

实验技术与方法

近红外光谱法快速鉴别转基因大豆模型优化

高慧宇,王竹,张雪松,王国栋

(中国疾病预防控制中心营养与健康所 国家卫生健康委员会微量元素
与营养重点实验室,北京 100050)

摘要:目的 应用近红外光谱(NIR)结合偏最小二乘判别分析(PLS-DA)建立转基因大豆的快速鉴别模型,并选择最优模型。方法 主成分分析(PCA)用于从光谱数据中提取相关特征并剔除异常样品。在试验中,94份样品用于构建模型,41份样品用作验证评估模型的效果。分别讨论样品形态(整粒和粉末)、波长范围和光谱预处理方法对所建模型判别正确率的影响。结果 粉末状大豆样品建模的效果好于整粒大豆样品。其中判定效果最好的模型,整粒大豆在 $9\ 403\sim 5\ 438\ \text{cm}^{-1}$ 范围内,采用二阶导数(2nd)处理光谱,模型的校正集和验证集的判定正确率均为100.00%;粉末状大豆在 $7\ 505\sim 4\ 597\ \text{cm}^{-1}$ 范围内,采用矢量归一化+一阶导数(SNV+1st)处理光谱,模型的校正集和验证集的判定正确率也均为100.00%。结论 通过选择样品形态、波长范围和光谱预处理方法可以优化鉴别模型,提高近红外判别模型的鉴别正确率。

关键词:近红外光谱;转基因大豆;偏最小二乘-判别分析;鉴别

中图分类号:R155 文献标识码:A 文章编号:1004-8456(2020)03-0244-06

DOI:10.13590/j.cjfh.2020.03.005

Model optimization for fast discrimination of transgenic soybeans using near-infrared spectroscopy

GAO Huiyu, WANG Zhu, ZHANG Xuesong, WANG Guodong

(Key Laboratory of Trace Element Nutrition of National Health Commission, National Institute for Nutrition and Health, Chinese Center for Disease Control and Prevention, Beijing 100050, China)

Abstract: Objective Near-infrared (NIR) spectroscopy and partial least squares-discriminant analysis (PLS-DA) were applied to discriminate soybean samples as being transgenic or non-transgenic. The rapid discrimination models for transgenic soybean were established, and the optimal model was selected. **Methods** Principal component analysis (PCA) was used to extract relevant features from the spectral data and remove anomalous samples. In experimental studies, 94 samples were used to build models and 41 samples were used as the validation to evaluate the performance of the developed models. The effects of sample morphology (intact or ground), wavelength range and spectral pretreatment method on the correctness of the model were discussed. **Results** Models for intact soybean samples obtain better judgment performance than models for ground samples. The best discriminant model for intact soybean samples possessed both 100.00% discriminant correct rate in calibration and validation sets at $9\ 403\sim 5\ 438\ \text{cm}^{-1}$ using second derivative (2nd). The best discriminant model for ground soybean samples also achieved both 100.00% discriminant correct rate in calibration and validation sets at $7\ 505\sim 4\ 597\ \text{cm}^{-1}$ using standard normal variate plus first derivative (SNV+1st). **Conclusion** By selecting sample morphology, wavelength range and spectral pretreatment method, the discrimination model can be optimized and the discriminant correct rate can be significantly improved.

Key words: Near-infrared spectroscopy; transgenic soybean; partial least squares-discriminant analysis; discrimination

大豆在许多国家和地区的传统饮食中发挥着

重要的作用,其脂肪含量低,是蛋白质、膳食纤维以及各种微量营养素和植物化学物的极好来源^[1]。研究^[2-3]显示,人们定期食用大豆和大豆制品,可降低患癌症的风险,改善心血管疾病或减少其他慢性疾病,因此,人们希望通过遗传技术^[4]提高大豆的产量或改变大豆的特性,例如提高种子的营养成分^[5-7](包括蛋白质、油和一些活性成分)、提高种子

收稿日期:2020-04-09

基金项目:国家转基因生物新品种培育重大专项(2016ZX08011005)

