

DOI: 10.19812/j.cnki.jfsq11-5956/ts.20241121001

引用格式: 罗晓宏, 王楠希, 陈红娟, 等. 基于化学计量学和近红外光谱法的油莎豆产地溯源[J]. 食品安全质量检测学报, 2025, 16(4): 178–184.

LUO XH, WANG NX, CHEN HJ, et al. Traceability of geographical origin of *Cyperus esculentus* based on chemometrics and near infrared spectroscopy [J]. Journal of Food Safety & Quality, 2025, 16(4): 178–184. (in Chinese with English abstract).

# 基于化学计量学和近红外光谱法的 油莎豆产地溯源

罗晓宏<sup>1</sup>, 王楠希<sup>1</sup>, 陈红娟<sup>1</sup>, 庄绪会<sup>1</sup>, 孙恬恬<sup>1</sup>, 肖巾秀<sup>2</sup>, 令狐羽珮<sup>2</sup>, 杨永坛<sup>1\*</sup>

(1. 国家粮食和物资储备局科学研究院, 北京 100037; 2. 北京农学院食品科学与工程学院, 北京 102200)

**摘要:** **目的** 采用近红外光谱技术对油莎豆进行分析, 并应用化学计量学中识别模式对油莎豆进行产地溯源。**方法** 采用近红外光谱法结合化学计量学软件, 对来自河北、湖南、山东、新疆、云南等地 408 份油莎豆样品进行产地溯源, 分别采用多元散射校正、多量标准化或多量标准化耦合去趋势算法 3 种光谱预处理方法和支持向量机(support vector machine, SVM)、簇类独立分类(soft independent modeling of class analogy, SIMCA)、正交偏最小二乘判别(orthogonal partial least squares discriminant analysis, OPLS-DA)、偏最小二乘判别(partial least squares discriminant analysis, PLS-DA)、和 K 最近邻算法(K-nearest neighbor algorithm, KNN)等 5 种识别模式进行产地识别。**结果** SVM、SIMCA、OPLS-DA、PLS-DA 和 KNN 等 5 种模式的建模识别率分别为 91.89%、94.47%、62.37%、65.32%和 100.00%。选择 KNN 作为产地识别模型, 分析不同预处理方法、数据预处理及样本距离对模型预测结果稳定性的影响, 筛选出最优模型参数。选用多元散射校正光谱预处理方式, 在 UV 标度化、Pareto 标度化、自动标度化或中心化任一种数据预处理条件下, 样本距离选用街区距离, 测试集识别率能达到 100.00%。**结论** 近红外光谱结合 KNN 模式的技术具有分析速度快、操作简单、样本预处理容易、无损、在线的定性定量分析等优点, 具有一定应用前景。

**关键词:** 油莎豆; 近红外光谱技术; K 最近邻算法; 产地溯源

## Traceability of geographical origin of *Cyperus esculentus* based on chemometrics and near infrared spectroscopy

LUO Xiao-Hong<sup>1</sup>, WANG Nan-Xi<sup>1</sup>, CHEN Hong-Juan<sup>1</sup>, ZHUANG Xu-Hui<sup>1</sup>,  
SUN Tian-Tian<sup>1</sup>, XIAO Jin-Xiu<sup>2</sup>, LINGHU Yu-Pei<sup>2</sup>, YANG Yong-Tan<sup>1\*</sup>

(1. Academy of National Food and Strategic Reserves Administration, Beijing 100037, China;  
2. Food Science and Engineering College, Beijing University of Agriculture, Beijing 102200, China)

**ABSTRACT: Objective** To analyze *Cyperus esculentus* by near infrared spectroscopy, and trace geographical

收稿日期: 2024-11-21

基金项目: 中央级公益性科研院所基本科研业务费项目(JY2408)

