

基于化学计量学和气相色谱法的艾叶及其混伪品蒙古蒿叶差异性特征成分和鉴别规律的研究*

张文静¹, 李海燕¹, 王晓伟¹, 王海波^{1**}, 李向阳¹,
李桂本¹, 张红伟¹, 耿怡玮¹, 杨元¹, 石岩^{2**}

(1. 河南省药品医疗器械检验院(河南省疫苗批签中心) 国家药品监督管理局中药材及饮片质量控制重点实验室, 郑州 450008;
2. 中国食品药品检定研究院, 北京 102629)

摘要 目的: 以气相色谱和化学计量学技术为化学分析和数据分析手段, 研究艾叶及其混伪品蒙古蒿叶的鉴别方法及化学成分分布规律。方法: 采用 Agilent HP-5 19091J(30 m × 0.32 mm, 0.25 μm) 色谱柱进行分离, 使用氢火焰离子检测器进行检测, 建立了气相色谱化学分析方法, 指认了其中 21 个化学成分, 使用该方法对 29 批样品进行测定。采用化学计量学相关技术对色谱数据进行相似度分析、相关性分析、聚类分析、主成分分析和正交偏最小二乘判别分析。结果: 色谱峰峰 20(母菊萹)、峰 3(桉油精)和峰 19[(1S, 8aα)-十氢-1, 4aβ-二甲基-7β-异丙烯基-1-萘酚]是艾叶与蒙古蒿叶的差异特征色谱峰(化学成分), 艾叶和蒙古蒿叶样品的峰 3 与峰 20 峰面积比值范围分别为 54.50 ~ 348.39 和 0.16 ~ 0.87, 峰 19 与峰 20 峰面积比值范围分别为 18.55 ~ 128.46 和 0.01 ~ 0.14, 正品与混伪品之间有巨大差异, 可用于艾叶与蒙古蒿叶的鉴别。结论: 本研究得出的艾叶与蒙古蒿叶的差异特征化学成分及其分布规律对艾叶及相关药品的研究和分析具有一定的参考意义。

关键词: 艾叶; 蒙古蒿叶; 混伪品; 气相色谱; 化学计量学; 数据分析; 数据预处理; 主成分分析; 正交偏最小二乘判别分析

中图分类号: R 917 文献标识码: A 文章编号: 0254-1793(2024)04-0649-14
doi: 10.16155/j.0254-1793.2024.04.12

Identification of Artemisiae Argyi Folium and Artemisiae Mongolica Folium based on gas chromatography and chemometric techniques*

ZHANG Wen-jing¹, LI Hai-yan¹, WANG Xiao-wei¹, WANG Hai-bo^{1**},
LI Xiang-yang¹, LI Gui-ben¹, ZHANG Hong-wei¹, GENG Yi-wei¹,
YANG Yuan¹, SHI Yan^{2**}

(1. Henan Institute for Drug and Medical Device Inspection(Henan Vaccine Issuance Center), NMPA Key Laboratory for Quality Control of Traditional Chinese Medicine (Chinese Materia Medica and Prepared Slices), Zhengzhou 450008, China; 2. National Institutes for Food and Drug Control, Beijing 102629, China)

Abstract Objective: To study and establish a method based on gas chromatography and chemometrics techniques

* 中国药品监管科学行动计划第二批重点项目(NMPAJGKX-2023-030); 河南省科技厅科技攻关项目(222102310110); 国家局药品监管科学体系建设重点项目-新技术新方法在中药质量控制中的应用(RS2024Z006-035)

** 通信作者 王海波 Tel:(0371)65566577; E-mail: haibowang99@163.com

石岩 Tel:(010)53852081; E-mail: shiyan@nifdc.org.cn

第一作者 Tel:(0371)65566039; E-mail: 878348987@qq.com

for distinguishing *Artemisiae Argyi Folium* and its adulterants *Artemisiae Mongolica Folium*. **Methods:** Gas chromatography method was established with Agilent HP-5 19091J (30 m × 0.32 mm, 0.25 μm) as chromatographic column, and hydrogen flame ion detector (FID) as detector. After the chemical composition of 21 chromatographic peaks in the chromatogram were identified, and the peak area data of the 21 chromatographic peaks in 29 batches of samples were determined. Similarity analysis, correlation analysis, cluster analysis, principal component analysis and orthogonal partial least squares - discriminant analysis were applied to analyze the chromatographic data. **Results:** The results of chemometric analysis indicated that tpeak 20 (chamazulene), peak 3 (1,8 - cineole) and peak 19 ((1S,8aα) - decahydro - 1,4aβ - dimethyl - 7β - isopropenyl - 1 - naphthol) were the differential characteristic chromatographic peaks between *Artemisiae Argyi Folium* and its adulterants *Artemisiae Mongolica Folium*. The ratios of the peak areas of peak 3 to peak 20 were in the ranges of 54.50 - 348.39 and 0.16 - 0.87 respectively, and the ratios of the peak areas of peak 19 to peak 20 were in the ranges of 18.55 - 128.46 and 0.01 - 0.14 respectively. These significant differences could be used for the identification of *Artemisiae Argyi Folium* and its adulterant *Artemisiae Mongolica Folium*. **Conclusion:** The research findings can be used for the identification of *Artemisiae Argyi Folium* and its adulterant *Artemisiae Mongolica Folium*, and these have certain reference significance for the research and analysis of *Artemisiae Argyi Folium* and related drugs.

Keywords: *Artemisiae Argyi Folium*; *Artemisiae Mongolica Folium*; adulterant; gas chromatography; chemometrics; data analysis; data preprocess; PCA; OPLS - DA

艾叶为菊科植物艾 *Artemisia argyi* Lévl. et Vant. 的干燥叶, 又称冰台、艾蒿、灸草^[1], 药用记录始载于汉末《名医别录》, 列为中品^[2], 具有温经止血、散寒止痛的功效, 主要用于吐血、衄血、崩漏、月经过多、胎漏下血、少腹冷痛、经寒不调、宫冷不孕的治疗, 外用可治疗皮肤瘙痒^[3]。艾叶主要含有挥发油类、黄酮类、萜类、苯丙素类、有机酸类、甾体类、多糖类及微量元素等成分, 现代研究证实这些成分具有抗菌、抗病毒、平喘镇咳、镇痛、抗炎、止血、抗凝等多种药理活性作用^[4-6]。

艾叶所属的菊科蒿属植物的分布较广, 据《中国植物志》记载, 菊科蒿属植物在我国有 186 个种、44 个变种^[7], 其中不乏艾叶的类同品。课题组在前期的调研工作中发现, 药材市场上存在将蒙古蒿叶冒充艾叶出售并使用的现象, 同时在种植基地里发现艾叶与蒙古蒿有伴生的情况, 若在采收时不注意区分, 艾叶样品常会混有蒙古蒿叶。

蒙古蒿 (*Artemisia mongolica* (Fisch. ex Besser) Nakai) 又名蒙蒿、狭叶蒿、水红蒿, 产于黑龙江、河北、陕西、河南、江苏等地^[7]。具有祛风散寒、散瘀消肿、理气安胎等功效^[8-9], 但不是《中华人民共和国药典》规定的艾叶药材来源植物, 毒副作用不明, 且

与艾叶功效不同, 不可代替艾叶入药, 若艾叶中混有蒙古蒿叶势必影响临床效果, 也无法保证人民群众用药的安全有效。对于艾叶和蒙古蒿叶的鉴别研究不多, 李欣等^[10]将艾叶与蒙古蒿叶的叶形特征进行比较区分, 李恩波等^[11]利用 DNA 条形码片段 ITS2 序列对艾叶及其混伪品进行了区分鉴别。这些方法为艾叶与蒙古蒿叶的鉴别提供了技术手段, 但由于艾叶与蒙古蒿叶形变异较大, 采收、晒干后叶片多皱缩、破碎, 仅依据外观性状难以将二者区分, 且这种基于叶形的鉴别对于粉碎后的样品无能为力; 对于 DNA 鉴别实验则多需要新鲜嫩叶, 操作也需要中药的分析人员具有一定的相关专业性。

艾叶中主要活性成分是挥发油, 气相色谱技术是研究挥发性成分的首选之一, 因此本研究建立了使用气相色谱分析艾叶及蒙古蒿叶的方法, 通过 NIST 质谱谱库检索的方式对色谱图中的 21 个色谱峰进行了指认, 其中 9 个色谱峰还通过对照品比对进行了确认; 由于样品数据较大, 本研究还借助适于处理复杂中药色谱数据并发现特征性差异组分的化学计量学技术^[12-14]对艾叶和蒙古蒿叶的色谱数据进行分析和研究, 明确了二者的化学差异所在, 为艾叶的质量分析与控制提供了数据参考和技术支撑。

1 材料

1.1 仪器

XPE205 十万分之一电子分析天平(梅特勒-托利多仪器(上海)有限公司);挥发油测定仪(上海那艾实验仪器有限公司);Agilent 7890B 气相色谱仪(安捷伦公司);赛默飞 TRACE1300 气质联用仪(赛默飞世尔科技公司)。