作者简介:高慧宇 女 助理研究员 研究方向为营养与食品卫生学 E-mail: gaohuiyu2520@163.com

通信作者:王竹 女 研究员 研究方向为营养与食品卫生学 E-mail: wzhu@163.com

的抗病和抗虫性^[8]或是在新的地理区域种植发展新品种等。

虽然转基因生物具有明显的优势,如对昆虫、杂草、疾病的抗性,以及更好的营养价值和更高的产量,但转基因生物的潜在风险,如对人类和动物健康的影响及其对环境的影响,使得政府组织和公众仍然犹豫或者反对转基因生物的培养和应用^[9-11],因此,需要实施质量控制,开发一些可靠有效的针对转基因产品的检测方法。

关于转基因产品的检测方法主要有基于蛋白质、DNA 或者光谱学和色谱学的几类方法^[10]。免疫印迹(western blot)、酶联免疫吸附测定(ELISA)^[12]和测流条(lateral flow strips)^[13]是基于蛋白质的检测方法。聚合酶链式反应(PCR)^[14-15]是最常用的基于 DNA 的检测方法,广泛应用于扩增靶 DNA。高效液相色谱(HPLC)和气相色谱(GC)^[10]也可以作为转基因产品检测的替代方法。总的来说,基于 DNA 的转基因鉴定方法具有足够的可信度,与其他方法比较具有更高的可靠性^[10,15]。但是,DNA 方法需要一些非常复杂和耗时的过程才能获得结果,因此很多实际情况(比如样品量大、需要快速检测、缺乏专业技术人员等)不适用。理想的识别技术应该是快速准确易于使用,并且成本比较低。

作为上述几种方法的替代方法,近红外光谱(NIR)技术结合化学计量学凭借其优于其他分析技术的优势已在不同的领域获得广泛应用。近红外光谱分析具有非破坏性且简单、快速、无需或只需简单的样品预处理,使得该技术非常适用于过程监测和质量控制^[16-17]。有机物分子具有多个重叠的近红外吸收带,这是由于 C-H、N-H、O-H 和 S-H 官能团的组合振动引起的^[18]。近红外光谱技术检测转基因产品的基础是可以识别由基因型变化引起的表型变化,最终导致有机分子键的变化^[10]。

近年来,相关研究^[11,19-20]利用近红外光谱技术结合化学计量学显示出对转基因大豆或其制品鉴别的良好效果。然而,这些研究都没有涉及到关于样品不同形态(整粒和粉末)、建模过程中光谱预处理方法和波长选择等具体因素对于鉴别模型判定正确率的影响。本研究将进一步比较不同因素对转基因大豆鉴别模型判定正确率的影响,以期找到最佳鉴别模型用于转基因大豆样品的鉴别。

1 材料与方法

1.1 材料

1.1.1 样品收集

收集 139 份大豆样品,包括 69 份转基因样品和

70 份非转基因样品。非转基因大豆样品来自东北三省大豆主产区,品种包括绥农、吉育和黑河等;转基因大豆为 ZH10-6,通过向大豆黄品种中转入 G2-EPSPS 和 GAT 双价基因获得,分别来自北京市、石家庄市和三亚市各 23 份。转基因和非转基因大豆样品仅在目标性状(抗除草剂特性)方面有差异,其他外观上均无明显差异。制备两种形态的大豆样品(整粒和粉末),选取相似粒径(约 5 mm)的大豆作整粒大豆样品,用水清洗,干燥后一份在室温(25 ℃,相对湿度 60%)条件下保存,另一份研磨成粉并过筛 0.45 mm,制成粉末状大豆样品,同样室温(25 ℃,相对湿度 60%)保存备用。

1.1.2 主要仪器

MPA 型傅立叶近红外光谱分析仪(德国 Bruker,扫描范围为 12 500~3 300 cm^{-1} ,分辨率 2 cm^{-1})。

1.2 方法

1.2.1 光谱采集

光谱仪开机 0.5 h 后进行仪器性能测试,性能测试通过后,开始采集样品光谱。取不少于样品杯 1/3 体积的大豆样品,用积分球漫反射模式扫描得到整粒或粉末状大豆样品的近红外吸收光谱图。扫描谱区 12 800~3 600 cm^{-1} ,分辨率 16 cm^{-1} ,每份样品重复装样扫描 2 次。OPUS 7.5 软件用于控制记录仪,收集光谱和导入数据。

