型最优正判率为 100.00%。前 3 个主成分之和达 75%以上，可解释原变量大部分信息。PCA-DA 辨识结果显示：20 个炉贝、18 个松贝、18 个青贝、19 个平贝样品分类均正确；不存在未分类的样品。见图 4。

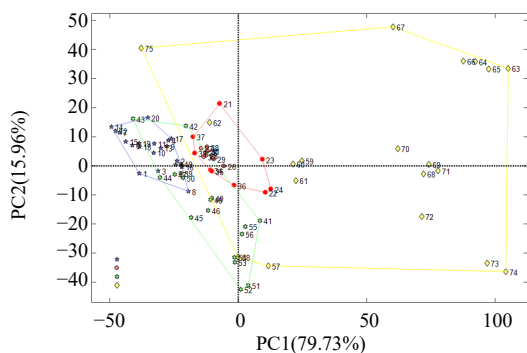


图 4 PCA-DA 模型得分图

Fig. 4 Principal component scores of PCA-DA model

2.4.2 PLS-DA 辨识模型及交互验证结果 参数设置：变量选择 9 个，识别模式选择线性；模型最优正判率为 100.00%。前 4 个主成分之和达 75%以上，可解释原变量大部分信息。PLS-DA 辨识结果显示：20 个炉贝分类正确；26、46、66、80 号未被分类。见图 5。

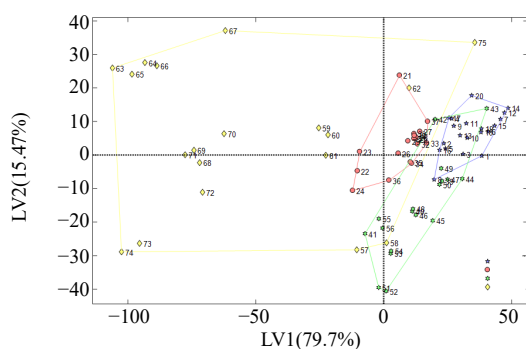


图 5 PLS-DA 模型主成分得分图

Fig. 5 Principal component scores of PLS-DA model

2.4.3 LS-SVM 辨识模型及交互验证结果 通过 LS-SVM R2009b-R2011a，载入矩阵 $X(80 \times 2075)$ 、矩阵 $Y(80 \times 1)$ ，同“2.3.2”项的方法一致，结果发现，线性核函数 (lin-kernel) 正判率为 100.00%，径向基核函数 (RBF-kernel) 正判率为 93.33%。因此，最终以线性核函数为最优函数，模型整体正判率为 100.00%。

2.5 M_3 方法最优辨识结果

M_3 法以正判率最高的模型结果为最终判别结果，真伪辨识模型正判率最高的 PLS-DA 与 PCA-DA 分类结果相同，不作分开表述，具体结果见表 3。

表 3 M_3 辨识方法的结果

Table 3 M_3 identification results

编号	真伪辨识	商品规格辨识
1~20	真品	炉贝
21~22	真品	松贝
23~24	伪品	未分类
25~40	真品	松贝
41	真品	青贝
42	真品	未分类
43	伪品	未分类
44	真品	青贝
45	真品	青贝
46~59	真品	青贝
60	真品	平贝
61~78	真品	平贝
79	真品	未分类
80	真品	炉贝