第一作者: 罗晓宏(1986—), 女, 硕士, 助理研究员, 主要研究方向为仪器分析方法开发及应用。E-mail: lxx@ags.ac.cn

\*通信作者: 杨永坛(1971—), 男, 博士, 研究员, 主要研究方向为食品加工与安全。E-mail: yyt@ags.ac.cn

origin of *Cyperus esculentus* by the identification model in chemometrics. **Methods** A total of 408 samples of *Cyperus esculentus* samples from Hebei, Hunan, Shandong, Xinjiang, and Yunnan were analyzed for provenance tracing using near-infrared spectroscopy and chemometric software, 3 kinds of spectral preprocessing methods including multiplicative scatter correction, standard normal variate transformation and standard normal variate transformation & detrending, were used respectively, and 5 kinds of recognition modes such as support vector machine (SVM), soft independent modeling of class analogy (SIMCA), orthogonal partial least squares discriminant analysis (OPLS-DA), partial least squares discriminant analysis (PLS-DA), and K-nearest neighbor algorithm (KNN) were used to identify the geographical origin. **Results** The modeling recognition rates of the 5 kinds of modes including SVM, SIMCA, OPLS-DA, PLS-DA, and KNN were 91.89%, 94.47%, 62.37%, 65.32%, and 100.00% respectively. The KNN was selected as the origin identification model, and the impact of different preprocessing methods, data preprocessing and sample distance on the stability of the model prediction results were analyzed in order to select the optimal model parameters. The prediction set recognition rate could reach 100.00% by using multiplicative scatter correction spectral preprocessing method, one of data preprocessing methods including UV, Pareto, automatic, or centering, and block distance as the sample distance. **Conclusion** The technology of near infrared spectroscopy combined with KNN mode has the advantages of fast analysis speed, simple operation, easy sample pretreatment, non-destructive, on-line qualitative and quantitative analysis, etc., and has a certain application prospect.

**KEY WORDS:** *Cyperus esculentus*; near infrared spectroscopy; K-nearest neighbor algorithm; geographical origin

## 0 引言

油莎豆(*Cyperus esculentus*)又名油莎草、虎坚果、铁荸荠等<sup>[1-2]</sup>, 可食用或作为油料作物, 营养物质丰富, 其地下块茎中含淀粉 25%~40%、油脂 20%~30%、糖 15%~20%、蛋白质 5%~10%、膳食纤维 8%~10%、维生素 C 和 E 8~14 mg/100 g, 以及丰富的矿物质(钾、磷、钠、钙、镁), 是一种的多用途经济作物<sup>[3]</sup>。油莎豆中的脂肪酸、淄醇和多种酚类化合物具有抗氧化、降血糖血脂、抗菌等生物活性<sup>[4]</sup>, 以及含有必需氨基酸的高活性蛋白<sup>[5]</sup>, 正在受到食品、营养保健和畜牧生产领域的广泛关注, 有开发成为功能食品潜力。

油莎豆原产于非洲, 自 1960 年引入我国, 近年来种植规模和加工量逐年增加, 在我国新疆、河北、湖南、山东、吉林和广东等地都有种植, 油莎豆品质由于受到基因遗传、气候和土壤等因素的影响, 营养成分会发生变化<sup>[6]</sup>, 不同产地的油莎豆品质不同, 导致市场的价格也会参差不齐, 一些不法商贩会趁机掺杂, 扰乱市场交易秩序, 因此, 建立油莎豆产地溯源就尤为重要。

目前, 产地溯源的方法有液相色谱法、气相色谱法和液相色谱-串联质谱法等<sup>[7-9]</sup>, 虽然上述方法测试准确度高, 但是检测成本高, 需要大量化学试剂和前处理, 很难实现快速检测。而近红外光谱在 12000~4000  $\text{cm}^{-1}$ , 可以检测样品中的 -CH、-OH 和 -NH 等化学键, 但由于仪器本身、环境、样品特性等因素的影响, 采集的光谱常出现噪声干扰、基线漂移、谱峰重叠等现象<sup>[10-11]</sup>, 为此, 需要引入化学计

量学方法, 并通过化学计量学方法分析其“指纹”特征, 可以实现全光谱和特定区域光谱的定性和定量分析, 从而实现产品的产地溯源。其中近红外光结合化学计量学技术因具有快速、高效和无损的检测特点, 适应当今市场快检需求, 已经在鳄嘴花、玉米油、黄精、海南山茶油、辣椒、柠檬和枸杞等<sup>[12-18]</sup>医药和农产品领域的产地鉴别中应用。

本研究拟利用傅里叶变换近红外漫反射光谱法结合化学计量学软件, 以 5 个省产地溯源识别准确率为指标, 采用 5 种识别模式算法, 建立产地与指纹图谱之间的关联, 实现油莎豆的近红外光谱定性分析模型, 提高产地溯源智能化和自动化分析水平。该技术不仅为油莎豆品牌保护, 而且也为产地溯源的应用研究提供理论依据。

## 1 材料与方 法

### 1.1 仪器与设备

MB3600 傅立叶近红外光谱仪(加拿大 ABB 公司); JP-300A 高速多功能粉碎机(永康市久品工贸有限公司); BPG-9056A 电热鼓风干燥箱(上海一恒科学仪器有限公司); Chempattern 化学计量学软件[科迈恩(北京)科技有限公司]。

