

有机与非有机大豆的营养成分分析及近红外鉴别

余新金¹, 熊倩¹, 甘蓓², 谢建华¹, 余强¹, 陈奕^{1*}

(1. 南昌大学食品科学与技术国家重点实验室, 南昌 330047; 2. 江西省产品质量监督检测院, 南昌 330029)

摘要: 目的 分析有机与非有机大豆的营养成分差异, 建立有机与非有机大豆的快速鉴别方法。**方法** 测定40份有机与非有机大豆的蛋白质、脂肪、水分、大豆苷和染料木苷含量并进行差异性分析; 采集大豆的近红外光谱, 基于主要营养成分以及近红外光谱结合不同预处理采用偏最小二乘法判别分析(partial least squares discriminant analysis, PLS-DA)建立有机与非有机大豆的快速鉴别模型; 通过PLS-DA模型筛选对于区分有机与非有机大豆有重要贡献的近红外光谱特征波段以建立最佳的快速鉴别模型。**结果** 有机与非有机大豆在蛋白质、水分和染料木苷含量上不存在显著性差异, 但有机大豆的脂肪和染料木苷含量显著高于非有机大豆。仅基于主要营养成分建立的鉴别模型显示出较低的模型参数和预测正确率, 而基于近红外光谱及其与主要营养成分结合建立的模型具有较高的模型参数和预测正确率。对于区分有机与非有机大豆有重要贡献的近红外光谱特征波段是5974~5372 cm⁻¹以及5064~4000 cm⁻¹, 这两个特征波段的近红外光谱吸收峰与脂肪、黄酮类物质以及蛋白质官能团的振动有关。基于近红外光谱特征波长结合Smoothing-SG预处理的模型显示出最优的模型参数(R^2X 为0.996, R^2Y 为0.735, Q^2 为0.677), 能够有效鉴别有机与非有机大豆。**结论** 由于种植模式不同, 有机与非有机大豆的营养成分存在差异, 有机大豆的营养价值高于非有机大豆。脂肪和染料木苷含量可以作为区分有机与非有机大豆的重要质量指标。近红外光谱技术在鉴别有机与非有机大豆上具有可行性, 为大豆的质量控制提供了技术基础。

关键词: 有机大豆; 非有机大豆; 营养成分; 近红外光谱; 鉴别

Nutrients analysis and near infrared identification of organic and non-organic soybeans

YU Xin-Jin¹, XIONG Qian¹, GAN Bei², XIE Jian-Hua¹, YU Qiang¹, CHEN Yi^{1*}

(1. State Key Laboratory of Food Science and Technology, Nanchang University, Nanchang 330047, China;
2. Jiangxi Provincial Product Quality Supervision Testing College, Nanchang 330029, China)

ABSTRACT: Objective To analyze the differences of nutrients between organic and non-organic soybeans and establish a rapid method for identifying organic and non-organic soybeans. **Methods** The protein, fat, moisture, daidzin and genistin content of forty organic and non-organic soybeans were determined and analyzed for the differences. The near infrared spectra of soybeans were collected, and a rapid discrimination model between organic and non-organic soybeans was established using partial least squares discriminant analysis (PLS-DA) based on the main nutrients and near infrared spectra combined with different pretreatments. The characteristic wavelengths of

基金项目: 食品科学与技术国家重点实验室自由探索资助课题项目(SKLF-ZZB-202115)

Fund: Supported by the Research Project of State Key Laboratory of Food Science and Technology (SKLF-ZZB-202115)

***通信作者:** 陈奕, 教授, 博士生导师, 主要研究方向为食品质量与安全。E-mail: chenyi-417@163.com

***Corresponding author:** CHEN Yi, Professor, State Key Laboratory of Food Science and Technology, Nanchang University, No.235, Nanjing East Road, Qingshanhu District, Nanchang 330047, China. E-mail: chenyi-417@163.com

near infrared spectra that contributed significantly to the distinction between organic and non-organic soybeans were screened by the PLS-DA model to establish the best rapid identification model. **Results** There was no significant difference between organic and non-organic soybeans in terms of protein, moisture and genistin content, but the fat and daidzin content of organic soybeans were significantly higher than those of non-organic soybeans. The identification model based only on major nutrients showed low model parameters and correct prediction, while models based on near infrared spectra and their combination with major nutrients had high model parameters and correct predictions. The characteristic wavelengths of near infrared spectra that contributed significantly to the distinction between organic and non-organic soybeans were $5974\text{--}5372\text{ cm}^{-1}$ and $5064\text{--}4000\text{ cm}^{-1}$. It was noted that the absorption peaks at these two characteristic wavelengths were related to the vibrations of functional groups of fat, flavonoids and protein. The model based on the characteristic wavelengths of near infrared spectra combined with Smoothing-SG pre-processing showed optimal model parameters (R^2X of 0.996, R^2Y of 0.735 and Q^2 of 0.677) to effectively identify organic and non-organic soybeans. **Conclusion** Due to different cultivation patterns, the nutrients of organic and non-organic soybeans are different, and the nutritional value of organic soybeans is higher than that of non-organic soybeans. The fat and daidzin content can be used as important quality indicators to distinguish between organic and non-organic soybeans. Near infrared spectra technology is practicable in identifying organic and non-organic soybeans, which provides a technical basis for the quality control of soybeans.