2.6 M_3 法的时效性计算

参照参考文献的方法^[34]进行计算，结果发现： M_3 法耗费的时间远少于 M_2 法 ($P < 0.01$)。

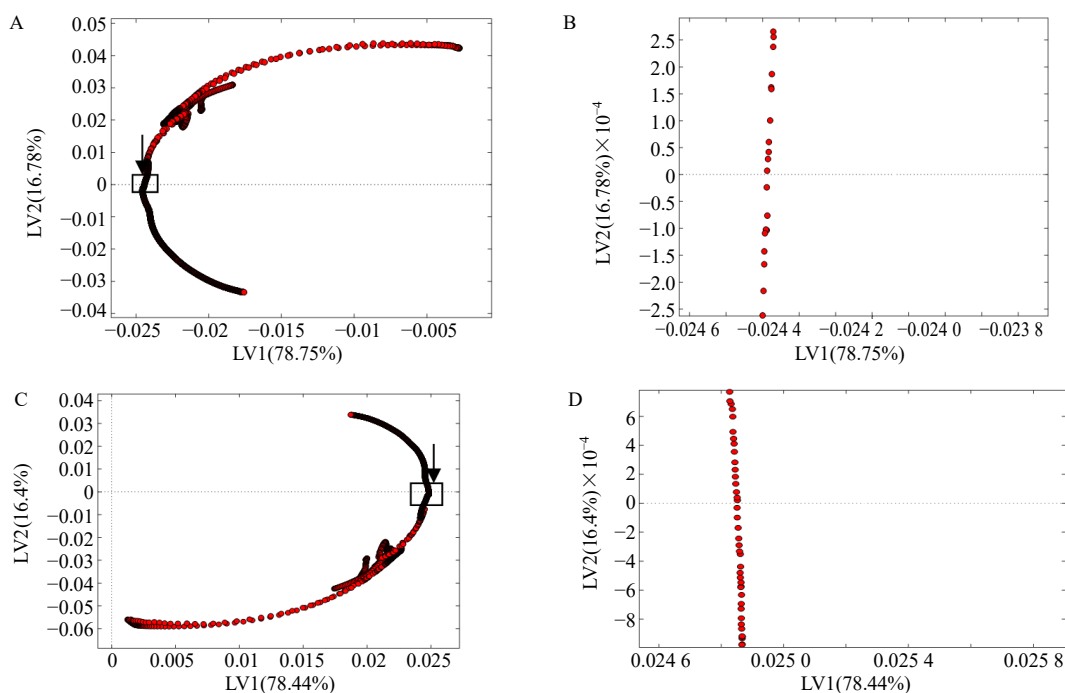
2.7 近红外传感器信息值 (变量) 对模型贡献度分析

真伪与多规格模型分类优选结果：真伪模型选择 PLS-DA 及 PCA-DA 辨识模型，多规格模型选择 PCA-DA、LS-SVM 辨识模型。图 6-B、D 图中编号仅代表波长对应的序号，因变量数太多不能完整显现出来最优波长数。图 7-B 图贡献率最大的波长范围是 $7\ 960.837 \sim 8\ 045.691\ \text{cm}^{-1}$ ，D 图贡献率最大的波长范围是 $8\ 045.691 \sim 8\ 091.975\ \text{cm}^{-1}$ 。图 7-B 中编号仅代表波长对应的序号，其对应最优波长范围是 $8\ 153.687 \sim 8\ 296.396\ \text{cm}^{-1}$ 。

3 讨论

3.1 被剔除样品的原因

80 个川贝母及其近似样品中经专家及药典鉴定后显示，23、24、43、61~80 号为伪品；商品规格鉴别结果显示，23、24、42、43、79 号共计 5 个样品因专家意见不一致，因此被剔除，剩余 75 个样品按商品规格分为炉贝、松贝、青贝、平贝。综上，

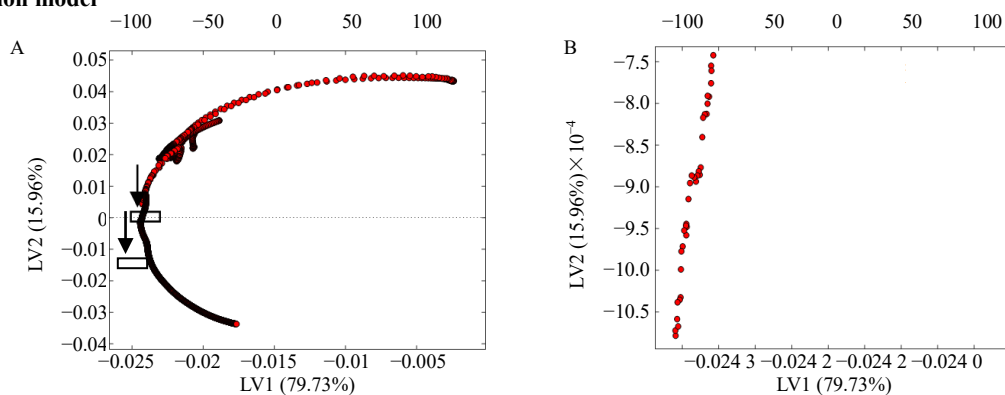


A-PCA-DA 真伪识别模型变量载荷图 B-PCA-DA 真伪识别模型变量载荷图最优波长放大图 C-PLS-DA 真伪识别模型变量载荷图 D-PLS-DA 真伪识别模型变量载荷图最优波长放大图