1.2.2 判别模型的建立

1.2.2.1 判别分析方法

本研究利用 OPUS 7.5 软件的 Quant 2 方法,采用偏最小二乘-判别分析(PLS-DA)进行判别分析。PLS-DA 是 y 二进制矩阵的 PLS 回归,可将类别与 x 矩阵区分开。该技术用于改善两组或多组样品之间的分离^[9]。当构建分类模型区分基于同一校准集的不同样品时,校准集中的每份样品会被分配一个虚拟变量作为参考值^[21](本试验中,设定非转基因大豆样品=1,转基因大豆样品=2)。未知样品的分类取决于 PLS 模型预测的 y 值。根据文献^[11-22],所有模型的识别阈值误差设置为 ± 0.5 ,因此,如果 y 值在 0.5~1.5 之间,非转基因样品被正确分类,否则样品被错误分类;同理, y 值在 1.5~2.5 之间时,转基因样品被正确分类。

1.2.2.2 光谱预处理方法和主成分分析(PCA)

本研究使用不同的光谱预处理方法,其目的是消除由于样品颗粒大小、均匀性等^[20]对光程的多重干扰或减少光谱峰重叠和消除基线漂移。主要有矢量归一化(SNV)、多元散射校正(MSC)、一阶导数(1st)、二阶导数(2nd)以及几种叠加的光谱预处理方法(如 SNV+1st, MSC+1st),以比较不同预处理

方法的模型鉴别效果。

作为目前近红外光谱分析中常用的化学计量法,PCA可以尽可能多的记录原始变量的方差,同时降低数据的维数并最大限度减少测量误差^[9]。具体来说,可以通过构造被称为主成分(PCs)的新的不相关变量从原始数据集中提取信息^[21]。每个频谱都有自己独特的分数,因此,光谱可以用因子空间中的PCA分数代替波长空间中的强度来表示^[22]。此后,这些主成分可用于研究样品之间的分组现象,特别应用于分类问题中。本研究利用PCA剔除样品光谱异常值,并随机选取正常光谱的70%作为校正集光谱,30%作为验证集光谱来构建判别模型。

1.2.3 模型评价

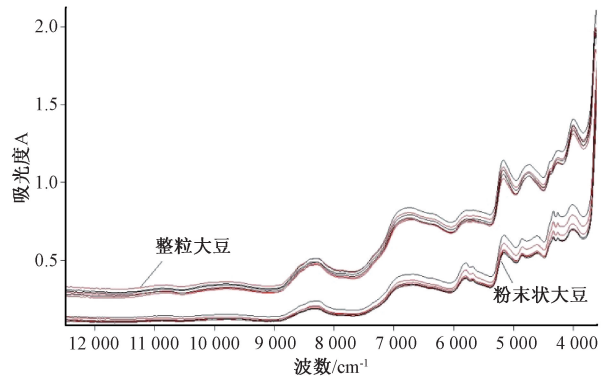
用校准和验证集的均方根误差(RMSE)和相关系数(*r*)评价PLS-DA模型。对于所有分析的参数,产生最小RMSE值的预处理方法被认为是最佳的。用于校准的RMSE为校准均方根误差(RMSEC),用于验证或预测的RMSE为验证均方根误差(RMSEV),用于校准的*r*为*r_c*,用于验证的*r*为*r_v*。

2 结果与分析

2.1 转基因和非转基因大豆的近红外光谱

整粒和粉末状的转基因和非转基因大豆样品的部分原始光谱图如图1所示。粉末状大豆样品的近红外光谱图形状与整粒大豆样品非常相似,只是整个光谱图相比整粒大豆的图谱下移,这可能与样品的粒径有关,导致近红外光谱中纵坐标值有差异。非转基因和转基因大豆的近红外光谱没有明显差异,各条谱线几乎重叠在一起,因此,不可能通过观察转基因和非转基因大豆样品的光谱差异来辨别样品,需要借助化学计量学工具结合样品的近红外光谱来判定。总体来说,这些光谱的特征峰有5000、6500和8500 cm⁻¹,主要是由于C-H以及三种主要化合物(蛋白质、脂肪和碳水化合物)的C-H

和O-H的振动引起的。而5000~4000 cm⁻¹之间的振动信号则是由于O-H(水)和C-C的伸缩振动引起的,高于4300 cm⁻¹的峰则是与C-H相关的键的组合振动引起的^[23]。