KEY WORDS: organic soybeans; non-organic soybeans; nutrients; near infrared spectra; identification

0 引言

大豆起源于中国,已有 3000 多年的种植历史,是食用油和植物蛋白的重要来源,具有很高的营养价值和经济效益^[1-3]。大豆含有丰富的蛋白质、脂肪、碳水化合物等营养物质,常被用于生产大豆油、豆腐、豆浆等产品^[4]。此外,由于富含异黄酮、酚酸、不饱和脂肪酸等生物活性物质,大豆被认为是一种有益于人类健康的食品^[5]。

近年来,随着消费者对食品安全和可持续发展问题的重视,有机农业已被公认为是非有机农业的一种可持续替代方式,其特点是不使用杀虫剂和合成肥料^[6]。消费者普遍认为有机食品比非有机食品更加安全健康,因此有机食品的市场份额和贸易量正在快速增长^[7]。一般来说,食品的营养和生物活性成分可能受到气候条件、土壤、栽培品种以及农业实践(有机或非有机种植模式)的影响^[6]。近年来,有机种植模式对食品化学成分的影响引起了研究人员的广泛关注。如 KAZIMIERCZAK 等^[8]研究了有机和非有机模式下生产的绿茶的化学成分差异,发现有机绿茶的黄酮醇和单宁酸含量明显更高,而非有机茶中的酚酸含量更高; BARRETO 等^[6]的研究结果表明有机农业模式能够有效改善黑豆的营养成分,提高酚类化合物含量; IGLESIAS-CARRES 等^[9]的研究表明与非有机种植相比,有机种植可以增加红葡萄中低聚原花青素和花青素的含量,但也伴随着游离黄烷醇和膳食纤维含量的降低; JIMÉNEZ 等^[10]研究发现有机初榨橄榄油的甜味和挥发性化合物含量呈现较高水平,而非有机初榨橄榄油的酚类含量更高。

基于以上的文献报道可以发现有机农业模式对不同食品中的不同组分存在不同的影响。然而,有机种植模式对大豆营养成分的影响还不清楚,因此有必要深入探究有机与非有机大豆的营养成分差异。

近红外光谱主要包含氢基团化学键伸缩振动倍频以及合频信息,具有检测成本低、重现性好、操作简单等特点^[11]。作为近年来快速发展起来的一种新型绿色无损检测手段,近红外技术已被广泛应用于质量分析和掺假控制等方面^[12-13]。有研究表明近红外技术在快速鉴别有机与非有机食品上具有可行性^[14-15]。目前,近红外技术已被应用于大豆化学成分的快速定量预测^[16-18]、转基因和非转基因大豆以及正常和病变大豆的快速鉴别^[19-20]。但利用近红外技术从营养成分角度快速鉴别有机与非有机大豆上是否具有可行性并不清楚。鉴于此,本研究拟通过国标法、光谱法和色谱法对比分析有机与非有机大豆主要营养成分和活性物质的差异,并基于主要营养成分和近红外光谱建立快速鉴别有机和非有机大豆的方法,以期为大豆的质量控制提供技术基础。

1 材料与方法

1.1 材料与试剂

1.1.1 材料

40 份大豆样品产自吉林、辽宁、黑龙江 3 个省份,其中有机大豆 18 份,非有机大豆 22 份,样本信息如表 1 所示。如图 1 所示,有机与非有机大豆在外观形状上没有明显的差异。实验前用粉碎机将大豆样品粉碎成粉末,置于干燥器中保存备用。

表 1 大豆样本信息
Table 1 Information of soybean samples

编号	产地	收获年份	类别
1~7	黑龙江	2018/2019	非有机
8~14	吉林	2018/2019	非有机
15~22	辽宁	2018/2019	非有机
23~29	黑龙江	2018/2019	有机
30~34	吉林	2018/2019	有机
35~40	辽宁	2018/2019	有机



图 1 有机(左)与非有机(右)大豆外观

Fig.1 Appearance of organic (left) and non-organic (right) soybeans

1.1.2 试剂

浓硫酸(纯度 95%~98%)、盐酸(纯度 36%~38%)、石油醚(30~60°C)、氢氧化钠、正己烷、五水硫酸铜、硫酸钾(分析纯)(西陇科学股份有限公司); 盐酸标准滴定溶液(0.1 mol/L, 深圳市博林达科技有限公司); 甲醇(色谱纯, 德国 Merck 公司); 二甲基亚砜(色谱纯, 上海阿拉丁生化科技股份有限公司); 大豆苷、染料木苷标准品(纯度 ≥98%, 上海源叶生物科技有限公司)。

1.2 仪器与设备

K9860 全自动凯氏定氮仪(山东海能科学仪器有限公司); Nicolet 5700 红外光谱仪(美国 Thermo Fisher Scientific 公司); KQ5200E 型超声波清洗器(昆山超声仪器有限公司); 博迅 HHS-21-6 电热恒温水浴锅(上海博迅实业有限公司医疗设备厂); DHG-9246A 型电热鼓风干燥箱(上海精宏实验设备有限公司); 1290 HPLC 高效液相色谱仪(美国 Agilent 公司); TU-1900 紫外可见分光光度计(北京普析通用仪器有限公司)。

1.3 实验方法

1.3.1 基本营养成分测定

参照 GB 5009.3—2016《食品安全国家标准 食品中水分的测定》中直接干燥法测定水分含量; 参照 GB 5009.6—2016《食品安全国家标准 食品中脂肪的测定》中索氏抽提法测定粗脂肪含量; 参照 GB 5009.5—2016《食品安全国家标准 食品中蛋白质的测定》中凯氏定氮法测定粗蛋白质含量。

1.3.2 大豆异黄酮含量测定

样品处理: 准确称取 0.2 g 大豆粉, 加入 5 mL 正己烷,

超声 10 min 后 4800 r/min 离心 5 min, 弃去上清液以除去脂肪。向沉淀中加入 3 mL 提取液(甲醇:二甲亚砜=1:1, V:V), 超声提取 45 min 后在室温下静置过夜。提取完成后离心, 收集上清液, 即为大豆提取液。取 1 mL 大豆提取液用 0.22 μm 滤膜过滤后通过高效液相色谱仪测定大豆苷和染料木苷的含量。

色谱条件: Kromasil C₁₈ 色谱柱(250 mm×4.6 mm, 5 μm); 流动相: 甲醇(A)-0.1%甲酸水(B); 进样量: 10 μL; 流速: 0.25 mL/min; 柱温: 30°C; 检测波长: 262 nm; 洗脱梯度: 0~20 min, 35%~55% A; 20~35 min, 55% A 等度洗脱。