1.2 试药

α -蒎烯(Stanford Analytical Chemicals Inc., 批号 AF210929-14);桉油精(中国食品药品检定研究院,批号 110788-202108);侧柏酮(Stanford Analytical Chemicals Inc., 批号 THJ-982234);樟脑(中国食品药品检定研究院,批号 110747-202011);龙脑(中国食品药品检定研究院,批号 110881-201709); α -萜品醇(Stanford Analytical Chemicals Inc., 批号 AL210810-20);4-萜品醇(中国食品药品检定研究院,批号 11196-201501);反式石竹烯(Stanford Analytical Chemicals Inc., 批号 CYH-207225);氧化石竹烯(Stanford Analytical Chemicals Inc., 批号 CHO-222076);乙酸乙酯为分析纯,水为去离子水。

艾叶与其混伪品样品来源见表 1,经河南省药品医疗器械检验院王晓伟副主任药师鉴定样品 S1~S20 为菊科植物艾的干燥叶,S21~S29 为菊科植物蒙古蒿的干燥叶。

1.3 软件与语言

TraceFinder 化学工作站(赛默飞世尔科技公司, version 4.1);SIMCA 多元变量统计分析软件(Umetrics 公司, version 14.1);Python 计算机编程语言(Python Software Foundation, version 3.8.8)。

2 方法与结果

2.1 色谱-质谱条件

2.1.1 色谱条件 采用 Agilent HP-5 19091J(30 m \times 0.32 mm, 0.25 μ m) 色谱柱,进样口温度 250 $^{\circ}$ C,程序升温(初始温度 60 $^{\circ}$ C,以 3 $^{\circ}$ C \cdot min $^{-1}$ 的速率升温至 120 $^{\circ}$ C,再以 15 $^{\circ}$ C \cdot min $^{-1}$ 的速率升温至 250 $^{\circ}$ C),载气为氮气(纯度 99.99%),流速 1 mL \cdot min $^{-1}$,进样量 1 μ L。检测器为氢火焰离子检测器(FID)。

气质联用确认色谱峰时,采用 Agilent HP-5MS 19091J(30 m \times 0.32 mm, 0.25 μ m) 色谱柱,其他色谱条件同上。

2.1.2 质谱条件 检测器为电子轰击(EI)离子源,离子源温度为 230 $^{\circ}$ C,电子轰击能量为 70 eV,采集

表 1 样品来源信息

Tab. 1 Information of samples

样品 (sample)	产地 (origin)	批号/收集日期 (batch number/collection date)	种类 (class)
S1	安徽(Anhui)	220106022c	艾叶(Artemisiae)
S2	河南(Henan)	Y220301	Argyi Folium)
S3	安徽(Anhui)	211201	
S4	河南(Henan)	201201	
S5	安徽(Anhui)	220101	
S6	河南(Henan)	220301	
S7	安徽(Anhui)	20211101	
S8	山东(Shandong)	220401	
S9	河南(Henan)	210901	
S10	湖北(Hubei)	220201	
S11	湖北(Hubei)	220301	
S12	河南(Henan)	11220311034A	
S13	四川(Sichuan)	201201	
S14	四川(Sichuan)	211001	
S15	河南(Henan)	Y220101	
S16	湖北(Hubei)	220301	
S17	河南(Henan)	21120502	
S18	安徽(Anhui)	220301	
S19	河南(Henan)	220401	
S20	陕西(Shanxi)	220301	
S21	河南(Henan)	2021年10月(October 2021)	蒙古蒿叶
S22	河南(Henan)	2022年4月(April 2022)	(Artemisiae)
S23	河南(Henan)	2022年4月(April 2022)	Mongolica
S24	河南(Henan)	2022年5月(May 2022)	Folium)
S25	河南(Henan)	2022年5月(May 2022)	
S26	河南(Henan)	2022年5月(May 2022)	
S27	河南(Henan)	2022年6月(June 2022)	
S28	河南(Henan)	2022年6月(June 2022)	
S29	河南(Henan)	2022年6月(June 2022)	

模式为全扫描,质量扫描范围 m/z 为 40~600,扫描速度为 2 788 u \cdot s $^{-1}$ 。

2.2 溶液的制备

2.2.1 供试品溶液 取样品适量,剪碎成约 0.5 cm 的碎片,精密称取 2.5 g,置圆底烧瓶中,加水 300 mL,连接挥发油测定器。自测定器上端加水使充满刻度部分,并溢流入烧瓶为止,再加乙酸乙酯 2.5 mL,连接回流冷凝管。加热至沸腾,再加热 5 h,放冷,分取乙酸乙酯液,置 10 mL 量瓶中,用乙酸乙酯分次洗涤测定器及冷凝管,转入同一量瓶中,用乙酸乙酯稀释至刻度,摇匀,即得。

2.2.2 混合对照品溶液 精密称取 α -蒎烯、桉油精、侧柏酮、樟脑、龙脑、4-萜品醇、 α -萜品醇、反式

石竹烯、氧化石竹烯的对照品适量,置于 10 mL 量瓶中,加乙酸乙酯定容,得到混合对照品溶液(α -蒎烯 $0.375 \text{ mg} \cdot \text{mL}^{-1}$ 、桉油精 $0.159 \text{ mg} \cdot \text{mL}^{-1}$ 、侧柏酮 $0.186 \text{ mg} \cdot \text{mL}^{-1}$ 、樟脑 $0.335 \text{ mg} \cdot \text{mL}^{-1}$ 、龙脑 $0.038 \text{ mg} \cdot \text{mL}^{-1}$ 、4-萜品醇 $0.815 \text{ mg} \cdot \text{mL}^{-1}$ 、 α -萜品醇 $0.105 \text{ mg} \cdot \text{mL}^{-1}$ 、反式石竹烯 $0.331 \text{ mg} \cdot \text{mL}^{-1}$ 、氧化石竹烯 $0.156 \text{ mg} \cdot \text{mL}^{-1}$)。

2.3 方法学考察

2.3.1 精密度试验 取样品(S1)适量,按照“2.2.1”项方法制成供试品溶液,在“2.1.1”项的色谱条件下连续进样 6 次。以 3 号峰(桉油精)为参照峰,计算得到各共有峰相对保留时间的 $\text{RSD} \leq 0.003\%$,相对峰面积的 $\text{RSD} \leq 1.9\%$,表明仪器精密度良好。

2.3.2 重复性试验 取样品(S1)6 份,按照“2.1.1”项方法平行制备供试品溶液,在“2.1.1”项的色谱条件下测定。以 3 号峰(桉油精)为参照峰,计算得到各共有峰相对保留时间的 $\text{RSD} \leq 0.14\%$,

相对峰面积的 $\text{RSD} \leq 1.8\%$,表明方法重复性良好。

2.3.3 稳定性试验 取样品(S1)适量,按照“2.1.1”项方法制成供试品溶液,分别于室温放置 0、2、4、6、8、10、12 h 后,在“2.1.1”项的色谱条件下连续进样。以 3 号峰(桉油精)为参照峰,计算得到各共有峰相对保留时间的 $\text{RSD} \leq 0.02\%$,相对峰面积的 $\text{RSD} \leq 1.21\%$,表明方法稳定性良好。

2.4 色谱峰的确证

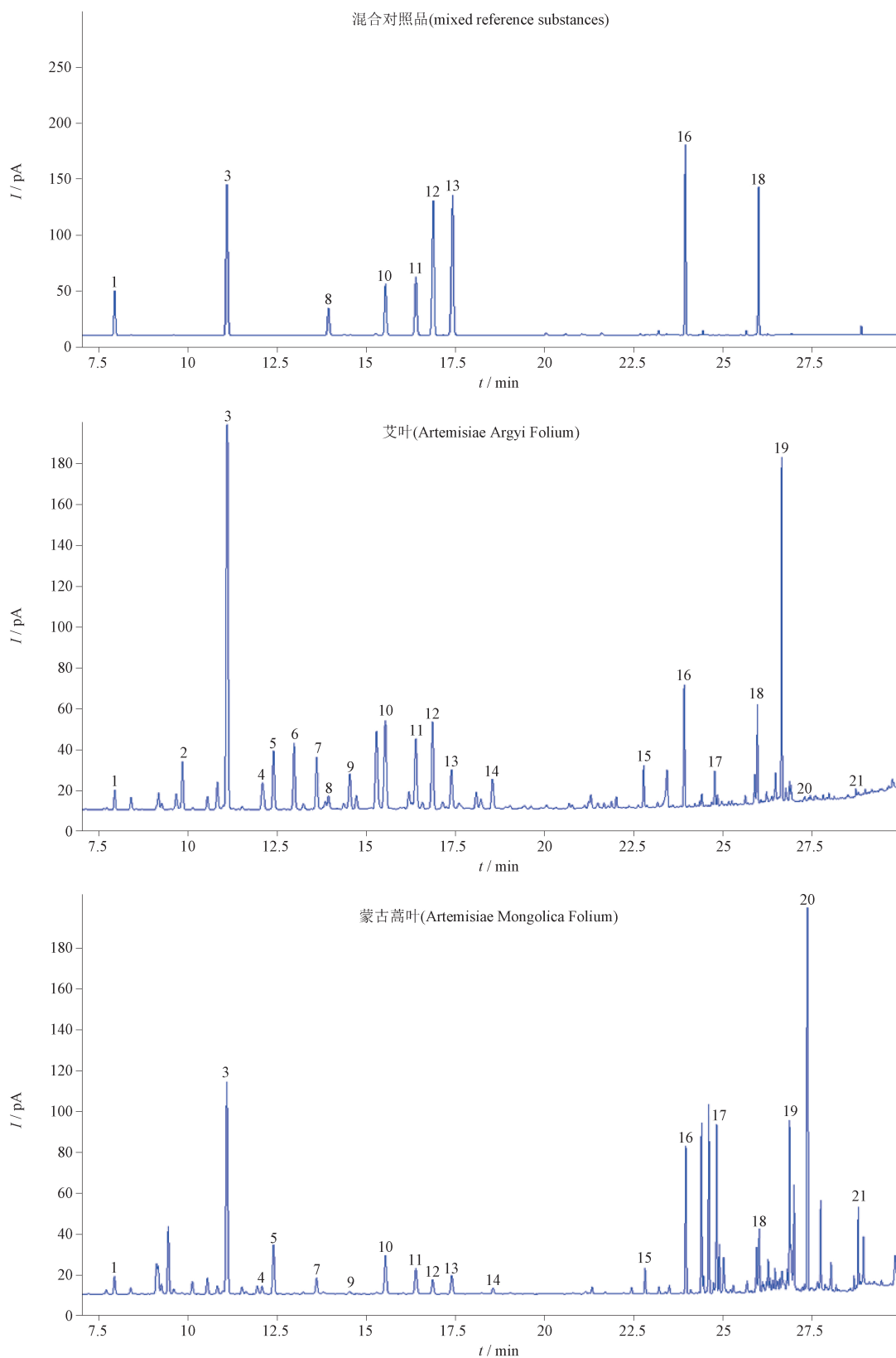
分别取混合对照品溶液、艾叶供试品溶液、蒙古蒿叶供试品溶液,按照气质联用色谱条件进样并检测,通过 Trace-Finder 4.1 化学工作站处理及 NIST 质谱谱库检索(色谱峰匹配度在 800 以上)可确定 21 个色谱峰成分。另外,经过与混合对照品溶液进行色谱峰出峰时间比对,进一步确定了 α -蒎烯、桉油精、侧柏酮、樟脑、龙脑、4-萜品醇、 α -萜品醇、反式石竹烯、氧化石竹烯色谱峰。21 个色谱峰的化学成分信息详见表 2,各类溶液色谱图见图 1。