A-PCA-DA True and False identification model Variable Loading Diagram B-PCA-DA True and false identification Model Variable Loading Diagram Optimal Wavelength Amplification Diagram C-PLS-DA True and false identification Model Variable Loading Diagram D-PLS-DA True and false identification Model Variable loading Diagram Optimal wavelength amplification diagram

图 6 真伪识别模型 PLS-DA、PCA-DA 贡献率最大的波段在潜变量上的载荷图

Fig. 6 PLS-DA and PCA-DA load diagram of the band with the largest contribution rate on latent variables of authenticity identification model



A-PCA-DA 规格识别模型变量载荷图 B-PCA-DA 规格识别模型变量载荷图最优波长放大图

A-variable load diagram of PCA-DA variety identification model B-optimal wavelength amplification diagram of variable-load diagram of PCA-DA variety identification model

图 7 规格识别模型的 PCA-DA 贡献率最大的波段在潜变量上的载荷图

Fig. 7 PCA-DA load diagram of the bands with the largest contribution rate on latent variables of variety identification model

22 号样品有 2 位专家认为掺杂少量的青贝，其余六位专家判定为炉贝，药典方法判定为松贝；23、24 号既为伪品，且在多规格分类中因外观形态与松贝

存在差异，不能准确的被归结为哪一类，因此未被分类；42、43 号人工鉴别中有 3~4 位专家不能给出鉴别结果，不符合人数 $\geq 3/4$ ，但药典结果鉴别出

42 为正品、43 为伪品,经过多次复检未能确定结果,最终以药典方法的结果作为标准。42 号虽被判为正品,但因外观形态与青贝不一致导致多个专家意见不同而未进行分类;43 号被判为伪品,且在因外观形态与青贝存在差异,不能准确的被归结为哪一类,因此未被分类;多规格分类中 79 号样品因 1 名专家认为是炉贝,2 名专家不确定是浙贝还是湖北贝母,1 名专家认为不属于平贝也不属于松贝,最终导致无法正确分类,标杆信息 Y 值无法确定,因此 79 号样品被剔除。

3.2 判错样品原因分析

真伪辨识 PCA-DA、PLS-DA 模型均认为 79 号为正品, M_1 法 4 个专家认为 79 号既不属于川贝母也不属于平贝母,说明待测样品中含有掺伪样品,使建立的模型准确率较低,导致 79 号样品判定错误。

3.3 未分类样品分析

PCA-DA、PLS-DA 模型是有监督的判别分析方法^[35-36],相对于无监督的 PCA,当样本组间差异小时,它也能实现良好的区分。规格辨识 PLS-DA 模型中正判率为 100%,26、46、66、80 号样品未被分类,但 M_1 法、 M_2 法均能够明确将其分类,表明 PLS-DA 模型不是规格辨识的最优模型。

3.4 近红外光谱法的不足之处

本研究运用傅里叶近红外光谱技术结合多种有指导的化学计量学方法,对 80 个川贝母待测样品分析,优选的模型正判率与药典方法相比无显著性差异 ($P < 0.01$),相对于人工鉴别大大缩短了时间,节省大量人力物力,实现了对复杂体系快速^[37]、无损定性分析。该技术同时也有不足之处:(1)分析需要建立模型,对模型建立要求较高;(2)模型传递技术尚不成熟,如果能建立成熟的模型传递^[38],使在一台仪器上建立的定性或定量校正模型可靠地移植到其他相同或类似的仪器上使用,就能减少建模所需的时间和费用。当然,随着光谱技术计算机技术与化学计量学相结合,使模型建立难度下降,模型准确度增强、模型传递技术更加成熟。今后,近红外光谱技术在中药质量控制领域将具有更加广阔的前景。