1.3.3 近红外光谱采集

取等量大豆粉置于石英器皿中, 压实。采用 InGaAs 检测器, 选用 CaF₂ 分束器和积分球漫反射方式。采集在 10000~4000 cm⁻¹ 范围内的近红外光谱, 分辨率为 8 cm⁻¹, 扫描次数为 64 次。每个样品扫描 3 次得到平均谱图。

1.4 数据分析

每个样品平行分析 3 次, 实验数据采用 SPSS 21.0 软件进行显著性分析, $P < 0.05$ 为差异显著。近红外光谱数据的标准正态变量变换(standard normal variable, SNV)、多元散射校正(multivariate scattering correction, MSC)、一阶导数、二阶导数以及卷积(savitzky-golay, S-G)平滑 5 种预处理采用 Unscrambler X 4 软件进行。主成分分析(principal component analysis, PCA)和偏最小二乘法判别分析(partial least squares discriminant analysis, PLS-DA)采用 Simca 13.0 软件进行。

2 结果与分析

2.1 主要营养成分差异分析

2.1.1 蛋白质、脂肪、水分含量差异分析

蛋白质和脂肪是大豆中的主要营养成分, 二者的含量是衡量大豆品质好坏的重要品质指标^[15]。此外, 水分含量是衡量大豆品质的另一个重要指标, 其对大豆在运输、储存过程中的稳定性至关重要, 我国国家标准规定大豆的水分含量应不大于 13%^[3]。本研究测定了 40 份大豆中的蛋白质、脂肪和水分含量, 并对其结果进行统计分析。如表 2 所示, 有机与非有机大豆在蛋白质和水分含量上不存在显著性差异($P > 0.05$), 而有机大豆的脂肪含量(19.95%)显著高于非有机大豆(18.73%) ($P < 0.05$)。张玉荣等^[4]测定了来自阿根廷的高油大豆以及来自内蒙古的国产普通大豆的主要营养成分, 与之相比本研究测得的大豆脂肪和水分含量与其相近, 但蛋白质含量略高; 顾强等^[3]测定了来自巴西、阿根廷、美国、乌拉圭 4 个国家大豆的质量指标, 与之相比本研究测得的蛋白质、脂肪和水分含量均较低, 这些差异可能是大豆的产地不同所带来的。

表2 有机与非有机大豆主要营养成分比较($n=3$)Table 2 Comparison of main nutrient compositions of organic and non-organic soybeans ($n=3$)

营养成分	有机大豆		非有机大豆	
	含量范围	含量	含量范围	含量
蛋白质/%	28.90~36.00	34.06±1.55	31.45~36.62	34.06±1.45
脂肪/%	18.00~21.90	19.95±1.13*	15.98~20.50	18.73±1.37
水分/%	7.57~10.61	8.88±0.74	7.57~10.61	8.73±0.79

注: *表示该指标在有机与非有机大豆中存在显著性差异($P<0.05$),表3同。

2.1.2 大豆异黄酮含量差异分析

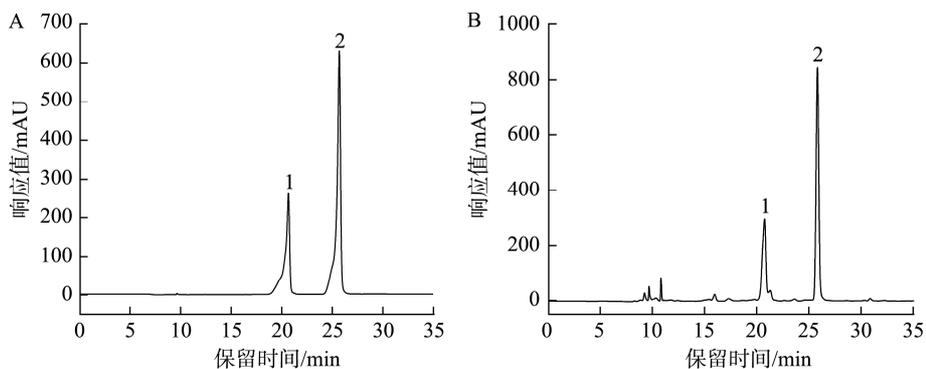
大豆异黄酮是存在于大豆中的次级代谢产物,具有抗胆固醇、调节肠道菌群、预防糖尿病和癌症等多种作用^[21]。大豆中天然存在的大豆异黄酮共有12种,分为游离型的甙元和结合型的糖甙两类,其中结合型的糖甙约占总量的97%~98%,大豆苷和染料木苷是大豆异黄酮中主要的糖甙型成分^[22-23]。本研究采用高效液相色谱法测定了40份大豆中大豆苷和染料木苷的含量。图2是大豆异黄酮标准品和大豆样品的液相色谱图,大豆苷和染料木苷在洗脱程序范围内可以很好地分离。大豆苷和染料木苷的标准曲线方程分别为 $Y=161.66X+177.97$ ($r^2=0.9998$)和 $Y=348.32X-51.46$ ($r^2=0.9993$),两种标准品的峰面积和浓度呈现出良好的线性关系。如表3所示,在有机和非有机大豆中染料木苷的含量均高于大豆苷的含量,这与费洪立等^[24]的研究结果是

一致的。有机与非有机大豆在染料木苷含量上不存在显著性差异($P>0.05$),但有机大豆的大豆苷含量(465.05 $\mu\text{g/g}$)显著高于非有机大豆(386.67 $\mu\text{g/g}$) ($P<0.05$),因此大豆苷含量可以作为区分有机与非有机大豆的重要质量指标之一。从实验结果可以发现,两种大豆异黄酮含量的值分布范围较广,这是由于大豆异黄酮的含量受环境因素影响很大,大豆异黄酮在特定环境的大豆中具有唯一性和特征性,这也使得大豆异黄酮常被作为鉴别大豆产地和品种的溯源指标^[24-25]。