注:黑线表示非转基因大豆;红线代表转基因大豆

图1 部分整粒和粉末状大豆的转基因和非转基因样品的近红外光谱图

Figure 1 Part of NIR spectra of non-transgenic and transgenic intact and ground soybeans

光谱范围的选取,整粒和粉末状大豆样品均选取了全谱12800~3960和9000~4000 cm⁻¹两段。其中选取全谱探索整个近红外光谱范围内建模的效果,而9000~4000 cm⁻¹段是包含了所有波峰信息的一段。通过“优化”的方法筛选出对于整粒和粉末状大豆建模效果最好的一段光谱,即整粒大豆在9403~5438 cm⁻¹范围内,粉末状大豆在7505~4597 cm⁻¹范围内。选取最常见的7种光谱预处理方法(包括无预处理),分别在三种波段下,建立模型,比较各模型的数据。

2.2 PLS-DA 建模样品分组情况

在12800~3600 cm⁻¹光谱范围内对整粒和粉末状大豆均进行PCA以检查所选样品的光谱可变性。如图2所示,整粒大豆样品中第一主成分(PC1)占其总变异的53.45%,PC2占45.03%;4份样品(样品号58、61、63和84)呈现高分值,表明这些样品很可能是异常的。对粉末状样品进行PCA,

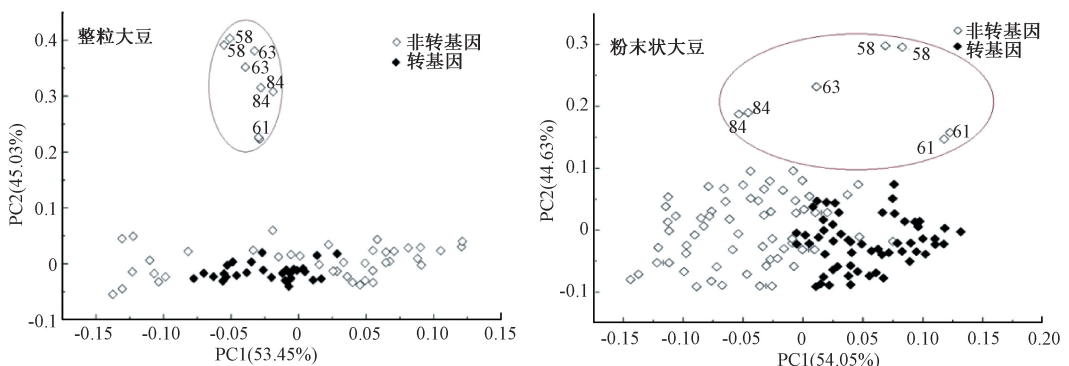


图2 PC1和PC2在整粒和粉末状大豆的分布情况

Figure 2 Scores on PC2 versus scores on PC1 for intact and ground soybean samples

也得到相似的结果,PC1 占其总变异的 54.05%,PC2 占 44.63%,4 份样品均为异常。样品分组情况见表 1,异常样品需要在建模前删除,以免对鉴别产生不好的影响^[11,24]。

表 1 整粒和粉末状大豆校正集和验证集样品分组情况

样品类型	基因型	样品份数	异常值	最终样品份数	校准集	验证集
整粒	转基因	69	0	69	56	13
	非转基因	70	4	66	38	28
粉末状	转基因	69	0	69	54	15
	非转基因	70	4	66	40	26