表 2 21 个共有峰信息

Tab. 2 Information of 21 common peaks

峰号 (peak No.)	化合物 (compound)	分子式 (molecular formula)	CAS 号 (CAS number)	匹配率 (match ratio)
1*	α -蒎烯(α -pinene)	$\text{C}_{10}\text{H}_{16}$	80-56-8	938
2	(3E)-2,5,5-三甲基-3,6-庚二烯-2-醇((3E)-2,5,5-trimethylhepta-3,6-dien-2-ol)	$\text{C}_{10}\text{H}_{18}\text{O}$	26127-98-0	887
3*	桉油精(1,8-cineole)	$\text{C}_{10}\text{H}_{18}\text{O}$	470-82-6	955
4	蒿酮(artemisyl ketone)	$\text{C}_{10}\text{H}_{16}\text{O}$	546-49-6	894
5	(2R,5R)-2-甲基-5-丙烷-2-基双环[3.1.0]己烷-2-醇((2R,5R)-2-methyl-5-propan-2-ylbicyclic [3.1.0] hexan-2-ol)	$\text{C}_{10}\text{H}_{18}\text{O}$	17699-16-0	938
6	3,3,6-三甲基庚-1,5-二烯-4-醇(3,3,6-trimethylheptyl-1,5-diene-4-ol)	$\text{C}_{10}\text{H}_{18}\text{O}$	27644-04-8	849
7	松油醇(terpineol)	$\text{C}_{10}\text{H}_{18}\text{O}$	7299-41-4	915
8*	侧柏酮(thujone)	$\text{C}_{10}\text{H}_{16}\text{O}$	546-80-5	914
9	(-)-宁酮((1)-ninone)	$\text{C}_{20}\text{H}_{32}\text{O}_2$	471-15-8	910
10*	樟脑(camphor)	$\text{C}_{10}\text{H}_{16}\text{O}$	464-48-2	905
11*	龙脑(borneol)	$\text{C}_{10}\text{H}_{18}\text{O}$	507-70-0	950
12*	4-萜品醇(4-carvomenthenol)	$\text{C}_{10}\text{H}_{18}\text{O}$	20126-76-5	909
13*	α -萜品醇(α -carvomenthenol)	$\text{C}_{10}\text{H}_{18}\text{O}$	98-55-5	912
14	(1R,5S)-rel-香芹醇[(1R,5S)-rel carvol]	$\text{C}_{10}\text{H}_{16}\text{O}$	1197-07-5	967
15	(-)- α -蒎烯(copaene)	$\text{C}_{15}\text{H}_{24}$	3856-25-5	843
16*	反式石竹烯(trans caryophyllene)	$\text{C}_{15}\text{H}_{24}$	87-44-5	949
17	α -石竹烯(α -caryophyllene)	$\text{C}_{15}\text{H}_{24}$	6753-98-6	880
18*	氧化石竹烯(caryophyllene oxide)	$\text{C}_{15}\text{H}_{24}\text{O}$	1139-30-6	919
19	(1S,8 α)-10-氢-1,4 α -二甲基-7 β -异丙烯基-1-萘酚[(1S,8 α)-decahydro-1,4 α -dimethyl-7 β -isopropenyl-1-naphthol]	$\text{C}_{15}\text{H}_{26}\text{O}$	5945-72-2	915
20	母菊萹(chamazulene)	$\text{C}_{14}\text{H}_{16}$	529-05-5	818
21	棕榈酸(palmitic acid)	$\text{C}_{16}\text{H}_{32}\text{O}_2$	57-10-3	908

注(note): * 为经对照品比对的色谱峰(the chromatographic peak confirmed by reference substance)



1 ~ 21. 峰号, 同表 2 (peak No. same as Tab. 2)

图 1 对照品和样品色谱图

Fig. 1 Chromatograms of reference substances and samples

2.5 无监督分析

2.5.1 相似度分析 将测得的 20 批正品艾叶样品 (S1 ~ S20) 的 21 个色谱峰数据取平均值, 记为 M , 然

后分别按照夹角余弦法和欧氏距离法求 29 批样品与 M 的相似度, 具体相似度值见表 3, 并绘制相关相似度散点图, 见图 2。

表 3 样品的相似度值
Tab. 3 Sample similarities

样品 (sample)	夹角余弦法 (angle cosine method)	欧氏距离法 (euclidean distance method)	样品 (sample)	夹角余弦法 (angle cosine method)	欧氏距离法 (euclidean distance method)
S1	0.98	249.68	S16	0.95	245.98
S2	0.96	247.28	S17	0.97	531.33
S3	0.94	265.87	S18	0.79	752.57
S4	0.98	167.64	S19	0.98	204.60
S5	0.98	143.41	S20	0.89	349.63
S6	0.97	268.36	S21	0.24	1 433.23
S7	0.99	142.59	S22	0.30	1 995.62
S8	0.97	271.43	S23	0.34	895.74
S9	0.99	127.60	S24	0.27	1 375.56
S10	0.97	185.87	S25	0.48	799.90
S11	0.97	205.51	S26	0.52	1 984.83
S12	0.96	269.82	S27	0.22	1 848.76
S13	0.99	148.78	S28	0.67	894.05
S14	0.98	185.28	S29	0.48	762.98
S15	0.90	335.28			