2.2 有机与非有机大豆鉴别分析

2.2.1 有机与非有机大豆的近红外光谱图

如图3所示,有机与非有机大豆对应的近红外光谱在4000~10000 cm^{-1} 范围内的光谱吸收峰的位置基本一致,但吸收峰的强度存在差异。在5600~5000 cm^{-1} 的吸收峰归属于与脂肪有关的C-H官能团(第一泛音和HC=CH形式)的振动^[26]。在5000~4500 cm^{-1} 左右的光谱峰对应于与蛋白质有关的N-H和C-O的拉伸振动^[27]。5600~6000 cm^{-1} 左右的吸收峰归因于C-H的二级倍频的特征吸收,6000~7000 cm^{-1} 左右的吸收峰与O-H或N-H二级倍频的特征吸收有关,8300 cm^{-1} 左右的吸收峰与C-H拉伸三级倍频的特征吸收有关^[28]。此外,5700 cm^{-1} 附近的谱带已被认为是与黄酮有关的重要谱带^[29]。总体上有机与非有机大豆的近红外光谱相似性较高,通过肉眼分析光谱很难直接鉴别大豆的类型,需要结合化学计量学进一步对光谱数据进行挖掘分析。



注: 1. 大豆苷; 2. 染料木苷。

图2 大豆异黄酮标准品(A)和大豆样品(B)的高效液相色谱图

Fig.2 High performance liquid chromatograms of soybean isoflavones standard (A) and soybean sample (B)

表3 有机与非有机大豆的异黄酮含量

Table 3 Isoflavone content of organic and non-organic soybean

物质名称	有机大豆		非有机大豆	
	含量范围	含量	含量范围	含量
大豆苷($\mu\text{g/g}$)	273.00~837.98	468.05±133.94*	258.46~606.94	386.67±104.92
染料木苷($\mu\text{g/g}$)	173.46~977.94	529.66±233.42	265.27~992.20	467.71±187.67

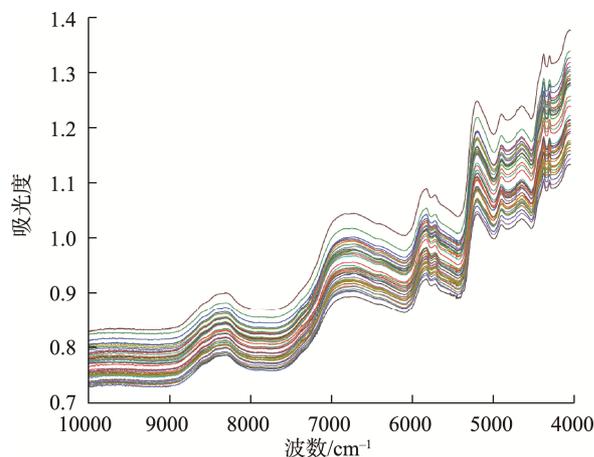


图 3 有机与非有机大豆近红外光谱图

Fig.3 Near infrared spectra of organic and non-organic soybeans

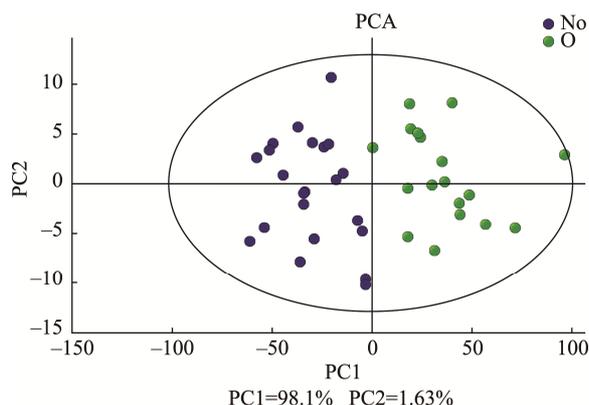
2.2.2 PCA 结果

PCA 是一种无监督的模式识别算法, 可以将原变量线性组合成为少数的新变量, 并且保证新变量能够不丢失原变量的数据特征, 从而达到数据降维的目的^[30]。本研究采用 PCA 对有机和非有机大豆的近红外光谱图进行聚类及初步检验, 结果如图 4 所示。前两个主成分的累积贡献率为 99.73%, 表明前两个主成分可以替代原变量 99.73% 的信息。从图 4 中可以非常直观地看出两种类型的大豆聚类为相互独立的 2 组, 分布于得分图的左右两侧, 表明前两个主成分可以将大豆分成两类。从 PCA 的结果可以发现有机大豆和非有机大豆的近红外光谱数据存在明显差异。

2.2.3 PLS-DA 结果

PLS-DA 是一种基于偏最小二乘回归算法的线性分类模型, 适用于多自变量并且强线性情况的分析和预测^[31]。除了样品本身的化学信息外, 收集的近红外光谱数据中还

可能包含其他不相关的信息和噪声, 比如仪器背景、噪声和杂散光等^[11]。光谱预处理是光谱数据分析中提取准确信息的关键步骤, 能够增强不同样本之间的细微差异, 提高光谱数据的可靠性。在 PCA 的基础上, 进一步利用有机与非有机大豆的主要营养成分以及近红外光谱结合多种预处理方式建立 PLS-DA 模型, 以实现两种类型大豆的快速鉴别预测。通常, 评价 PLS-DA 模型的参数包括 R^2X 、 R^2Y 、 Q^2Y 、交叉验证方差分析中的 $P_{CV-ANOVA}$ 值以及外部验证的正确率。其中, R^2X 和 R^2Y 分别表示模型对 X 和 Y 矩阵的解释能力, Q^2Y 表示模型的预测能力^[32]。将大豆样品按照 3:1 的比例随机分成校正集和验证集, 其中校正集 30 个样品 (17 个非有机大豆, 13 个有机大豆), 验证集 10 个样品 (5 个非有机大豆, 5 个有机大豆)。PLS-DA 模型建立后, 将 10 个验证集样本导入软件进行外部验证, 模型参数及验证结果如表 4 所示。在建立的 13 个模型中, 共有 5 个模型的外