### 1.2 样品采集与处理

#### 1.2.1 样品采集

收集新疆、云南、湖南、河北、山东等地油莎豆。时间集中在 2023 年 10 月 13 日和 2023 年 11 月 10 日期间, 每

个采样地点进行区域设计,来进行随机样品采集,每个采样点收集 3 份 1 kg 的油莎豆,记录采样信息并整理编号。

### 1.2.2 样品预处理方法

将采集的油莎豆按照地区整理,按产地进行编号,用超纯水对前处理后的油莎豆籽粒进行流动水清洗,清洗后晾晒,烘干至水分 5% 以下,再进行粉碎处理,过 50 目不锈钢筛,得到油莎豆粉末样本,存入密封袋,于 -20 °C 保存,所有样本采用统一处理方式。

### 1.2.3 近红外光谱采集方法

利用近红外光谱仪漫反射,设备开机预热 30 min,钢砷检测器(InAsTEC)增益为 80.36,实验所处的环境温度为 (25±1) °C,相对湿度为 20%~30%,将油莎豆粉末样本均匀的放入 20 mL 玻璃小瓶中,然后安放于旋转支架,启动光谱仪进行近红外光谱扫描,光谱采集设置参数分辨率 32 cm<sup>-1</sup>,扫描次数 64,波长范围 10000~4000 cm<sup>-1</sup>。托盘中样本紧密程度保持基本一致,扫描过程中要避免样品间的交叉污染,每次扫描后要用无尘纸擦净样品杯,每个样品均扫描背景。

## 1.3 数据处理

采用 Chempattern 2020(x86) 计量学软件绘制红外光谱图,并利用软件中光谱预处理和不同算法的模式识别模块对近红外实验数据进行建模分析。其中光谱数据处理时常用的预处理方法很多,如多元散射校正、多量标准化(standard normal variate transformation, SNV)和多量标准化耦合去趋势算法(standard normal variate transformation & detrending, SNV+Detrending)等<sup>[19]</sup>。采用计量学软件中模式识别模块计算训练集和测试集识别结果,将数据结果导出至 Excel 2021 中,应用 EXACT 函数判别识别结果真假,用 COUNTIFS 函数计算识别正确结果个数。识别分类图采用计量学软件 Chempattern 制作,采用 Excel 2021 绘制识别率表,采用编程语言 Python(版本 3.9.13)绘制接收者操作特征(receiver operating characteristic, ROC)曲线和统计出曲线下面积(area under the curve, AUC)值。

模型预测正确率计算公式(1)如下。

$$\text{预测识别率}/\% = \frac{\text{预测正确产地数量}}{\text{所有预测产地数量}} \times 100\% \quad (1)$$

K 最近邻算法(K-nearest neighbor algorithm, KNN)模式中涉及的欧式距离、街区距离计算公式(2)和(3)如下。

N 维空间的两个点 A (X<sub>1</sub>, X<sub>2</sub>, …, X<sub>N</sub>) 和 B (Y<sub>1</sub>, Y<sub>2</sub>, …, Y<sub>N</sub>)。

$$d(A, B) = \sqrt{\sum_{i=1}^N (X_i - Y_i)^2} \quad (2)$$

N 维空间的两个点 A (X<sub>1N</sub>, Y<sub>1N</sub>) 和 B (X<sub>2N</sub>, Y<sub>2N</sub>)。

$$d(A, B) = \sum_{i=1}^N |X_{2i} - X_{1i}| + \sum_{i=1}^N |Y_{2i} - Y_{1i}| \quad (3)$$

## 2 结果与分析

### 2.1 试验样品的选取与划分

采用留一交叉验证<sup>[20]</sup>,该方法是将数据训练集中的每一个样本都作为单独的验证集,而其余的样本作为训练集进行模型训练。将随机选择 2/3 的样本用于训练集,1/3 的样本用于测试集。总样本数 408 个,其中训练集样本为 271 个,测试集 137 个,详见表 1。

表 1 样品数据集的分类结果  
Table 1 Classification results of sample data sets

编号	分类名称	归属	总样本 /个	训练集 /个	测试集 /个
1	河北	训练集	30	30	-
2	湖南	训练集	15	15	-
3	山东	训练集	21	21	-
4	新疆	训练集	175	175	-
5	云南	训练集	30	30	-
6	河北预测	测试集	15	-	15
7	湖南预测	测试集	9	-	9
8	山东预测	测试集	9	-	9
9	新疆预测	测试集	89	-	89
10	云南预测	测试集	15	-	15
合计			408	271	137