表 2 整粒大豆样品 PLS-DA 建模各指标的统计量

Table 2 Statistics for intact soybean samples using PLS-DA models

预处理方法	波长/(cm^{-1})	校正集($n=94$)			验证集($n=41$)		
		$RMSEC$	r_c	正确率/%	$RMSEV$	r_v	正确率/%
无	12 800~3 960	0.180	0.930 6	98.40	0.233	0.866 0	95.12
	9 000~4 000	0.216	0.899 1	96.81	0.233	0.870 6	93.90
	9 403~5 438	0.217	0.897 8	96.28	0.216	0.889 9	100.00
SNV	12 800~3 960	0.178	0.932 1	97.34	0.216	0.890 0	96.34
	9 000~4 000	0.213	0.901 8	95.21	0.243	0.858 3	95.12
	9 403~5 438	0.232	0.882 7	95.74	0.200	0.906 8	100.00
MSC	12 800~3 960	0.200	0.913 5	97.87	0.207	0.900 1	97.56
	9 000~4 000	0.208	0.906 5	95.74	0.243	0.859 0	96.34
	9 403~5 438	0.232	0.882 6	95.74	0.198	0.908 7	100.00
1st	12 800~3 960	0.138	0.960 1	99.47	0.156	0.944 1	100.00
	9 000~4 000	0.182	0.929 3	97.34	0.147	0.950 5	100.00
	9 403~5 438	0.161	0.944 8	98.40	0.150	0.948 5	100.00
2nd	12 800~3 960	0.166	0.941 7	98.94	0.126	0.963 9	100.00
	9 000~4 000	0.174	0.935 5	98.94	0.143	0.953 8	100.00
	9 403~5 438	0.120	0.969 6	100.00	0.147	0.950 9	100.00
SNV+1st	12 800~3 960	0.175	0.934 4	98.94	0.136	0.958 2	100.00
	9 000~4 000	0.214	0.900 6	98.94	0.144	0.952 9	100.00
	9 403~5 438	0.167	0.941 0	98.94	0.121	0.967 0	100.00
MSC+1st	12 800~3 960	0.154	0.949 6	98.94	0.136	0.958 0	100.00
	9 000~4 000	0.184	0.927 6	98.94	0.129	0.962 3	100.00
	9 403~5 438	0.158	0.947 2	98.40	0.124	0.965 4	100.00

通常以 $RMSE$ 越小、 r 值越大判定为较好的模型。对于整粒大豆样品,光谱预处理为 2nd,光谱范围为 $9\ 403\sim 5\ 438\ \text{cm}^{-1}$,得到的鉴别模型校正集的 $RMSEC$ 最小(0.120),且 r_c 值较大(0.969 6),校正集和验证集判定正确率均为 100.00%,故此模型的效果最好。对于粉末状大豆样品,光谱预处理为 SNV+1st,光谱范围为 $7\ 505\sim 4\ 597\ \text{cm}^{-1}$,得到的鉴别模型校正集 $RMSEC$ 最小(0.075 3), r_c 值较大(0.988 3),校正集和验证集的判定正确率均为 100.00%,故此模型的效果最好。图 3 为判定效果最好的两个模型中参考值和预测值的关系图。

2.3 PLS-DA 模型建立和参数优化

选取不同光谱范围、利用不同光谱预处理方法得到整粒和粉末状大豆样品 PLS-DA 模型的鉴别结果,见表 2 和 3。整粒和粉末状大豆分别建立鉴别模型,判定正确率均在 95.00% 以上,效果较好。通过选取适当的光谱预处理方法和光谱范围可以提高模型的预测效果,粉末状大豆建模的结果整体强于整粒大豆样品建模的效果,可能是粉末状大豆的近红外光谱包含了比整粒大豆更多的信息,从而减少了模型预测与真实情况的差异。

3 讨论

结合近年来涉及到的对转基因大豆及其制品鉴别的相关研究^[11,19-20],LUNA 等^[11]对 80 份大豆油进行转基因和非转基因的鉴别,通过对三种建模法的比较,得出 PLS-DA 建模是最优的建模方法,因此本试验直接采用 PLS-DA 建模,在判定正确率方面有较大提高。吴江等^[19]是利用 BP 神经网络技术构建模型,也应用了 PCA 对 90 份样品建模、22 份样品预测,判定正确率为 100%。本试验中建模样品为 94 份,预测样品为 41 份,除了建模方法不同外,本试验还详细对比了各个因素对于建模效果的影响(包括样品形态、光谱选择范围和光谱预处理方