注: No 为非有机大豆; O 为有机大豆。

图 4 有机与非有机大豆的 PCA 得分图

Fig.4 PCA score plot for organic and non-organic soybeans

表 4 基于近红外光谱和营养成分建立的 PLS-DA 模型的参数及验证结果

Table 4 Parameters and validation results of the PLS-DA model based on near infrared spectra and nutrient composition

数据源	光谱预处理方式	R^2X	R^2Y	Q^2	$P_{CV-ANOVA}$	正确率/%
近红外光谱	-	0.997	0.732	0.662	1.41×10^{-5}	100
	SNV	0.565	0.519	0.480	1.46×10^{-4}	90
	MSC	0.565	0.519	0.480	1.46×10^{-4}	90
	一阶导数	0.241	0.994	0.769	6.85×10^{-5}	90
	二阶导数	0.195	0.998	0.737	0.72	60
	S-G 平滑	0.980	0.713	0.663	4.27×10^{-7}	100
营养成分	-	0.191	0.309	0.0912	0.27	60
	-	0.977	0.713	0.663	4.25×10^{-7}	100
营养成分结合 近红外光谱	SNV	0.563	0.519	0.480	1.45×10^{-4}	90
	MSC	0.563	0.519	0.480	1.45×10^{-4}	90
	一阶导数	0.240	0.994	0.767	6.68×10^{-5}	100
	二阶导数	0.195	0.999	0.740	0.71	60
	S-G 平滑	0.977	0.713	0.663	4.25×10^{-7}	100

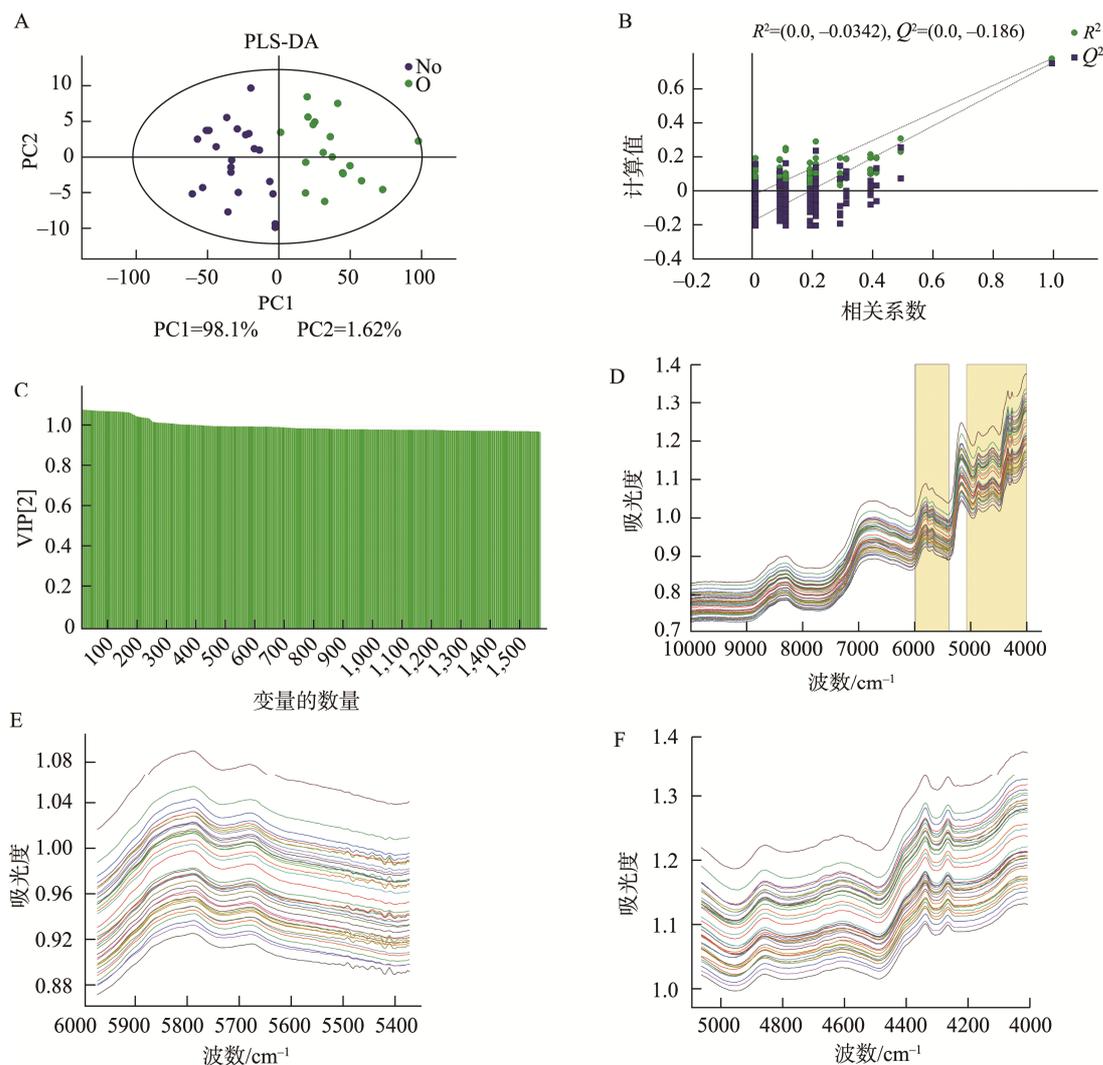
注: -表示无预处理; 营养成分指蛋白质、脂肪、水分、大豆苷、染料木苷含量。

部验证正确率达到了 100%。与基于原始近红外光谱建立的模型相比, S-G 平滑预处理对模型的鉴别能力没有明显的影响, 而 SNV、MSC、一阶导数和二阶导数预处理反而会降低鉴别模型的性能, 可能采用 SNV、MSC 等算法丢失了部分有用信息。仅基于主要营养成分建立的鉴别模型显示出较低的模型参数以及预测正确率, 而将主要营养成分与近红外光谱结合后作为输入值所建立的模型虽然具有较高的模型参数和预测正确率, 但是与近红外光谱作为单一输入值的模型相比, 其鉴别效果并没有得到很大改善, 说明有机和非有机大豆的成分差异不仅仅体现在本研究所检测的这几个营养成分上, 还需要深入挖掘其他差异性特征标志物, 同时也证明了近红外光谱结合化学计量学方法对有机和非有机大豆进行分类的潜力。