注: -代表没有参与相应数据集。

### 2.2 近红外原始光谱图

采集 5 个省共 408 个样本的近红外光谱图,由图 1 可以看出,5 个产地的油莎豆官能团出峰位置在谱图上非常接近,共有特征峰的波长为 4212.2236、4289.3706、4505.3823、4983.6934、5384.8579、5739.7730、7359.8203 cm<sup>-1</sup>等,分别对应 C-H 伸缩振动、C-H 伸缩振动或 N-H 伸缩振动、C-H 相关的振动、O-H 伸缩振动或 C-H 伸缩振动、C-H 伸缩振动或 N-H 伸缩振动、C-H 伸缩振动或 O-H 伸缩振动、C-H 或 N-H 伸缩振动等,样品出峰位置基本一致,仅仅依靠肉眼无法实现区分,这可能是由于不同产地油莎豆化学组成非常相似,另外人为操作不当、仪器误差等都会对光谱产生影响,样品的近红外光谱信号也会受到杂散光、噪声、基线漂移等因素影响,以至于最后分析结果较差<sup>[21]</sup>。因此,在运用近红外光谱联合化学计量学判别模型进行分析之前,选择分析波长 9000~4000 cm<sup>-1</sup>,需要对原始

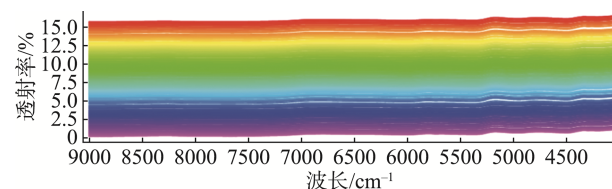


图 1 近红外原始光谱图  
Fig.1 Near-infrared original spectrogram

近红外光谱图进行预处理, 即采用数学统计的方法消除或减弱一些干扰因素对光谱的影响, 以便于更好的对近红外光谱进行解析说明, 以此提高模型分析的可靠性和准确性。

### 2.3 不同模型比较与验证

#### 2.3.1 模型建立

将原始光谱数据导入 Chempattern 软件中, 分为训练组和测试组, 支持向量机 (support vector machine, SVM)<sup>[22]</sup>、簇类独立分类 (soft independent modeling of class analogy, SIMCA)<sup>[23]</sup>、正交偏最小二乘判别 (orthogonal partial least squares discriminant analysis, OPLS-DA)<sup>[24]</sup>、偏最小二乘判别 (partial least squares discriminant analysis, PLS-DA)<sup>[25]</sup> 和 KNN<sup>[26]</sup> 等 5 种识别模式对油莎豆产地溯源建模, 具体操作如下: 分别把河北样本赋值为 1, 湖南样本赋值为 2, 山东样本赋值为 3, 新疆样本赋值为 4, 云南样本赋值为 5, 即测定值位于赋值区间, 就把该样品产地归类到赋值的省份, 详见图 2。5 种模型训练集识别率详见表 2,

在红外谱图未经光谱预处理情况下, 数据预处理为自动标度化, 5 种模式训练集识别效果不一, 训练集识别率最好的是 KNN, 其次是 SIMCA, OPLS-DA 建模识别效果最差, 其中支持向量机建模识别率为 91.89%, SIMCA 识别率为 94.47%, OPLS-DA 和 PLS-DA 识别率偏低, 可能原因模型样品中没有存在很好的线性关系, PLS-DA 属于多元线性识别方法, 不能很好地解决非线性问题, PLS-DA 模式更适合基中红外光谱的不同省份样品的识别<sup>[27]</sup>。采用留一交叉验证方法, 确定 KNN 的  $k$  值参数为 2, KNN 训练集识别率为 100.00%, 验证集识别率为 100.00%。KNN 模式取得更好识别结果的原因可能是近红外光谱峰重叠和噪声干扰, 造成了近红外数据存在严重的非线性, 而 KNN 分类算法是一种惰性机器学习算法, 具备增量学习特性, 适用于非线性分类问题<sup>[28]</sup>。另外, KNN 模式已在茉莉花、大米识别中取得很好效果<sup>[29-30]</sup>, 本研究中 KNN 也被选为油莎豆产地溯源识别模式。