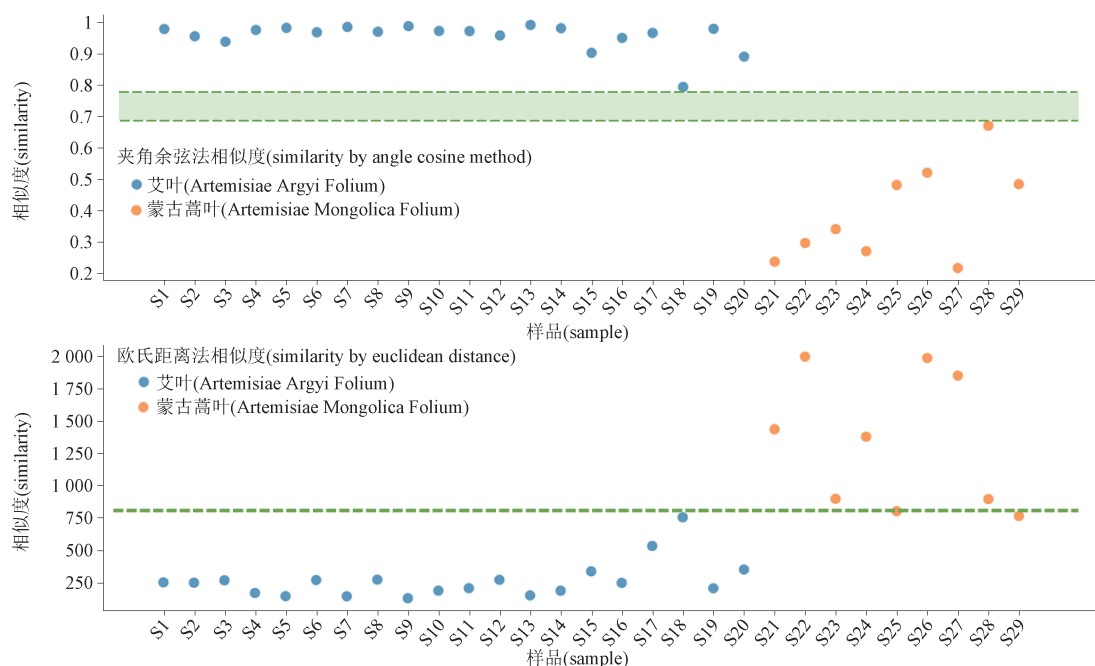


图 2 样品相似度散点图

Fig. 2 Scatter plots of similarity analysis for samples

2.5.2 相关性分析 为了从色谱峰 [峰 1 ~ 峰 21 (p1 ~ p21)] 之间相关性的角度比较分析正品艾叶及

其混伪品蒙古蒿叶的差异, 对这 2 类样品分别进行变量 Pearson 相关系数分析, 结果以热度图表示, 见图 3。

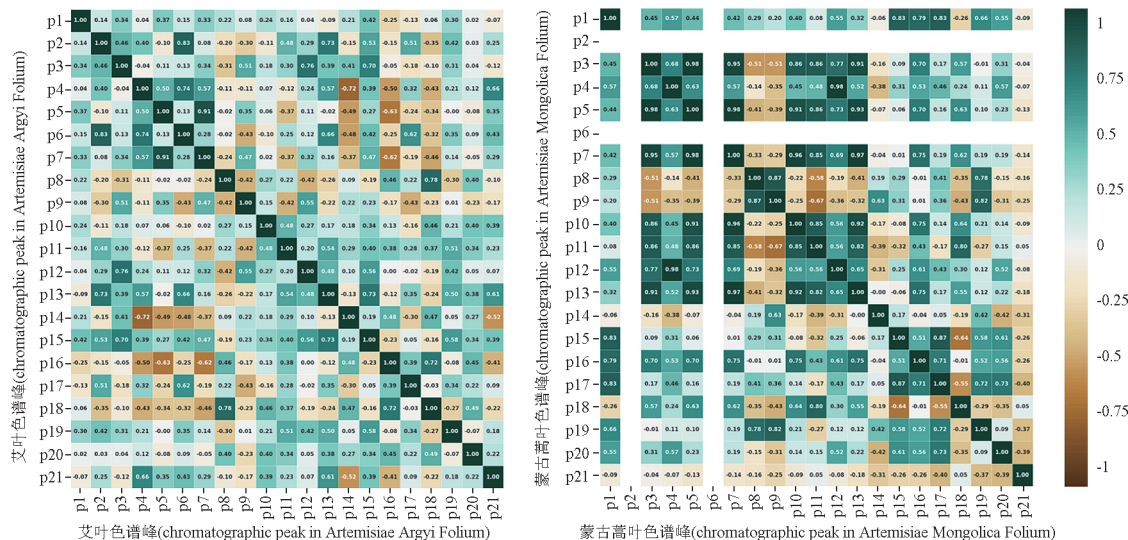


图3 样品的色谱峰间相关性热度图

Fig. 3 Heat maps of chromatographic peak to peak correlation from samples

将数据矩阵倒置处理,进行 Pearson 相关系数分析,可得出 29 批样品之间的相关性信息,结果同样

以热度图表示,见图 4。

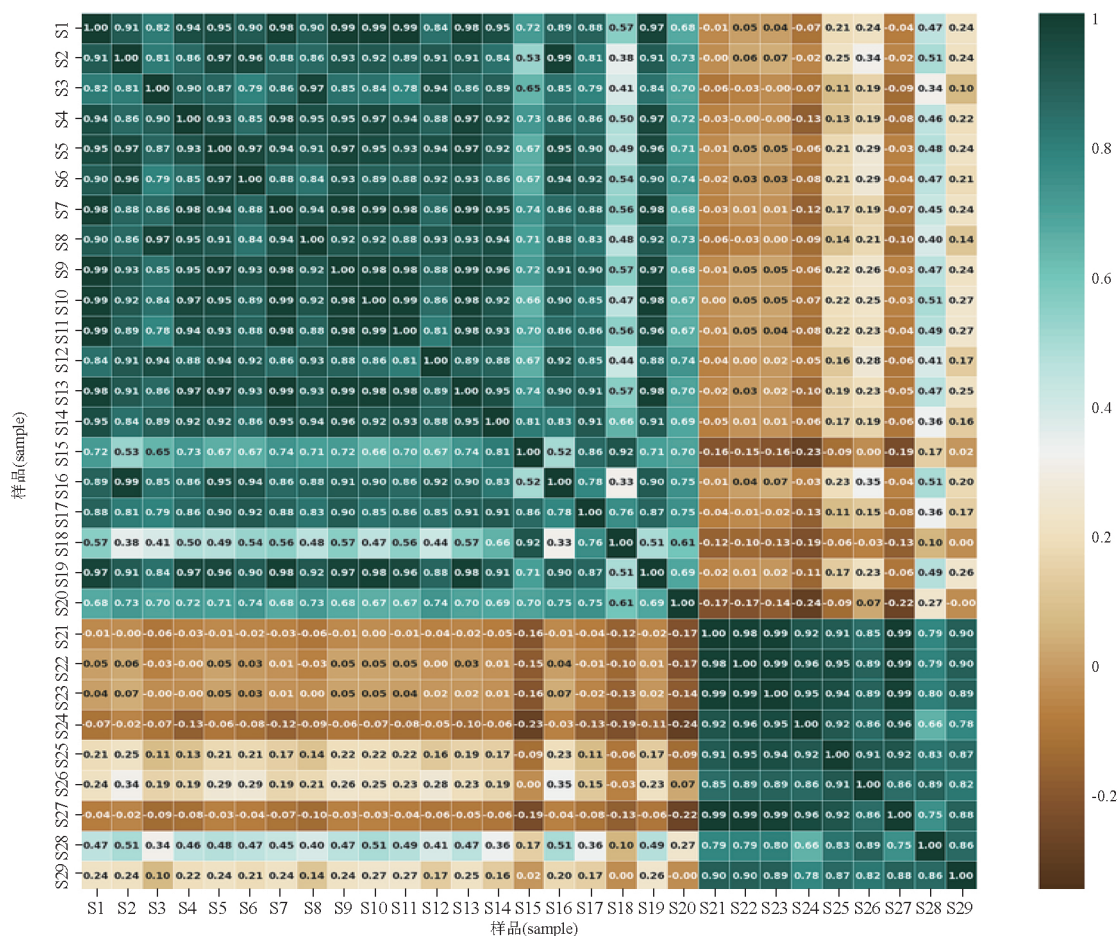


图4 样品间相关性热度图

Fig. 4 Heat maps of sample to sample correlation

2.5.3 聚类分析 将各批样品的 21 个色谱峰数据 L2 范数归一化预处理后,进行聚类分析,采用 Ward

法和欧式距离度量,结果数据聚类热点图及树状图见图 5。

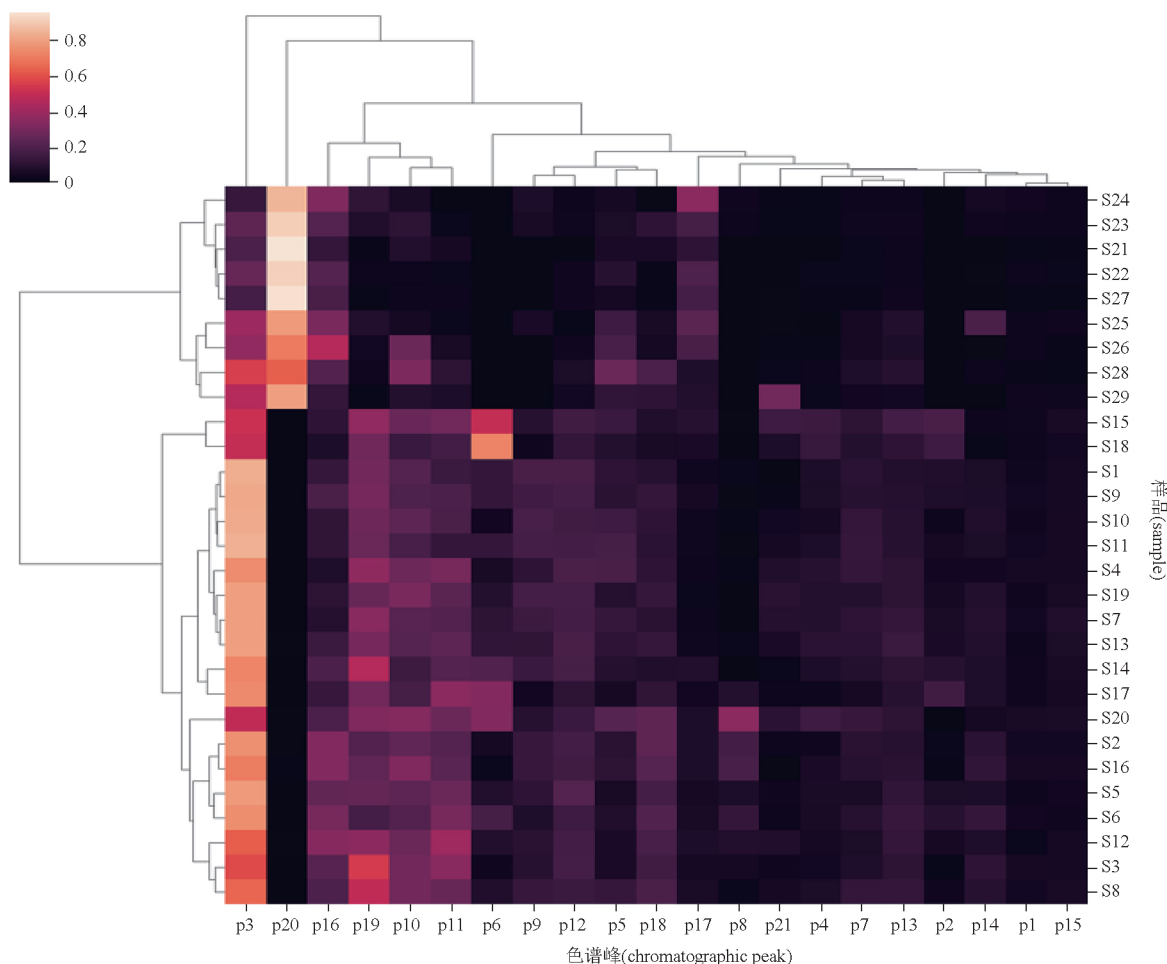


图 5 聚类分析热图

Fig. 5 Heat maps of cluster analysis

2.5.4 主成分分析 对各批样品的 21 个色谱峰原数据及 L2 范数归一化预处理后的数据分别进行主成分分析,结果各主成分累计方差曲线分别见图 6 - A 和图 7 - A;样品在前 2 个主成分(PC1 和 PC2)平面的分布分别见图 6 - B 和图 7 - B;累计方差达 95% 总方差的主成分载荷分布分别见图 6 - C 和图 7 - C。