2.2.4 特征波长的筛选

为进一步挖掘有机与非有机大豆近红外光谱与营养

成分之间的联系, 接下来采用 PLS-DA 模型进一步筛选了有机与非有机大豆的特征波长。PLS-DA 模型中的变量投影重要度(variable importance for the projection, VIP)表征了在分类过程中各个变量的重要性, 通常 VIP 值大于 1 的变量对于不同样本组的区分中起着重要贡献^[31]。如图 5A 所示有机与非有机大豆样本在 PLS-DA 得分图中能够明显区分, 这与 PCA 分析的结果是一致的。该模型的 R^2X 为 0.997; R^2Y 为 0.769; Q^2Y 为 0.743, 均大于 0.5, 表明模型稳定性好^[33]。采用交叉验证方差分析检验模型的可靠性, 该模型的 $P_{CV-ANOVA}$ 为 $8.15 \times 10^{-10} < 0.05$, 表明模型在统计上显著有效。从置换检验结果(图 5B)可以发现最右端原始鉴别模型的 R^2 和 Q^2 均大于左边任意一个 Y 变量随机排列模型的值, 且 R^2 的截距为 -0.0342 , Q^2 的截距为 -0.186 , 表明回归模型具有良好的稳定性和重复性。



注: No 为非有机大豆; O 为有机大豆。

图 5 PLS-DA 模型的得分图(A)、置换检验结果(B)、VIP 值图(C)、有机与非有机大豆近红外光谱的特征波长(D)以及特征波长放大图(E 和 F)

Fig.5 Score plot (A), substitution test result (B), VIP value plot of PLS-DA model (C), characteristic wavelengths of organic and non-organic soybean near infrared spectra (D) and amplification plot of characteristic wavelengths (E and F)

基于建立的 PLS-DA 模型筛选对于有机与非有机大豆分类有重要贡献的波长。提取 PLS-DA 模型各个变量的 VIP 值, 共得到 433 个 VIP 值大于 1 的变量, 这些变量对应近红外光谱的波长范围是 5974~5372 cm^{-1} 以及 5064~4000 cm^{-1} (图 5D)。经过分析发现, 得到的两个特征波长范围的近红外光谱吸收峰与脂肪、黄酮类物质以及蛋白质官能团的振动有关, 这从侧面验证了主要营养成分分析的结果, 即有机与非有机大豆在脂肪和大豆苷含量上存在显著差异 ($P < 0.05$)。这也提示未来还需进一步对大豆中的黄酮类化合物和脂肪组成进行全面定性和定量测定, 以准确筛选出有机和非有机大豆的差异化化合物。

2.2.5 基于特征波长的 PLS-DA 模型

基于筛选的特征波长结合不同的预处理建立了有机

与非有机大豆的鉴别模型, 模型参数及验证结果如表 5 所示。在建立的 6 个模型中共有 3 个模型的外部验证正确率达到了 100%。与表 4 中基于近红外光谱全谱建立的模型相比, 基于特征波长建立的模型显示出更高的模型参数。在所有鉴别模型中, 基于近红外光谱特征波长经过 Smoothing-SG 预处理的模型显示出最优的模型参数 R^2X 为 0.996, R^2Y 为 0.735, Q^2 为 0.677, 表明 5974~5372 cm^{-1} 和 5064~4000 cm^{-1} 两个波长范围可以作为鉴别有机与非有机大豆的特征波段。总之, 基于有机与非有机大豆近红外光谱的特征波长结合 S-G 平滑预处理建立的 PLS-DA 模型具有最优的模型参数, 能够有效鉴别有机与非有机大豆。

表 5 基于特征波长建立的 PLS-DA 模型的参数及验证结果
Table 5 Parameters and validation results of the PLS-DA model based on the characteristic wavelength

数据源	光谱预处理方式	R^2X	R^2Y	Q^2	$P_{CV-ANOVA}$	正确率/%
特征波长	-	0.995	0.735	0.677	2.39×10^{-9}	100
	SNV	0.47	0.35	0.259	1.76×10^{-2}	60
	MSC	0.469	0.351	0.259	1.75×10^{-2}	60
	一阶导数	0.265	0.769	0.707	6.40×10^{-8}	100
	二阶导数	0.111	0.986	0.477	0.008	90
	S-G 平滑	0.996	0.735	0.677	2.39×10^{-7}	100

注: -表示无预处理。

3 结 论

本研究分析了有机与非有机大豆的营养成分差异, 有机与非有机大豆在蛋白质、水分以及染料木苷含量上不存在显著性差异, 但有机大豆的脂肪和大豆苷含量显著高于非有机大豆, 这些结果表明有机大豆比非有机大豆具有更高的营养价值。基于营养成分以及近红外光谱建立了有机与非有机大豆的鉴别模型, 大豆近红外光谱涵盖的信息足以鉴别有机与非有机大豆, 而仅基于检测的主要营养成分难以实现两种大豆的鉴别。进一步筛选了对区分有机与非有机大豆有重要贡献的近红外光谱特征波段 (5974~5372 cm^{-1} 和 5064~4000 cm^{-1}), 并基于特征波长结合 S-G 平滑预处理建立了最优的鉴别模型, 能够有效鉴别有机与非有机大豆。总之, 由于种植模式不同, 有机与非有机大豆的营养成分存在差异, 近红外光谱技术在鉴别有机与非有机大豆上具有可行性, 为大豆的质量控制提供了技术基础。在今后的研究中需要进一步探究近红外技术在解决不同品质大豆掺假问题上的应用, 并从增加样品数量和优化判别方法的角度提高鉴别模型的精度和稳定性。