表3 粉末状大豆样品 PLS-DA 建模各指标的统计量

Table 3 Statistics for ground soybean samples using PLS-DA models

预处理方法	波长/(cm^{-1})	校正集($n=94$)			验证集($n=41$)		
		RMSEC	r_c	正确率/%	RMSEV	r_v	正确率/%
无	12 800~3 960	0.141	0.958 2	99.47	0.180	0.928 0	96.34
	9 000~4 000	0.123	0.968 7	99.47	0.129	0.963 2	97.73
	7 505~4 597	0.109	0.975 3	99.47	0.111	0.973 1	97.73
SNV	12 800~3 960	0.123	0.968 6	100.00	0.129	0.963 4	98.86
	9 000~4 000	0.113	0.973 6	100.00	0.127	0.964 4	100.00
	7 505~4 597	0.085	0.985 1	100.00	0.091	0.982 0	100.00
MSC	12 800~3 960	0.117	0.971 4	100.00	0.137	0.958 8	97.73
	9 000~4 000	0.109	0.975 5	100.00	0.126	0.965 5	100.00
	7 505~4 597	0.089 2	0.983 6	100.00	0.089 9	0.982 4	100.00
1st	12 800~3 960	0.098 2	0.980 1	100.00	0.102	0.977 2	100.00
	9 000~4 000	0.108	0.976 0	100.00	0.096 1	0.979 9	100.00
	7 505~4 597	0.083 8	0.985 5	100.00	0.093 8	0.980 9	100.00
2nd	12 800~3 960	0.094 3	0.981 6	100.00	0.099 4	0.978 5	100.00
	9 000~4 000	0.103	0.978 0	100.00	0.110	0.973 8	100.00
	7 505~4 597	0.095 7	0.981 1	100.00	0.101	0.977 6	100.00
SNV+1st	12 800~3 960	0.095 3	0.981 3	100.00	0.103	0.977 0	100.00
	9 000~4 000	0.093 7	0.981 9	100.00	0.088 1	0.983 1	100.00
	7 505~4 597	0.075 3	0.988 3	100.00	0.072 6	0.988 6	100.00
MSC+1st	12 800~3 960	0.099 3	0.979 6	100.00	0.094 7	0.980 5	100.00
	9 000~4 000	0.093 0	0.982 1	100.00	0.104	0.976 5	100.00
	7 505~4 597	0.076 9	0.987 8	100.00	0.089 8	0.982 5	100.00

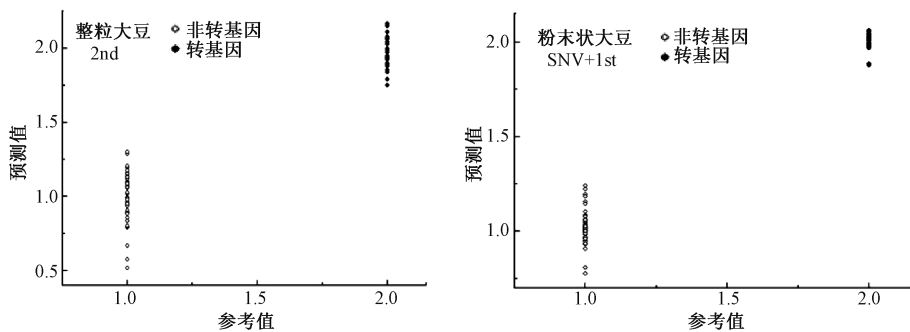


图3 整粒和粉末状大豆样品效果最好的鉴定模型中参考值和预测值关系图

Figure 3 Predicted vs reference value plot of the best PLS-DA model for intact and ground soybean samples

法等,其中大部分模型的判定正确率都能达到95.00%以上。王海龙等^[20]研究采用的是高光谱成像技术,此技术与近红外光谱技术比较,弥补了检测对象的空间信息,采用PLS-DA建模,通过提取特征光谱探讨了非转基因亲本大豆品种鉴别、转基因大豆及其亲本大豆的品种鉴别等,但是没有讨论光谱预处理方法对于模型鉴别效果的影响。

本试验基于近红外光谱技术,结合PCA,对转基因与非转基因大豆进行快速鉴别研究,分别探讨

了样品形态(整粒和粉末状)、近红外光谱波长选择以及光谱预处理等对鉴别模型判定正确率的影响。结果表明粉末状大豆样品所建模型的判定正确率普遍高于整粒大豆样品的模型。同时,通过提取合适光谱范围以及选择合适光谱预处理方法可以明显提高判别模型的正确率。

参考文献

[1] 李峰,徐小洲,张文伟,等.大豆营养价值及品质分析[J].