3.5 真伪与多规格分类正判率比较

川贝母真伪模型正判率大小为 PLS-DA=PCA-DA>LS-SVM,最终优选 PCA-DA、PLS-DA 为真伪最优判别模型;多规格分类模型正

判率大小为 PCA-DA=PLS-DA=LS-SVM,但因 PLS-DA 模型中未被分类数太多,故最终优选 PCA-DA、LS-SVM 为最优判别模型。

综合比较,近红外辨识正判率与传统人工经验鉴别无显著性差异,近红外辨识不如传统经验辨识快,但较药典检测时间极显著缩短。综上所述,近红外技术结合多元统计方法快速准确的辨识川贝母真伪和商品规格,可为中药品质评价及质量鉴别提供一种新思路和新技术。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 李瑞琦,吴翠,徐靓,等.基于传统性状客观化分析的川贝母质量评价[J].中国药理学杂志,2020,55(1):14-18.
- [2] 仰铁锤,谢慧敏,谢慧淦,等.聚合酶链式反应-限制性内切酶多态法检查川贝母的掺伪情况[J].华西药理学杂志,2020,35(3):265-269.
- [3] Rodionova O Y, Fernández P J A, Baeten V, et al. Chemometric non-targeted analysis for detection of soybean meal adulteration by near infrared spectroscopy [J]. *Food Control*, 2021, 119: 107459.
- [4] 崔治家,马艳珠,张小荣,等.川贝母化学成分和药理作用研究进展及质量标志物的预测分析[J].中草药,2021,52(9):2768-2784.
- [5] 兰青阔,赵新,陈锐,等.基于 LAMP 的中药材川贝母真伪鉴别方法的建立[J].药物分析杂志,2019,39(3):551-556.
- [6] 中国药典[S].一部.2020:126.
- [7] 翟映红,于国强,杨永茂,等.从川贝母中检出平贝母的理化鉴别方法[J].华西药理学杂志,2017,32(6):633-635.
- [8] 时圣明,潘明佳,王洁,等.分子鉴定技术在中药中的应用[J].中草药,2016,47(17):3121-3126.
- [9] 罗达龙,黄林杰,黄琳.实时荧光定量 PCR 对川贝母的鉴别应用[J].中国药师,2016,19(6):1068-1070.
- [10] 吴雪莉,王毅,赵筱萍.基于荧光探针的消渴安中 DPP-4 抑制剂的筛选与鉴定[J].中国中药杂志,2016,41(7):1241-1245.
- [11] 刘晶,王曙,辛贵忠,等.HPLC-ELSD 测定太白贝母和瓦布贝母中贝母辛的含量[J].中国药理学杂志,2010,45(13):1032-1034.
- [12] 杨靖雯,高秀强,田清存,等.10 种中药材快速经验鉴别方法总结[J].中药材,2018,41(12):2792-2794.
- [13] 中国药典[S].四部.2020:29.
- [14] Arndt M, Drees A, Ahlers C, et al. Determination of the geographical origin of walnuts (*Juglans regia* L.) using