参考文献

- [1] KIM EH, OH SW, LEE SY, *et al.* Comparison of the seed nutritional composition between conventional varieties and transgenic soybean overexpressing *Physaria* FAD3-1 [J]. *J Sci Food Agric*, 2021, 101(6): 2601–2613.
- [2] 张勇, 李雪, 汪雪芳, 等. 基于脂肪酸组成的进口大豆鉴别技术研究[J]. *食品安全质量检测学报*, 2020, 11(8): 2375–2379.
ZHANG Y, LI X, WANG XF, *et al.* Research on identification of imported soybean based on fatty acid composition [J]. *J Food Saf Qual*, 2020, 11(8): 2375–2379.
- [3] 顾强, 乙小娟, 丁萍, 等. 不同原产国大豆质量指标差异的研究[J]. *食品安全质量检测学报*, 2021, 12(6): 2068–2072.
GU Q, YI XJ, DING P, *et al.* Study on the differences of quality items of soybeans imported from different countries [J]. *J Food Saf Qual*, 2021, 12(6): 2068–2072.
- [4] 张玉荣, 寇含笑, 吴琼, 等. 高油大豆储藏过程中的营养品质变化研究[J]. *食品科技*, 2022, 47(2): 207–214.
ZHANG YR, KOU HX, WU Q, *et al.* Changes on nutritional quality of high-oil soybeans during storage [J]. *Food Sci Technol*, 2022, 47(2): 207–214.

- [5] YUN DY, KANG YG, YUN B, *et al.* Distinctive metabolism of flavonoid between cultivated and semiwild soybean unveiled through metabolomics approach [J]. *J Agric Food Chem*, 2016, 64(29): 5773–5783.
- [6] BARRETO NMB, PIMENTA NG, BRAZ BF, *et al.* Organic black beans (*Phaseolus vulgaris* L.) from Rio de Janeiro State, Brazil, present more phenolic compounds and better nutritional profile than nonorganic [J]. *Foods*, 2021, 10(4): 900.
- [7] 郭会平, 陆维盈, 牛宇戈, 等. 运用超高效液相色谱技术区分有机和非有机种植模式的樱桃番茄[J]. *上海交通大学学报(农业科学版)*, 2017, 35(6): 20–25, 30.
- GUO HP, LU WY, NIU YG, *et al.* Differentiating organic and conventional cherry tomatoes by using ultra performance liquid chromatography [J]. *J Shanghai Jiaotong Univ (Agric Sci Ed)*, 2017, 35(6): 20–25, 30.
- [8] KAZIMIERCZAK R, HALLMANN E, RUSACZONEK A, *et al.* Polyphenols, tannins and caffeine content and antioxidant activity of green teas coming from organic and non-organic production [J]. *Renew Agric Food Syst*, 2013, 30(3): 263–269.
- [9] IGLESIAS-CARRERES L, MAS-CAPDEVILA A, BRAVO FI, *et al.* A comparative study on the bioavailability of phenolic compounds from organic and nonorganic red grapes [J]. *Food Chem*, 2019, 299: 125092.
- [10] JIMÉNEZ B, RIVAS A, LORENZO ML, *et al.* Chemosensory characterization of virgin olive oils obtained from organic and conventional practices during fruit ripening [J]. *Flavour Frag J*, 2017, 32(4): 294–304.
- [11] 余梅, 李嘉仪, 范伟, 等. 基于近红外光谱仪与模式识别方法的不同年份陈皮无损鉴别研究[J]. *食品研究与开发*, 2021, 42(19): 171–178.
- YU M, LI JY, FAN W, *et al.* Combination of near-infrared spectroscopy and partial least squares discriminant analysis in detecting the quality of *Panax notoginseng* [J]. *Food Res Dev*, 2021, 42(19): 171–178.
- [12] 李长滨, 牛畅炜, 苏丽, 等. 不同产地山药的近红外鉴别和差异分析[J]. *食品研究与开发*, 2022, 43(15): 175–181.
- LI CB, NIU CW, SU L, *et al.* Identification and variance analysis of Chinese yam from different origins by near infrared spectroscopy [J]. *Food Res Dev*, 2022, 43(15): 175–181.
- [13] 刘珈羽, 李峰庆, 郭换, 等. 白及粉品种近红外快速定性鉴别模型的建立[J]. *成都中医药大学学报*, 2018, 41(1): 34–37.
- LIU JY, LI FQ, GUO H, *et al.* Establishment of rapid NIRS qualitative identification model of *Rhizoma Bletilla* powder varieties [J]. *J Chengdu Tradit Chin Med Univ*, 2018, 41(1): 34–37.
- [14] GONZALEZ-MARTIN MI, HERNANDEZ-HIERRO JM, REVILLA I, *et al.* Differentiation of organic and non-organic ewe's cheeses using main mineral composition or near infrared spectroscopy coupled to chemometric tools: A comparative study [J]. *Talanta*, 2011, 85(4): 1915–1919.
- [15] LIU N, PARRA HA, PUSTJENS A, *et al.* Evaluation of portable near-infrared spectroscopy for organic milk authentication [J]. *Talanta*, 2018, 184: 128–135.
- [16] 王铎, 张力, 范素杰, 等. 大豆主要脂肪酸含量近红外模型的建立[J]. *大豆科学*, 2017, 36(2): 295–299.
- WANG D, ZHANG L, FAN SJ, *et al.* Establishment of NIRS model for the main fatty acids in soybean [J]. *Soybean Sci*, 2017, 36(2): 295–299.
- [17] 李琳琳, 金华丽, 崔彬彬, 等. 基于近红外透射光谱的大豆蛋白质和粗脂肪含量快速检测[J]. *粮食与油脂*, 2014, 27(12): 57–60.
- LI LL, JIN HL, CUI BB, *et al.* Rapid determination of soybean protein and crude fat content by near-infrared transmittance spectroscopy [J]. *Cere Oil*, 2014, 27(12): 57–60.
- [18] 温冰消, 刘卫国, 李虹桥, 等. 基于近红外法的鲜食大豆品质快速分析技术[J]. *分子植物育种*, 2018, 16(12): 4062–4067.
- WEN BX, LIU WG, LI HQ, *et al.* Rapid analysis of edamame quality by near infrared spectroscopy [J]. *Mol Plant Breed*, 2018, 16(12): 4062–4067.
- [19] 王海龙, 杨向东, 张初, 等. 近红外高光谱成像技术用于转基因大豆快速无损鉴别研究[J]. *光谱学与光谱分析*, 2016, 36(6): 1843–1847.
- WANG HL, YANG XD, ZHANG C, *et al.* Fast identification of transgenic soybean varieties based near infrared hyperspectral imaging technology [J]. *Spectrosc Spectral Anal*, 2016, 36(6): 1843–1847.
- [20] 李凯楠, 沈广辉, 叶文武, 等. 基于近红外高光谱成像技术识别紫斑和霉变大豆的方法[J]. *南京农业大学学报*, 2022, 45(4): 691–698.
- LI KN, SHEN GH, YE WW, *et al.* Identification method of soybean seeds with purple spot and moldy based on near infrared hyperspectral imaging technology [J]. *J Nanjing Agric Univ*, 2022, 45(4): 691–698.
- [21] 王富豪, 黄璐, 薛晨晨, 等. 不同品种大豆异黄酮组分及体外抗氧化活性分析[J]. *食品工业科技*, 2021, 42(17): 247–255.
- WANG FH, HUANG L, XUE CC, *et al.* Analysis on the isoflavone profile and antioxidant activity in vitro of different soybean varieties [J]. *Sci Technol Food Ind*, 2021, 42(17): 247–255.
- [22] 晁芳芳, 李森, 王悦. HPLC法测定大豆提取物中大豆异黄酮含量的研究[J]. *粮食加工*, 2012, 37(5): 79–81.
- CHAO FF, LI S, WANG Y. Research of determination of isoflavone in extracting of soybean by HPLC [J]. *Grain Proc*, 2012, 37(5): 79–81.
- [23] QI SW, XU DP, CHEN QW, *et al.* Indirect competitive chemiluminescence immunoassay for glycosides daidzin and genistin in soy products [J]. *Food Anal Method*, 2017, 11(2): 367–373.
- [24] 费洪立, 阮长青, 李志江, 等. 大豆异黄酮作为大豆溯源特征指标的筛选[J]. *中国油脂*, 2022, 47(2): 148–152.
- FEI HL, RUAN CQ, LI ZJ, *et al.* Screening of soybean isoflavone as characteristic index of soybean traceability [J]. *Chin Oil Fat*, 2022, 47(2): 148–152.
- [25] 刘文静. 基于大豆异黄酮特征的大豆产地溯源研究[D]. 大庆: 黑龙江八一农垦大学, 2018.
- LIU WJ. Origin traceability of soybean based on soybean isoflavone features [D]. Daqing: Heilongjiang Bayi Agricultural University, 2018.
- [26] AMANAH HZ, TUNNY SS, MASITHOH RE, *et al.* Nondestructive prediction of isoflavones and oligosaccharides in intact soybean seed using fourier transform near-infrared (FT-NIR) and fourier transform infrared