2.6 监督分析

使用正交偏最小二乘判别分析(orthogonal partial least squares - discriminant analysis, OPLS - DA)对 29 批样品的 L2 范数归一化预处理后的色谱数据进行监督分析,以进一步明确艾叶与其混伪品蒙古蒿叶之间最具特征性的化学成分。OPLS - DA 模型 $R^2 X$ 、 $R^2 Y$ 及 Q^2 分别为 0.927、0.507 和 0.504,

200 次置换检验表明无过拟合,样品的最小二乘法正交分布图见图 8 - A;对峰 1 ~ 峰 21 (p1 ~ p21) 进行 VIP 排序,排序结果见图 8 - B。

3 讨论

3.1 化学分析实验部分

在对艾叶与蒙古蒿叶样品进行性状分析时发现,二者都具有特殊气味,并且艾叶散发的清香气味明显与蒙古蒿叶不同,结合艾叶中大量含有挥发油的相关研究^[15],本研究将艾叶与其混伪品蒙古蒿叶的挥发性成分作为二者的化学鉴别分析的研究切入点。

在确定挥发性成分提取方法时,分别考察了正己烷超声法、正己烷加热回流法、水蒸气蒸馏法,

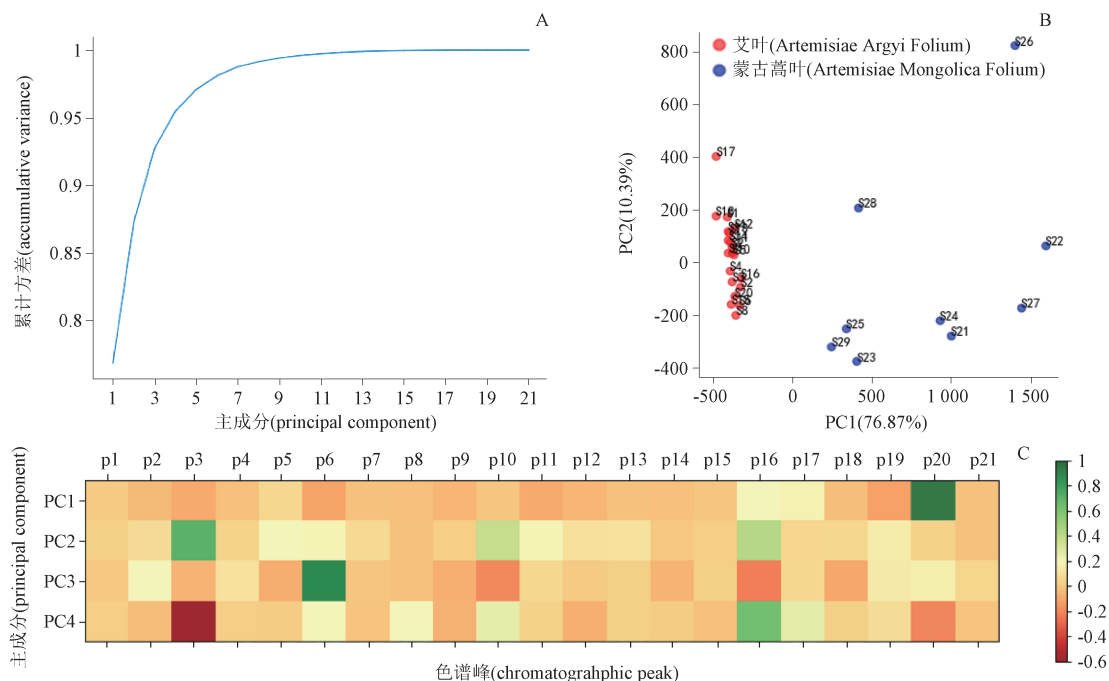


图 6 未经预处理的数据主成分分析

Fig. 6 Principal component analysis of unprocessed data

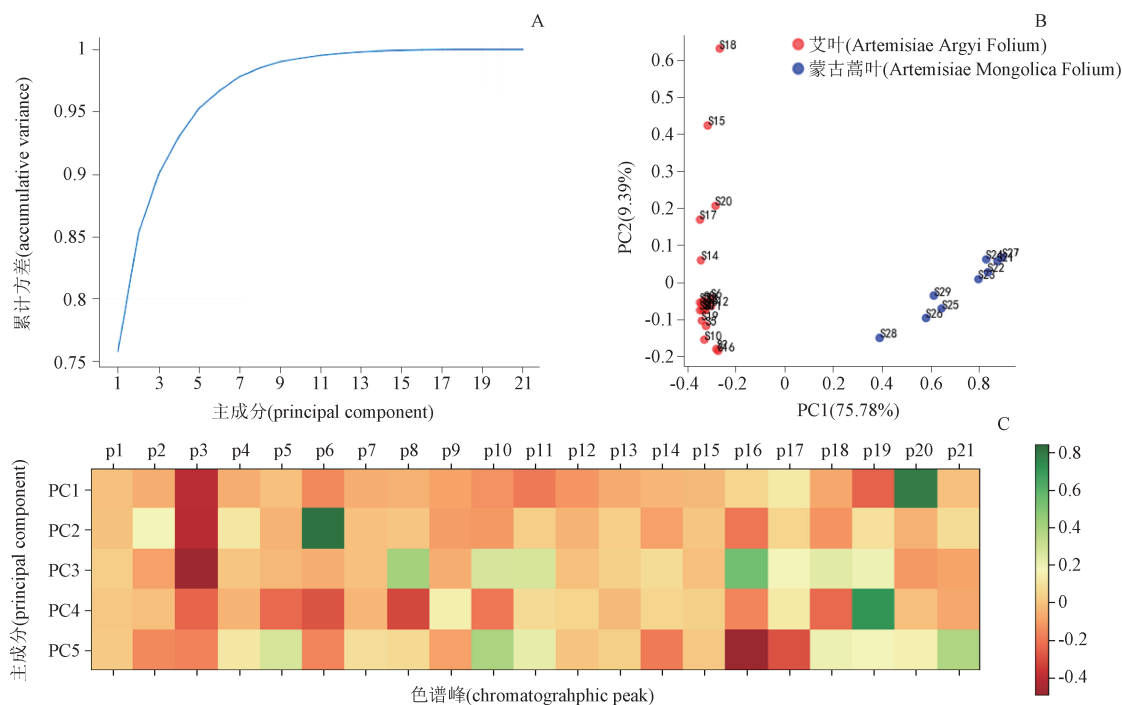


图 7 归一化预处理后的数据主成分分析

Fig. 7 Principal component analysis of normalized preprocessed data

结果发现水蒸气蒸馏法提取所得挥发性成分最丰富,可以得到更多的化合物的色谱信息,因此选择该法作为样品提取方法。样品的气相色谱在 7 min 前

几乎没有色谱峰出现,属于无效的色谱分析时间段,因此本研究在气相色谱条件设置时仅记录了 7 min 后的色谱信息。

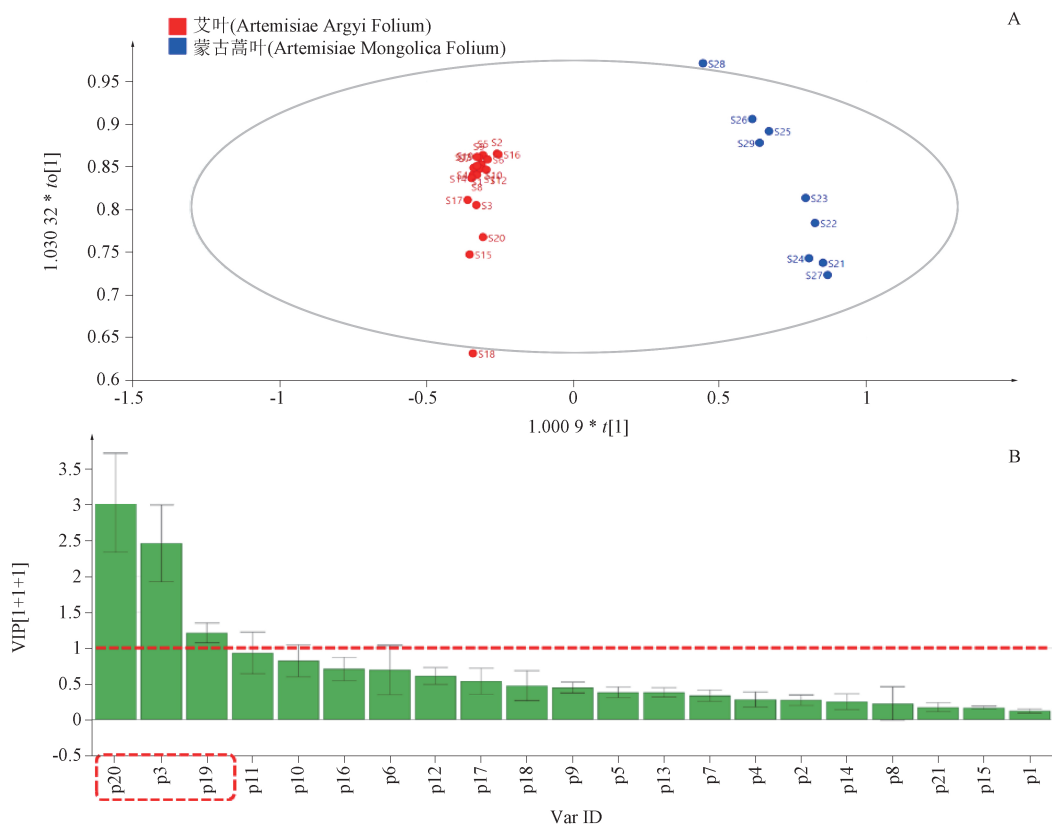


图 8 OPLS-DA 图
Fig. 8 Plots of OPLS-DA

3.2 数据分析实验部分

3.2.1 无监督分析 相似度分析是中药色谱分析时常用的分析方法,其中对于相似度的计算有多种度量方法^[12]。本研究分别使用了其中常用的夹角余弦法和欧氏距离法对艾叶及其混伪品蒙古蒿叶色谱数据与艾叶样品平均数据 M 进行相似度的计算。由相似度计算结果绘制的散点图(图 2)可知,艾叶与蒙古蒿叶的相似度具有较明显的差异。夹角余弦法中,S18 的相似度是 0.79,属艾叶中相似度最低的样品;蒙古蒿叶样品中 S28 的相似度为 0.67,是蒙古蒿叶中相似度最高的样品;从本研究样品来看,艾叶和蒙古蒿叶样品的夹角余弦法相似度鉴别窗口可达 0.12。使用欧氏距离法度量相似度时,艾叶样品中距离 M 最远(即相似度最低)的样品为 S18,欧氏距离为 752.56;蒙古蒿叶样品中以 S29 距离 M 最近,距离为 762.98;从本研究样品来看,艾叶和蒙古蒿叶样品的欧氏距离法相似度鉴别窗口为 10.42。关于色谱数据相似度的分析原理在文献[12]中已有详细的讲解,在此不再赘述。不同的相似度度量方法代表

了不同的数学思想,方法不同,尽管趋势一样,详细的结果也可能会有些许差异,本研究相似度的研究就诠释了这一点。采用 2 种度量方法得到的结果均表明艾叶样品中 S18 对 M 的相似度最低(欧氏距离最远),而蒙古蒿叶样品的相似度结果的比较却有不同,S28 对 M 的夹角余弦法相似度最大,而 S29 则距离 M 的欧氏距离最近。欧氏距离法的相似度鉴别窗口绝对数值大于夹角余弦法,但是图 2 表明夹角余弦法的相似度鉴别窗口更大,更易于相似度鉴别艾叶和蒙古蒿叶的操作。

相关性分析是测量分析定量数据之间关系情况的分析方法,多用于分析变量间关系情况,其中多为直接计算变量间相关程度,即简单相关分析。简单相关分析常使用 Pearson 相关系数、Kendall 相关系数和 Spearman 相关系数等,其中最常用的为 Pearson 相关系数^[16-17]。本研究使用了 Pearson 相关系数分别对艾叶和蒙古蒿叶的色谱峰数据进行了分析,试图从化学成分间的相关性方面揭示艾叶和蒙古蒿叶的生物种属异同。如图 3 所示,艾叶样品的各色谱峰