- 中国食品工业, 2016(4):56-57.
- [2] MESSINA M. Soy and health update: evaluation of the clinical and epidemiologic literature[J]. *Nutrients*, 2016,8(12):754.
- [3] CHANG A S, YEONG B Y, KOH W P. Symposium on plant polyphenols: nutrition, health and innovations, June 2009[J]. *Nutrition Reviews*, 2010,68(4):246-252.
- [4] BROOKES G, BARFOOT P. Global income and production impacts of using GM crop technology 1996-2014[J]. *GM Crops & Food*, 2016,7(1):38-77.
- [5] NATARAJAN S, LUTHRIA D, BAE H, et al. Transgenic soybeans and soybean protein analysis: an overview[J]. *Journal of Agricultural and Food Chemistry*, 2013, 61 (48): 11736-11743.
- [6] 王凤敏, 赵青松, 史晓蕾, 等. 转基因大豆食品的发展现状[J]. *大豆科技*, 2019(4):49-52.
- [7] 王邵宇. 转基因大豆的发展及其风险探究[J]. *粮食科技与经济*, 2018,43(7):117-119.
- [8] BALES C, ZHANG G R, LIU M H, et al. Mapping soybean aphid resistance genes in PI 567598B [J]. *Theoretical and Applied Genetics*, 2013,126(8):2081-2091.
- [9] LUNA A S, DA SILVA A P, PINHO J S A, et al. A novel approach to discriminate transgenic from non-transgenic soybean oil using FT-MIR and chemometrics [J]. *Food Research International*, 2015,67(11):206-211.
- [10] ALISHAHI A, FARAHMAND H, PRIETO N, et al. Identification of transgenic foods using NIR spectroscopy: a review [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2010,75(1):1-7.
- [11] LUNA A S, DA SILVA A P, PINHO J S A, et al. Rapid characterization of transgenic and non-transgenic soybean oils by chemometric methods using NIR spectroscopy [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2013,100(2):115-119.
- [12] ALBRIGHT V C, HELLMICH R L, COATS J R. A review of cry protein detection with enzyme-linked immunosorbent assays[J]. *Journal of Agricultural and Food Chemistry*, 2016, 64 (11): 2175-2189.
- [13] HUANG X, ZHAI C C, YOU Q M, et al. Potential of cross-priming amplification and DNA-based lateral-flow strip biosensor for rapid on-site GMO screening[J]. *Analytical and Bioanalytical Chemistry*, 2014,406(17):4243-4249.
- [14] BOGANI P, MINUNNI M, SPIRITI M M, et al. Transgenes monitoring in an industrial soybean processing chain by DNA-based conventional approaches and biosensors [J]. *Food Chemistry*, 2009,113(2):658-664.
- [15] DEISINGH A K, BADRIE N. Detection approaches for genetically modified organisms in foods [J]. *Food Research International*, 2005,38(6):639-649.
- [16] 卞中悦, 贾若玉, 郭倩倩, 等. 基于近红外光谱的掺伪乳制品鉴别模型优化[J]. *自然科学*, 2019,7(3):96-105.
- [17] 欧家杰, 姜仕程, 张成, 等. 基于可见-近红外光谱技术的红茶等级判别研究[J]. *应用物理*, 2019,9(5):233-242.
- [18] HACISALIHOGU G, LARBI B, SETTLES A M. Near-infrared reflectance spectroscopy predicts protein, starch, and seed weight in intact seeds of common bean (*Phaseolus vulgaris* L.) [J]. *Journal of Agricultural and Food Chemistry*, 2010, 58 (2): 702-706.
- [19] 吴江, 黄富荣, 黄才欢, 等. 近红外光谱结合主成分分析和BP神经网络的转基因大豆无损鉴别研究[J]. *光谱学与光谱分析*, 2013,33(6):1537-1541.
- [20] 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.
- [21] ZHANG H G, YANG Q M, LU J G. Classification of washing powder brands using near-infrared spectroscopy combined with chemometric calibrations [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2014, 120 (11): 625-629.
- [22] XIE L J, YING Y B, YING T J, et al. Discrimination of transgenic tomatoes based on visible/near-infrared spectra [J]. *Analytica Chimica Acta*, 2007,584(2):379-384.
- [23] PLANS M, SIMÓ J, CASAÑAS F, et al. Near-infrared spectroscopy analysis of seed coats of common beans (*Phaseolus vulgaris* L.): a potential tool for breeding and quality evaluation [J]. *Journal of Agricultural and Food Chemistry*, 2012,60(3):706-712.
- [24] CUI H F, YE Z H, XU L, et al. Automatic and rapid discrimination of cotton genotypes by near infrared spectroscopy and chemometrics [J]. *Journal of Analytical Methods in Chemistry*, 2012(1):793468.