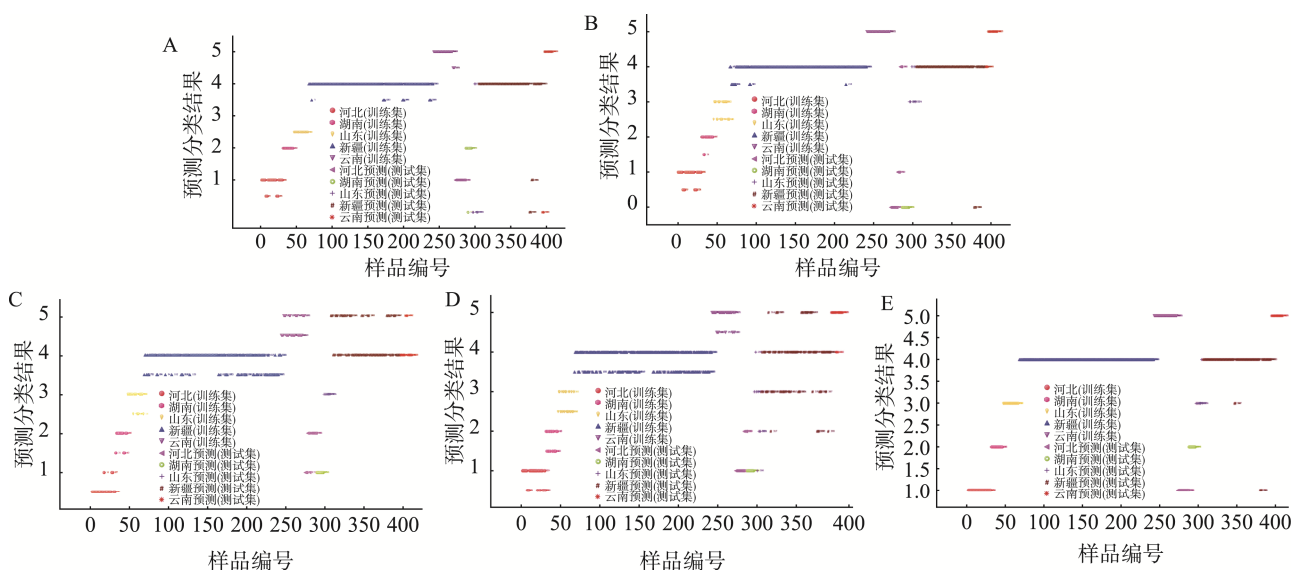


图 2 SVM (A)、SIMCA (B)、OPLS-DA (C)、PLS-DA (D)、KNN (E) 识别结果  
Fig.2 Results of SVM (A), SIMCA (B), OPLS-DA (C), PLS-DA (D), KNN (E) classification

表 2 5 种模式建模识别结果

Table 2 Results of identification based on 5 kinds of modeling patterns to model

模式	光谱预处理	数据预处理	训练集识别率/%	验证集识别率/%
SVM	无	自动标度化	91.89	85.24
SIMCA	无	自动标度化	94.47	90.77
OPLS-DA	无	自动标度化	62.37	66.42
PLS-DA	无	自动标度化	65.32	83.39
KNN	无	自动标度化	100.00	100.00

#### 2.3.2 模型优化

选择 KNN 作为识别模式, 选择 MSC、SNV 和 SNV+Detrending 等 3 种光谱预处理方式和 1 种未经处理方式, 研究了不同光谱预处理对模型识别率影响, 结果表明, SNV+Detrending 光谱预处理方式对模式识别率结果最小, 其他处理方式对 KNN 模型识别率影响一致, 见表 3。

KNN 模式下, 选择 MSC 作为光谱预处理方式, 比较 5 种数据预处理对模型的影响, 由表 4 可知, 5 种数据预处理对河北、湖南和云南的识别率能达到 100.00%, 整体而言, 均一化的数据前处理会降低模型识别率, 其他 4 项识别结果一致。在优化模型数据预处理时, 可以优选 UV 标度化、Pareto