- near-infrared spectroscopy and chemometrics [J]. *Foods*, 2020, 9(12): 1860.
- [15] Yan H, Pu Z J, Wang Y J, *et al.* Rapid qualitative identification and quantitative analysis of *Flos Mume* based on Fourier transform near infrared spectroscopy [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2021, 249: 119344.
- [16] Mishra P, Herrmann I, Angileri M. Improved prediction of potassium and nitrogen in dried bell pepper leaves with visible and near-infrared spectroscopy utilising wavelength selection techniques [J]. *Talanta*, 2021, 225: 121971.
- [17] 周冰谷, 花振新, 杨荣, 等. 近红外光谱技术在食品微生物检测中的应用饲料技术 [J]. 广东饲料, 2019, 28(8): 44-46.
- [18] 陶瑞, 史智佳, 贡慧, 等. 傅里叶变换近红外光谱技术快速检测金枪鱼新鲜度 [J]. 肉类研究, 2017, 31(4): 43-49.
- [19] 宋雪健, 钱丽丽, 张东杰, 等. 基于漫反射傅里叶变换近红外光谱技术对不同年份的大米产地溯源检测 [J]. 食品科学, 2017, 38(18): 286-291.
- [20] 蒋宏霖, 刘会杰, 王学杰, 等. 近红外光谱技术在烟叶化学分析中的应用 [J]. 科技与创新, 2019(18): 153-155.
- [21] 陈佳乐, 金叶, 陈红英, 等. 川芎药材的近红外多指标快速质量评价 [J]. 中草药, 2016, 47(6): 1004-1009.
- [22] 雷晓晴, 李耿, 王秀丽, 等. 基于近红外光谱法快速测定丹参中 5 种成分模型的建立 [J]. 中草药, 2018, 49(11): 2653-2661.
- [23] 战皓, 吴宏伟, 张东, 等. 近红外光谱法测定不同产地黄芪中毛蕊异黄酮葡萄糖苷和黄芪甲苷含量 [J]. 光谱学与光谱分析, 2017, 37(5): 1391-1396.
- [24] 马越, 刘永强, 门宇凤, 等. 近红外光谱法快速测定芍药根中芍药苷和芍药内酯苷 [J]. 中草药, 2018, 49(3): 700-707.
- [25] 闫珂巍, 王福, 梅国荣, 等. 基于近红外光谱技术快速定性鉴别广陈皮模型的建立 [J]. 中草药, 2015, 46(20): 3096-3099.
- [26] 张敏, 韩正洲, 严萍, 等. 基于近红外漫反射光谱的鸡血藤药材定性分析 [J]. 世界科学技术—中医药现代化, 2014, 16(7): 1491-1496.
- [27] 徐亚, 李慧, 冯学峰, 等. 基于显微和近红外方法定性鉴别中药材威灵仙不同基原品种的比较研究 [J]. 中国中药杂志, 2015, 40(7): 1278-1282.
- [28] 余驰, 姜红, 刘爱萍. 近红外漫反射光谱法建立厚朴药材的定性模型 [J]. 药物分析杂志, 2009, 29(4): 656-658.
- [29] 胡钢亮, 陈瑞珍, 程柯, 等. 近红外漫反射光谱快速检测川贝母中浙贝母的掺入量 [J]. 药物分析杂志, 2005, 25(2): 150-152.
- [30] 杨复森, 武卫红, 王宁, 等. 基于 AOTF-近红外光谱技术的川贝母药材即时快速鉴别研究 [J]. 中成药, 2013, 35(1): 135-140.
- [31] Hua H M, Yi Y, Li H L, *et al.* Monitoring the extraction process of acidic polysaccharides in *Poria cocos* by near infrared spectroscopy combined with chemometrics [J]. *Infrared Phys Technol*, 2021, 113: 103613.
- [32] 雷艳辉, 李会军, 李萍. 川贝母药材生物碱成分 HPLC-ELSD 特征图谱的研究 [J]. 中成药, 2014, 36(7): 1477-1481.
- [33] 应雨棋, 张洋洋, 虞立, 等. 基于 LS-SVM 模型的丹参有效成分超声提取工艺优化 [J]. 中药材, 2020, 43(4): 936-941.
- [34] 刘瑞新, 郝小佳, 张慧杰, 等. 基于电子眼技术的中药川贝母真伪及规格的快速辨识研究 [J]. 中国中药杂志, 2020, 45(14): 3441-3451.
- [35] Ramírez-Elías M G, Guevara E, Zamora-Pedraza C, *et al.* Assessment of mezcal aging combining Raman spectroscopy and multivariate analysis techniques [J]. *Biomed Spectrosc Imaging*, 2017, 6(1/2): 75-81.
- [36] Wong K H, Razmovski-Naumovski V, Li K M, *et al.* Differentiating *Puerariae Lobatae Radix* and *Puerariae Thomsonii Radix* using HPTLC coupled with multivariate classification analyses [J]. *J Pharm Biomed Anal*, 2014, 95: 11-19.
- [37] 雷晓晴, 王秀丽, 李耿, 等. 近红外光谱法快速测定当归中 7 种成分的含量 [J]. 中草药, 2019, 50(16): 3947-3954.
- [38] 梁晨. 近红外光谱多元校正模型传递方法的研究 [D]. 北京: 北京化工大学, 2016.

[责任编辑 时圣明]