- (FT-IR) spectroscopic techniques [J]. *Foods*, 2022, 11(2): 232.
- [27] FERREIRA DS, GALÃO OF, PALLONE JAL, *et al.* Comparison and application of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for determination of quality parameters in soybean samples [J]. *Food Control*, 2014, 35(1): 227–232.
- [28] 李颖, 马雨辰, 刘萌, 等. 近红外光谱技术结合偏最小二乘判别分析检测三七品质[J]. *食品安全质量检测学报*, 2022, 13(12): 3923–3929.
- LI Y, MA YC, LIU M, *et al.* Combination of near- infrared spectroscopy and partial least squares discriminant analysis in detecting the quality of *Panax notoginseng* [J]. *J Food Saf Qual*, 2022, 13(12): 3923–3929.
- [29] WANG P, ZHANG H, YANG HL, *et al.* Rapid determination of major bioactive isoflavonoid compounds during the extraction process of kudzu (*Pueraria lobata*) by near-infrared transmission spectroscopy [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2015, 137: 1403–1408.
- [30] 王文秀, 彭彦昆. 基于近红外光谱的冷鲜肉-解冻肉的判别研究[J]. *食品安全质量检测学报*, 2014, 5(3): 754–760.
- WANG WX, PENG YK. Discrimination of fresh and frozen-thawed meat based on near infrared spectroscopy [J]. *J Food Saf Qual*, 2014, 5(3): 754–760.
- [31] 黄婷, 杨亚娇, 刘茗铭, 等. 基于化学计量学的四川不同产区浓香型原酒判别分析[J]. *酿酒*, 2022, 49(1): 64–70.
- HUANG T, YANG YJ, LIU MM, *et al.* Discriminant analysis of Luzhou-flavor original liquor from different producing areas in Sichuan based on chemometrics methods [J]. *Liquor Making*, 2022, 49(1): 64–70.
- [32] TRIBA MN, MOYEC LL, AMATHIEU R, *et al.* PLS/OPLS models in metabolomics: The impact of permutation of dataset rows on the K-fold cross-validation quality parameters [J]. *Mol Biosyst*, 2015, 11(1): 13–19.
- [33] 汪海斌, 张楠楠, 洪稳稳, 等. 密蒙花中 4 种成分的含量测定及其主成分、偏最小二乘判别分析[J]. *安徽中医药大学学报*, 2020, 39(4): 86–89.
- WANG HB, ZHANG NN, HONG WW, *et al.* Content determination of four constituents in *Buddleja officinalis* and its principal component analysis and partial least squares-discriminant analysis [J]. *J Anhui Univ Chin Med*, 2020, 39(4): 86–89.

(责任编辑: 于梦娇 韩晓红)

作者简介



余新金, 硕士研究生, 主要研究方向为食品加工与安全。
E-mail: 632445160@qq.com



陈奕, 教授, 博士生导师, 主要研究方向为食品质量与安全。
E-mail: chenyi-417@163.com