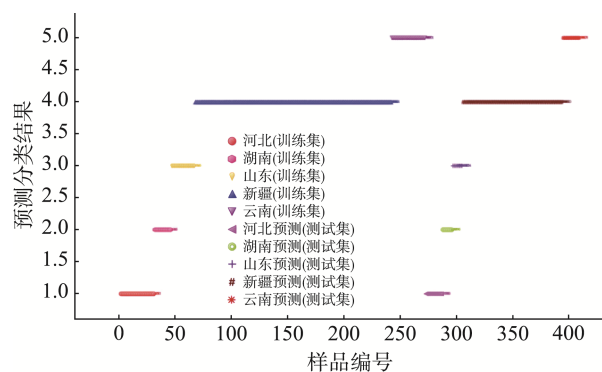


图3 KNN 最优识别结果

Fig.3 Optimal results of KNN classification

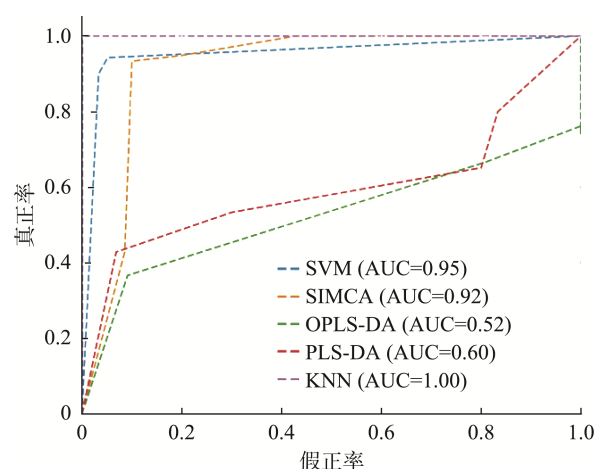


图4 5种算法在油莎豆产地溯源中 ROC 曲线和 AUC 值

Fig.4 ROC curves and AUC values for geographical origin of *Cyperus esculentus* by 5 kinds of algorithms

### 3 结论

近红外光谱结合化学计量学软件,对河北、湖南、山东、新疆和云南等5个省份油莎豆进行产地溯源研究,确定了KNN为产地溯源识别模型,优化该模式算法的参数,确定最优模型参数为:光谱预处理方式选择MSC,数据预处理方式UV标度化、Pareto标度化、自动标度化或中心化中任一种,样本距离选择街区距离,模型识别率都能达到100.00%。KNN模式下样本距离中街区距离要优于其他距离识别效果,涉及到算法机制需要后续深入研究。近红外光谱结合KNN模式的技术用于油莎豆产地溯源,极大提高光谱的数据处理与建模分析水平,这种技术具有分析速度快、操作简单、样本预处理容易、无损、在线的定性定量分析等优点,具有一定应用前景。

### 参考文献

- [1] 郭庆. 油莎豆产业发展展望[J]. 农业技术与装备, 2023(8): 51-53.  
GUO Q. Prospects for the development of *Cyperus beans* industry [J]. Agricultural Technology & Equipment, 2023(8): 51-53.
- [2] 王瑞元, 王晓松, 相海. 一种多用途的新兴油料作物——油莎豆[J]. 中国油脂, 2019, 44(1): 1-4.  
WANG RY, WANG XS, XIANG H. A versatile emerging oilseed crop-tiger nut [J]. China Oils and Fats, 2019, 44(1): 1-4.
- [3] 杨向东, 李子勇. 我国油莎豆产业发展现状、潜力及对策[J]. 中国油料作物学报, 2022, 44(4): 712-717.  
YANG XD, LI ZY. Tiger nut industry in China: Current status of development, potential and adaptive suggestions [J]. Chinese Journal of Oil Crop Sciences, 2022, 44(4): 712-717.
- [4] 丑义宣, 李柯洁, 闵琰, 等. 油莎豆的营养成分、生物活性及其应用研究进展[J]. 食品安全质量检测学报, 2023, 14(15): 222-230.  
CHOU YX, LI KJ, MIN Y, et al. Research progress in nutritional composition, biological and application of *Cyperus esculentus* [J]. Journal of Food Safety & Quality, 2023, 14(15): 222-230.
- [5] LOPEZ-CORTES I, SALAZAR-GARCIA D, MALHERIRO R, et al. Chemometrics as a tool to discriminate geographical origin of *Cyperus esculentus* L. based on chemical composition [J]. Industrial Crops and Products, 2013, 51: 19-25.