呈较强正相关(Pearson 相关系数 ≥ 0.75)的色谱峰:峰 2 与峰 6(Pearson 相关系数 0.83)、峰 3 与峰 12(Pearson 相关系数 0.76)、峰 5 与峰 7(Pearson 相关系数 0.91)、峰 8 与峰 18(Pearson 相关系数 0.78),没有发现呈较强负相关(Pearson 相关系数 ≤ -0.75)的色谱峰,仅有相对较强负相关的色谱峰:峰 4 与峰 14(Pearson 相关系数 -0.72)、峰 5 与峰 16(Pearson 相关系数 -0.63)、峰 7 与峰 16(Pearson 相关系数 -0.62)。蒙古蒿叶样品的各色谱峰呈较强正相关(Pearson 相关系数 ≥ 0.75)的色谱峰:峰 1 与峰 15(Pearson 相关系数 0.83)、峰 1 与峰 16(Pearson 相关系数 0.79)、峰 1 与峰 17(Pearson 相关系数 0.83)、峰 3 与峰 5(Pearson 相关系数 0.98)、峰 3 与峰 7(Pearson 相关系数 0.95)、峰 3 与峰 10(Pearson 相关系数 0.86)、峰 3 与峰 11(Pearson 相关系数 0.86)、峰 3 与峰 13(Pearson 相关系数 0.93)、峰 4 与峰 12(Pearson 相关系数 0.98)、峰 5 与峰 7(Pearson 相关系数 0.98)、峰 5 与峰 10(Pearson 相关系数 0.91)、峰 5 与峰 11(Pearson 相关系数 0.86)、峰 5 与峰 13(Pearson 相关系数 0.93)、峰 7 与峰 10(Pearson 相关系数 0.96)、峰 7 与峰 11(Pearson 相关系数 0.85)、峰 7 与峰 13(Pearson 相关系数 0.97)、峰 7 与峰 16(Pearson 相关系数 0.75)、峰 8 与峰 9(Pearson 相关

系数 0.87)、峰 8 与峰 19(Pearson 相关系数 0.78)、峰 9 与峰 19(Pearson 相关系数 0.82)、峰 10 与峰 11(Pearson 相关系数 0.85)、峰 10 与峰 13(Pearson 相关系数 0.92)、峰 10 与峰 16(Pearson 相关系数 0.75)、峰 11 与峰 13(Pearson 相关系数 0.82)、峰 11 与峰 18(Pearson 相关系数 0.80),同样没有发现呈较强负相关(Pearson 相关系数 ≤ -0.75)的色谱峰,仅有相对较强负相关的色谱峰:峰 9 与峰 11(Pearson 相关系数 -0.67)、峰 15 与峰 18(Pearson 相关系数 -0.64)。艾叶与蒙古蒿叶样品的色谱峰呈较强正相关关系图见图 9。该图清晰显示了艾叶相对于蒙古蒿叶而言,色谱峰间较强正相关情况较少,且二者色谱峰较强正相关性相同的情况仅有峰 5 与峰 7。由色谱峰相关性分析结果可知,艾叶与蒙古蒿叶虽有较多相同的内在的化学成分,但这些化学成分之间生成与代谢的关系却有很大的区别。忽略色谱数据的样品来源,按照 21 个色谱峰进行样品相关性分析,Pearson 相关系数热度图如图 4 所示,艾叶与蒙古蒿叶样品可以准确地按照热度图的颜色进行区分,可见按照艾叶和蒙古蒿叶分开进行色谱峰 1~峰 21(p1~p21)相关性分析是合理的,图 3 和图 9 在一定程度上可以反映两者色谱峰相关性差异。

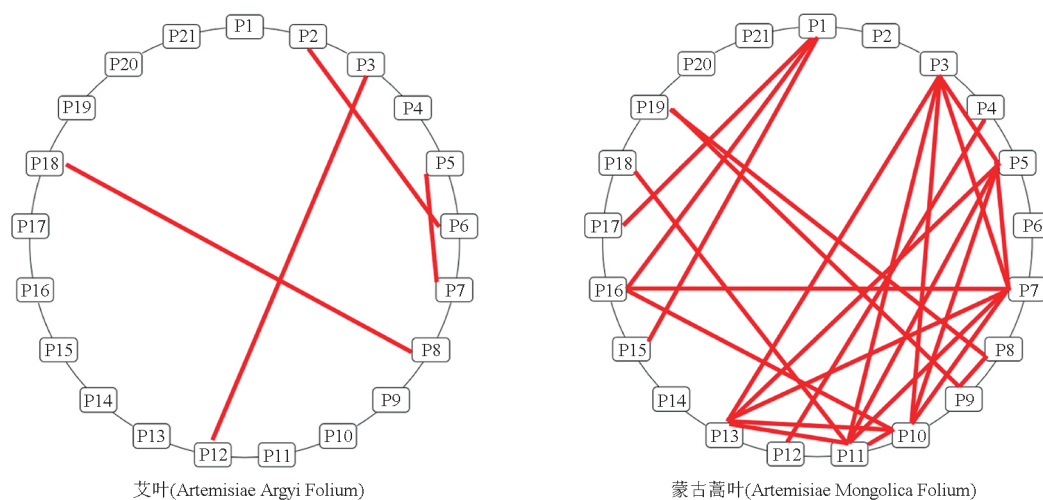


图 9 艾叶与蒙古蒿叶较强正相关性示意图

Fig. 9 Schematic diagram of strong positive correlation of samples from *Artemisiae Argyi Folium* and *Artemisiae Mongolica Folium*