- [6] 刘玉兰, 王小宁, 舒垚, 等. 不同产地油莎豆性状及组成分析研究[J]. 中国油脂, 2020, 45(8): 125-129.  
LIU YL, WANG XN, SHU Y, et al. Character and composition of *Cyperus esculentus* from different origins [J]. China Oils and Fats, 2020, 45(8): 125-129.
- [7] 李籽杉, 杜泽飞, 杨晓, 等. 基于超高效液相色谱与红外光谱的金银花药材产地溯源[J]. 医药导报, 2023, 42(7): 1067-1073.  
LI ZS, DU ZF, YANG X, et al. Geographic origin tracing of *Lonicera japonica* by FTIR and UPLC [J]. Herald of Medicine, 2023, 42(7): 1067-1073.
- [8] ZAPPI A, MELUCCI D, SCARAMAGLI S, et al. Botanical traceability of unifloral honeys by chemometrics based on head-space gas chromatography [J]. European Food Research and Technology, 2018, 244(12): 2149-2157.
- [9] 贺光云, 侯雪, 闫志农, 等. 基于超高效液相色谱-四极杆串联飞行时间质谱的绿茶产地溯源研究[J]. 农产品质量与安全, 2021(5): 63-68.  
HE GY, HOU X, YAN ZN, et al. The geographical origin traceability of green tea by ultra-high performance liquid chromatography-quadropole time-of-flight mass spectrometry [J]. Quality and Safety of Agro-products, 2021(5): 63-68.
- [10] 白铁成, 王涛, 陈佑启, 等. 南疆沙尘区骏枣叶片水分含量检测的近红外光谱预处理方法对比(英文)[J]. 光谱学与光谱分析, 2019, 39(4): 1323-1328.  
BAI TC, WANG T, CHEN YQ, et al. Comparison of near-infrared spectrum pretreatment methods for jujube leaf moisture content detection in the sand and dust area of southern xinjiang [J]. Spectroscopy and Spectral Analysis, 2019, 39(4): 1323-1328.
- [11] 李俊杰, 王学重, 王焱宇, 等. 近红外光谱用于硝基苯纯度快速定量分析方法研究[J]. 应用化工, 2022, 51(1): 277-280.  
LI JJ, WANG XZ, WANG YY, et al. Rapid quantitative analysis of nitrobenzene purity by near infrared spectrum [J]. Applied Chemical Industry, 2022, 51(1): 277-280.
- [12] YU FY, MA JF, QI Y, et al. Geographical traceability of *Clinacanthus nutans* with near-infrared spectroscopy and chemometrics [J]. American Journal of Analytical Chemistry, 2022, 13(2): 63-77.

- [13] ZHANG H, HU X, LIU L, *et al.* Near infrared spectroscopy combined with chemometrics for quantitative analysis of corn oil in edible blend oil [J]. *Spectrochimica acta, Part A. Molecular and Biomolecular Spectroscopy*, 2022, 270: 120841.
- [14] ZHANG J, ZHANG J, ZHONG Z. ATR-FTIR and NIR spectra combined with chemometrics and convolutional neural networks for identification of polygonati rhizome [J]. *Journal of Molecular Structure*, 2024, 1318: 139449.
- [15] DENG Z, FU J, YANG M, *et al.* Geographical origin identification of Hainan camellia oil based on fatty acid composition and near infrared spectroscopy combined with chemometrics [J]. *Journal of Food Composition and Analysis*, 2024, 125: 105730.
- [16] CASTELL A, ARROYO-MANZANARES N, LOPEZ-GARCIA I, *et al.* Authentication strategy for paprika analysis according to geographical origin and study of adulteration using near infrared spectroscopy and chemometric approaches [J]. *Food Control*, 2024, 161: 1–11.
- [17] RUGGIERO L, AMALFITANO C, DIVAIO C, *et al.* Use of near-infrared spectroscopy combined with chemometrics for authentication and traceability of intact lemon fruits [J]. *Food Chemistry*, 2022, 375: 131822.
- [18] SHEN TT, ZOU XB, SHI JY, *et al.* Determination geographical origin and flavonoids content of goji berry using near-infrared spectroscopy and chemometrics [J]. *Food Analytical Methods*, 2016, 9(1): 68–79.
- [19] 孙旭东, 廖琪城, 韩熹, 等. 基于电磁振动上料的茶梗和昆虫异物近红外光谱和荧光图像在线检测研究[J]. *光谱学与光谱分析*, 2023, 43(1): 100–106.
- SUN XD, LIAO QC, HAN X, *et al.* Research on online detection of tea stalks and Insect foreign bodies by near-infrared spectroscopy and fluorescence image combined with electromagnetic vibration feeding [J]. *Spectroscopy and Spectral Analysis*, 2023, 43(1): 100–106.
- [20] 白鹏利, 王钧, 尹焕才, 等. 基于偏最小二乘留一交叉验证法的近红外光谱建模样品选择方法的研究[J]. *食品安全质量检测学报*, 2017, 8(1): 182–186.
- BAI PL, WANG J, YIN HC, *et al.* Study on the sample selection methods for building calibration model by near infrared spectroscopy based on partial least squares-leave one out-cross validation [J]. *Journal of Food Safety & Quality*, 2017, 8(1): 182–186.
- [21] TORRELLA IC, GUAMIS B, TRUJILLO JA, *et al.* Characterization and comparison of tigernuts (*Cyperus esculentus* L.) from different geographical origin: Physico-chemical characteristics and protein fractionation [J]. *Industrial Crops and Products*, 2015, 65: 406–414.
- [22] YUAN WD, JIANG HZ, SUN MM, *et al.* Geographical origin identification of Chinese tomatoes using long-wave fourier-transform near-infrared spectroscopy combined with deep learning methods [J]. *Food Analytical Methods*, 2023, 16: 664–676.
- [23] 李跑, 谭惠珍, 谢淑娥, 等. 基于近红外光谱技术有监督模式识别的青皮产地溯源分析[J]. *轻工学报*, 2024, 39(2): 54–59.
- LI P, TAN HZ, XIE SE, *et al.* Traceability analysis of *Pericarpium citri reticulatae viride* origin based on near infrared spectroscopy technology and supervised pattern recognition [J]. *Journal of Light Industry*, 2024, 39(2): 54–59.
- [24] 高美美, 黄建猷, 翟雅南, 等. 指纹图谱和多成分定量结合化学模式识别法评价不同产地杉木叶质量[J]. *药物分析杂志*, 2024, 44(5): 882–892.
- GAO MM, HUANG JY, ZHAI YN, *et al.* Quality evaluation of leaves of *Cunninghamia lanceolata* from different habitats by fingerprint and multi-component qualification combined with chemical pattern recognition [J]. *Chinese Journal of Pharmaceutical Analysis*, 2024, 44(5): 882–892.
- [25] 宋雪健, 钱丽丽, 周义, 等. 近红外漫反射光谱技术对小米产地溯源的研究[J]. *食品研究与开发*, 2017, 38(11): 134–139.
- SONG XJ, QIAN LL, ZHOU Y, *et al.* Study on traceability of millet origin by near infrared diffuse reflectance spectroscopy [J]. *Food Research and Development*, 2017, 38(11): 134–139.
- [26] 刘庭恺, 胡子康, 龙婉君, 等. 基于近红外和中红外光谱的杜仲产地溯源[J]. *化学试剂*, 2022, 44(7): 952–959.
- LIU TK, HU ZK, LONG WJ, *et al.* Geographical origin traceability of *Eucommia cortex* based on near and mid infrared spectroscopy [J]. *Chemical Reagents*, 2022, 44(7): 952–959.
- [27] LU CH, XIANG BR, HAO G, *et al.* Rapid detection of melamine in milk powder by near infrared spectroscopy [J]. *Journal of Near Infrared Spectroscopy*, 2009, 17(2): 59–67.
- [28] 崔萌, 段宝忠, 程蕾, 等. 基于近红外光谱结合化学计量学方法的山里红产地溯源分析[J]. *中草药*, 2024, 55(14): 4897–4906.
- CUI M, DUAN BZ, CHENG L, *et al.* Geographical origin traceability of *Crataegus pinnatifida* var. major based on near-infrared spectroscopy combined with chemometrics [J]. *Chinese Traditional and Herbal Drugs*, 2024, 55(14): 4897–4906.
- [29] 张淑芳, 雷蕾, 雷顺新, 等. 近红外漫反射光谱的茉莉花产地溯源[J]. *光谱学与光谱分析*, 2023, 43(11): 3389–3395.
- ZHANG SF, LEI L, LEI SX, *et al.* Traceability of geographical origin of *jasmine* based on near infrared diffuse reflectance spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2023, 43(11): 3389–3395.
- [30] 刘晓欢, 刘翠玲, 孙晓荣, 等. 基于傅里叶红外光谱技术的大米产地溯源快速判别方法研究[J]. *食品科技*, 2021, 46(4): 244–249.
- LIU XH, LIU CL, SUN XR, *et al.* Study on fast identification method of rice origin traceability based on fourier transform infrared spectroscopy technology [J]. *Food Science and Technology*, 2021, 46(4): 244–249.
- [31] GERTZ C, GERTZ A, MATTHAUS B, *et al.* A systematic chemometric approach to identify the geographical origin of olive oils [J]. *European Journal of Lipid Science and Technology*, 2019, 121(12): 190028.

(责任编辑: 安香玉 蔡世佳)