基于欧氏距离的聚类分析结果显示,本研究的色谱数据可以将样品按照艾叶与蒙古蒿叶种类准确区分开,与欧氏距离法度量的相似度分析结果具有

一致性。而色谱峰聚类表明峰 3、峰 10、峰 11、峰 16、峰 19 和峰 20 属于较特殊的一系列色谱峰。

在主成分分析时,分析数据的处理采用了 2 种

方式,即原始色谱峰的峰面积数据和经过 L2 范数归一化预处理后的色谱峰的峰面积数据。由图 6 - A 和图 7 - A 的主成分累计方差曲线可知,以累计方差 $\geq 95\%$ 为限,数据未经预处理和归一化预处理后的主成分数分别为 4 个和 5 个;图 6 - B 和图 7 - B 显示数据经过归一化预处理后在第一和第二主成分(PC1 和 PC2)二维平面上分布按品种更加集中,且无论数据是否经过归一化预处理,品种的分布差异只表现在第一主成分(PC1)差异上,艾叶样品的 PC1 均为负值,蒙古蒿叶样品的 PC1 均为正值;图 6 - C 和图 7 - C 为数据变量的主成分载荷热度图,该图显示数据未经预处理和归一化预处理后数据变量的主成分载荷具有较明显差异,数据未经预处理时仅有峰 20 在 PC1 上具有较大的载荷,而归一化预处理后则可见峰 3、峰 19 和峰 20 在 PC1 上具有较大的载荷。主成分分析结果表明,无论数据是否经过归一化预处理,样品在 PC1 和 PC2 组成的二维平面上的分布都具有不同种类的分布特征,但是数据经过归一化预处理后样品的分布特征更加显著,并且 PC1 的色谱峰变量载荷的区分也更加明显。

3.2.2 监督分析 OPLS - DA 作为代谢组学数据分

析中常用算法,其特点是通过偏最小二乘法和正交化处理去除自变量中与分类标签无关的信息,使得分类信息集中体现在一个潜在变量上,可用于不同种类样品的差异特征变量的发现^[18-21]。图 8 - B 显示,色谱峰中仅有峰 20、峰 3 和峰 19 的 VIP 值 >1 ,提示这 3 个色谱峰可能是样品分类的差异性特征变量,这点与主成分分析中的变量载荷情况基本一致。VIP 值 <1 但是仅低于峰 20、峰 3 和峰 19 的色谱峰为峰 11、峰 10 和峰 16,这些与聚类分析结果也具有 consistency。

3.2.3 鉴别特征的筛选

艾叶与蒙古蒿叶的色谱峰峰面积数据中,蒙古蒿叶样品(S21 ~ S29)均未检出峰 6,而艾叶样品(S1 ~ S20)则都有峰 6 检出,如果直接从数据上看,无疑会认为峰 6 是艾叶与蒙古蒿叶样品差异的特征色谱峰,然而通过化学计量学分析技术与算法进行分析,表明二者差异性的特征色谱峰中并无峰 6。艾叶样品(S1 ~ S20)的峰 6 峰面积组内波动非常大,RSD 高达 131.35%,这种波动可从峰 6 峰面积柱状图(图 10)直观体现,这可能是峰 6 未被算法选作差异特征色谱峰的原因。由此可见,仅仅从数据表面得出,而未接受数学和统计学分析检验的结论可能并不可靠。

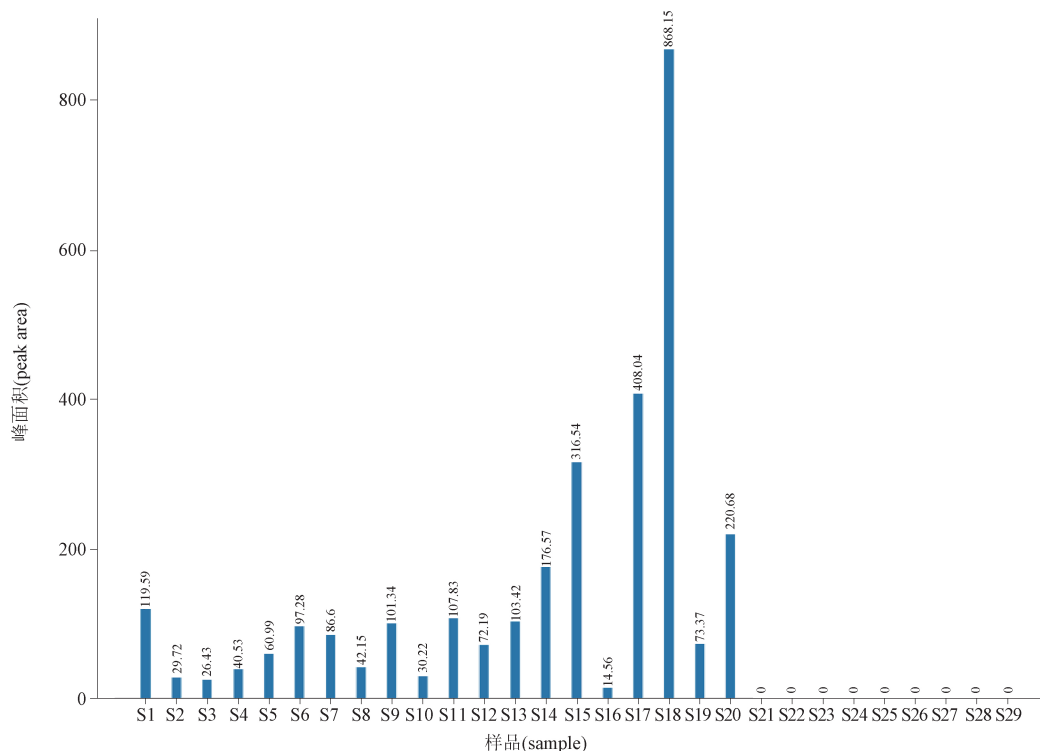


图 10 峰 6 峰面积柱状图

Fig. 10 Peak area bar chart of p6

峰 20、峰 3 和峰 19 是由以上各类化学计量学算法分析得出的艾叶与蒙古蒿叶的差异特征色谱峰,考虑到样品取样量及仪器进样体积等差异的影响,实际鉴别操作中可使用设定峰面积比值范围的方法。峰 20、峰 3 和峰 19 相互之间峰面积比值有 3 种情况,峰面积比值的柱状图见图 11。该图显示,艾叶样品(S1 ~ S20)与蒙古蒿叶样品(S21 ~ S29)的峰 3 与峰 20 峰面积比值范围分别为 54.50 ~ 348.39 和 0.16 ~ 0.87;峰 19 与峰 20 峰面积比值范围分别为 18.55 ~ 128.46 和 0.01 ~ 0.14,峰 3 与峰 19 峰面积比值范围分别为 1.06 ~ 4.44 和 1.15 ~ 38.94。艾叶同蒙古蒿叶的峰 3

与峰 20、峰 19 与峰 20 的 2 组峰面积比值范围具有较大的差异,而峰 3 与峰 19 峰面积比值范围则有重叠,由此可知峰 3 与峰 20、峰 19 与峰 20,这 2 组峰面积比值可用作艾叶和蒙古蒿叶的鉴别,并且艾叶和蒙古蒿叶的这 2 组比值范围都有较大的差异,较大的数值差异窗口更加表明这种鉴别方法具有较大的特征性。而峰 20、峰 3 和峰 19 对应的化学成分为母菊萹、桉油精和(1S,8 α) - 十氢 - 1,4 α β - 二甲基 - 7 β - 异丙烯基 - 1 - 萘酚,在实际鉴别应用中可使用对照品定位之后比较色谱峰的峰面积比值。而这 3 个化学成分比值的差异,也暗示艾与蒙古蒿生物体内代谢方式与途径的不同。

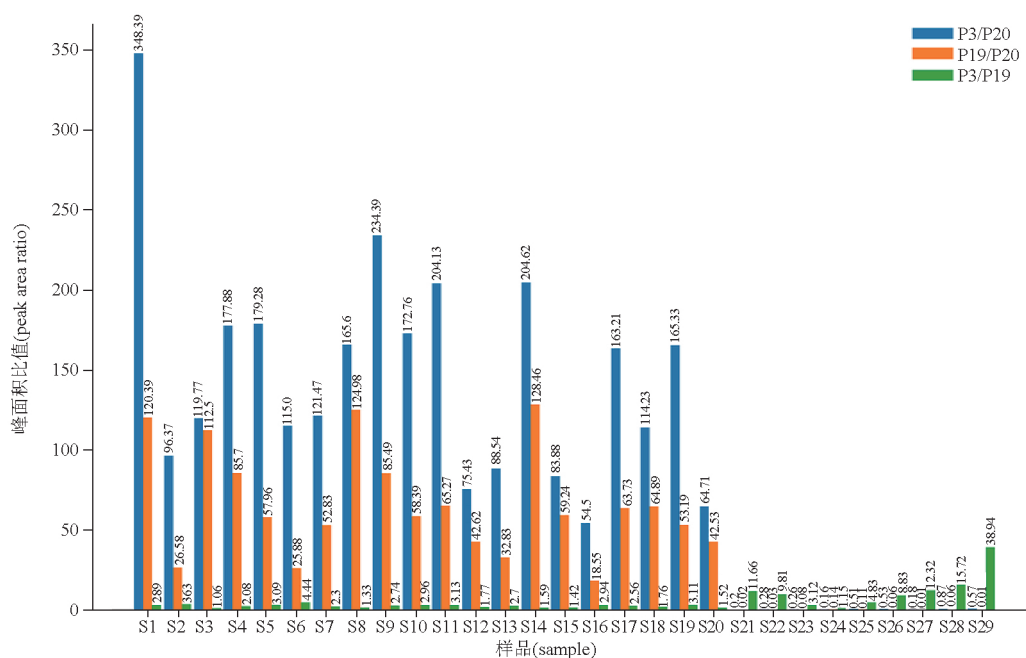


图 11 峰 20、峰 3 和峰 19 峰面积比值柱状图

Fig. 11 Peak area ratios bar chart of peak 20, peak 3 and peak 19

4 结论

本研究建立了艾叶与蒙古蒿叶气相指纹图谱,对其中的 21 个色谱峰通过对照品和 NIST 质谱谱库检索进行了确定与指认。对得到的各样品的色谱数据进行了数据分析实验,并从数据是否进行了预处理的角度进行了简单探讨,明确了预处理技术在数据分析实验中的重要性。对艾叶和蒙古蒿叶样品的色谱数据分析实验主要分为无监督分析和监督分析,无监督分析结果表明色谱数据的确存在因样品不同的差异性,这种差异不仅仅体现在每批样品数据的本身,如聚类分析和相似度分析等,还体现在 21 个色谱峰间相关性的不一致,如色谱峰的相关性分

析。由于样品批数较少,并不适合建立和训练相对复杂的基于常用机器学习的模型进行未知样品鉴别,因此在监督分析中,本研究采用了代谢组学研究中寻找差异特征代谢物时十分常用的 OPLS - DA 法,得到差异特征色谱峰峰 20 (母菊萹)、峰 3 (桉油精)和峰 19 [(1S,8 α) - 十氢 - 1,4 α β - 二甲基 - 7 β - 异丙烯基 - 1 - 萘酚],而这一结果与主成分分析法的变量载荷结果可以相互印证。最后,通过计算和比较这 3 个化学成分色谱峰的峰面积比值,确定了其中 2 组色谱峰的峰面积比值,通过这一结果可以准确鉴别艾叶和其混伪品蒙古蒿叶。考虑到样品批数有限,本研究未设立艾叶和蒙古蒿叶比值的

限度,这2组比值限度需要下一步大量样品收集和数据的测定。本研究的相关成果,不仅为艾叶真伪鉴别检验工作和相关药品监管提供了重要的技术支撑,而且对于类似情景的数据分析也具有一定的参考意义。

参考文献

- [1] 南京中医药大学. 中药大辞典[M]. 上海: 上海科学技术出版社, 2005: 801
Nanjing University of Chinese Medicine. Dictionary of Traditional Chinese Medicine[M]. Shanghai: Shanghai Scientific and Technical Publishers, 2005: 801
- [2] 聂麟, 刘畅, 单承莺. 艾草的本草考证及资源分布[J]. 中国野生植物资源, 2019, 38(4):93
NIE W, LIU C, SHAN CY. Textual research and resources distribution of wormwood[J]. *Chin Wild Plant Res*, 2019, 38(4):93
- [3] 中华人民共和国药典 2020 年版. 一部[S]. 2020: 91
ChP 2020. Vol I[S]. 2020: 91
- [4] 兰晓燕, 张元, 朱龙波, 等. 艾叶化学成分、药理作用及质量研究进展[J]. 中国中药杂志, 2020, 45(17):4017
LAN XY, ZHANG Y, ZHU LB, *et al.* Research progress on chemical constituents from *Artemisiae Argyi Folium* and their pharmacological activities and quality control[J]. *China J Chin Mater Med*, 2020, 45(17):4017
- [5] 袁林祥, 吴航宇, 邱彩玲. 艾叶的药用活性成分、药理作用及临床应用浅析[J]. 当代医药论丛, 2020, 18(2):171
YUAN LX, WU HY, QIU CL. A brief analysis of pharmaceutical active ingredients, pharmacological action and clinical application of *Artemisia argyi*[J]. *Contemp Med Forum*, 2020, 18(2):171
- [6] 张雪琳, 陈新旺, 吴毅明. 近 10 年来艾叶挥发油的化学成分及药理活性研究进展[J]. 中华中医药学刊, 2021, 39(5):111
ZHANG XL, CHEN XW, WU YM. Research progress on chemical constituents and pharmacological activities of volatile oil of *Aiye (Artemisia argyi)* [J]. *Chin Arch Tradit Chin Med*, 2021, 39(5):111
- [7] 中国科学院中国植物志编辑委员会. 中国植物志. 第 76 卷. 第二分册[M]. 北京: 科学出版社, 1991: 2,111
Editorial Committee of Flora of China. *Flora of China*. Vol 76(2) [M]. Beijing: Science Press, 1991: 2,111
- [8] 其乐木格, 包文强, 何祥, 等. 气质联用技术对蒙古蒿挥发油化学组成的鉴定[J]. 内蒙古民族大学学报(自然科学版), 2021, 36(2):140
QI LMG, BAO WQ, HE X, *et al.* Identification of the chemical composition of the volatile oil from *Artemisia mongolica* by GC-MS[J]. *J Inner Mongolia Minzu Univ (Nat Sci)*, 2021, 36(2):140
- [9] 董岩, 祁伟, 肖传勇. 蒙古蒿挥发油化学成分及其抑菌活性分析[J]. 中国药学杂志, 2013, 48(16):1359
DONG Y, QI W, XIAO CY. Analysis of chemical constituents of volatile oils from *Artemisia mongolica* and their antimicrobial activities[J]. *Chin Pharm J*, 2013, 48(16):1359
- [10] 李欣, 隋秀竹. 艾叶及其伪品蒙古蒿叶的鉴别[J]. 中国中药杂志, 1991, 16(8):458
LI X, SUI XZ. Identification of *Artemisia argyi* leaf and its counterfeit *Artemisia mongolica* leaf [J]. *China J Chin Mater Med*, 1991, 16(8):458
- [11] 李恩波, 孙稚颖. 艾叶及其常见混伪品的分子鉴定[J]. 中国药房, 2013, 24(43):4037
LI EB, SUN ZY. Molecular identification of *Artemisia argyi* and its adulterants[J]. *China Pharm*, 2013, 24(43):4037
- [12] 熊婧, 李慧勇, 李广生, 等. 化学计量学在中药色谱分析中的应用探讨[J]. 药物分析杂志, 2021, 41(10):1681
XIONG J, LI HY, LI GS, *et al.* Discussion on the application of chemometrics in chromatographic analysis of traditional Chinese medicine[J]. *Chin J Pharm Anal*, 2021, 41(10):1681
- [13] 彭琴, 邹福贤, 许少华, 等. 基于化学计量学的指纹图谱法鉴别金线莲及其混伪品[J]. 药学研究, 2021, 40(3):149
PENG Q, ZHOU FX, XU SH, *et al.* Identification of *Anoectochilus roxburghii* and its adulterants using HPLC fingerprints based on chemometrics[J]. *J Pharm Res*, 2021, 40(3):149
- [14] 曹秀楠, 张蒙蒙, 王维, 等. 基于 UPLC 指纹图谱的防己及其混伪品鉴别研究[J]. 中药材, 2022, 45(11):2671
CAO XN, ZHANG MM, WANG W, *et al.* Identification of *Stephaniae Tetrandrae Radix* and adulterants based on UPLC fingerprint[J]. *J Chin Med Mater*, 2022, 45(11):2671
- [15] 张蒙, 于丹, 崔磊, 等. 艾叶挥发油化学成分研究进展[J]. 现代生物医学进展, 2019, 19(4):777
ZHANG M, YU D, CUI L, *et al.* Research progress of chemical constituents of volatile oil from *Artemisia argyi*[J]. *Prog Mod Biomed*, 2019, 19(4):777
- [16] SANDUSKY P, RAFTERY D. Use of semiselective TOCSY and the pearson correlation for the metabonomic analysis of biofluid mixtures: application to urine[J]. *Anal Chem*, 2005, 77(23):7717
- [17] WANG Z, LUO P, CHENG L, *et al.* Hapten-antibody recognition studies in competitive immunoassay of α -zearalanol analogs by computational chemistry and pearson correlation analysis[J]. *J Mol Recognit*, 2011, 24(5):815
- [18] WOLD S, ANTTI H, LINDGREN F, *et al.* Orthogonal signal correction of near-infrared spectra [J]. *Chemometr Intell Lab*, 1998, 44(1):175
- [19] TRYGG J, WOLD S. Orthogonal projections to latent structures (O-PLS)[J]. *J Chemom*, 2002, 16(3):119
- [20] BIAGIONI JD, ASTLING PD, GRAF P, *et al.* Orthogonal projection to latent structures solution properties for chemometrics and systems biology data[J]. *J Chemom*, 2011, 25(9):514
- [21] YIN S, WANG G, GAO H. Data-driven process monitoring based on modified orthogonal projections to latent structures[J]. *IEEE Trans Control Syst Technol*, 2016, 24(4):1480

(本文于 2023 年 8 月 31 日收到)