

大豆粗蛋白、粗脂肪含量近红外检测模型建立及可靠性分析

闫 龙¹, 蒋春志¹, 于向鸿², 杨春燕¹, 张孟臣¹

(¹河北省农林科学院粮油作物研究所, 国家大豆改良中心石家庄分中心, 河北省遗传育种重点实验室, 河北 石家庄 050031; ²中国农业科学院研究生院, 北京 10081)

摘 要:大豆是重要的植物蛋白和食用油来源, 目前常用的蛋白质和脂肪化学分析方法不仅操作步骤繁琐, 而且籽粒粉碎而不能延续后代, 在大豆品质改良的后代选育过程中的应用受到限制。为明确快速、简便、非破坏籽粒的近红外检测方法利用的可行性。利用 4 个建模样品集, 采用偏最小二乘法建立了 3 个近红外大豆粗蛋白、粗脂肪籽粒检测模型和 1 个粉末检测模型。通过重新采集 8 个不同蛋白质和油份含量的大豆样品检测分析, 并送检权威检测部门进行化学分析比较, 分析近红外检测的可靠性。通过稳定性、一致性等分析表明, 合适的建模样品集是正确建模的前提, 在所建 3 个籽粒模型中, 以含 415 个大豆材料的样品集所建模型 (M6) 可靠性最好; 分析结果还表明, 近红外检测结果与化学分析结果一致, 在需要保存籽粒完好的大豆杂交分离世代或大量样品检测时, 用近红外检测代替化学分析是可行的。

关键词:大豆; 近红外; 蛋白质; 脂肪

中图分类号: S565.1

文献标识码: A

文章编号: 1000-9841(2008)05-0833-05

Development and Reliability of Near Infrared Spectroscopy (NIS) Models of Protein and Oil Content in Soybean

YAN Long¹, JIANG Chun-zhi¹, YU Xiang-hong², YANG Chun-yan¹, ZHANG Meng-chen¹

(¹Institute of Cereal and Oil Crops, Hebei Academy of Agricultural and Forestry Sciences, Shijiazhuang Branch Center of National Center for Soybean Improvement, Key Laboratory of Crop Genetics and Breeding, Shijiazhuang 050031, Hebei; ²Graduate School of Chinese Academy of Agricultural Sciences, Beijing 10081, China)

Abstract: Soybean provides a major source of protein and oil. The present analysis method of protein and oil is time-consuming and destructive, and its application in high oil and high protein breeding is restricted, hence, rapid and nondestructive analysis method is required for quality improvement. This research was done in order to validate the reliability of the near infrared spectroscopy (NIS) models which was rapid, simple and nondestructive. Near infrared spectroscopy of 4 sample pools were employed to develop regression models depending on least squares. They were 3 seed model and 1 powder model. Depending on re-sampling, comparison was done among 7 results including 3 chemistry results from different elite testing-center and 4 NIS models' results. The result showed that the suitable sample pool was necessary for excellent models. M6 which depended on a sample including 415 samples was the best one among 3 NIS seed model. Correlation among different methods were obviously consistent. Hence, using NIS models to substitute chemistry method was reliable especially for large scale nondestructive testing.

Key words: Soybean; NIS; Protein; Oil

大豆籽粒约含粗蛋白 40%, 粗脂肪 20%, 是重要的植物蛋白和食用油来源^[1-2]。提高大豆籽粒中蛋白和脂肪含量是品质改良的重要目标之一, 准确测定其粗蛋白、粗脂肪含量是前提。目前, 国家颁布

的粗蛋白标准测定方法是基于凯氏定氮法的化学分析 (GB2905-1982), 粗脂肪标准方法是基于索氏提取法的化学分析 (GB2906-1982)。尽管通常认为其可靠性较高, 但操作步骤繁琐、籽粒粉碎而不能延续

收稿日期: 2008-04-02

基金项目: 国家重点基础研究发展计划 (973 计划) 资助项目 (2004CB117203); 国家高技术研究发展计划 (863 计划) 资助项目 (2006AA10Z1B3、2006AA100104); 国家支撑计划资助项目 (2006BAD13B05)。

作者简介: 闫龙 (1979-), 男, 在读博士, 研究实习员, 研究方向为大豆遗传育种。E-mail: Dragonyan1979@yahoo.com.cn。

通讯作者: 张孟臣, 研究员。E-mail: zhangmengchen@hotmail.com。

后代。在大豆品质改良过程中,需要对大量种质资源、突变体、杂交后代材料及时鉴定、筛选和分析,同时希望具有优良品质的籽粒经分析后能保持完好,以便进一步繁殖^[3-4]。因此,迫切需要一种准确、快速、简便、非破坏籽粒的检测方法。

近红外光谱技术(Near Infrared Spectroscopy, NIS)是20世纪80年代后期发展起来的一项物理测试技术^[5]。它利用有机物在近红外光谱区的振动吸收而快速测定样品中多种化学成分含量,蛋白质、脂肪、糖类等含有的各种含氢基团(CH、OH、NH、SH等)的倍频与合频谱带恰好落在近红外区,以得到这些有机分子含氢基团的特征振动信息,从而测定其化学成分的含量^[6-7],已成为水稻、小麦、玉米等农作物品质分析的重要手段^[8-10]。与化学分析相比,近红外光谱技术是一项间接的分析方法,它测定样品成分含量的方法是建立在化学分析法或其它仪器测定基础之上,是一种可“再生”的测定方法。因此,任何一台近红外光谱仪对每种组分或每种参数都要单独定标。所以,利用近红外检测大豆籽粒中粗蛋白和粗脂肪含量,关键是构建正确的模型,并验证模型的可靠性^[11]。

研究旨在利用不同的建模样品集分别构建近红外检测模型,以化学分析结果作参照,探求近红外检测代替化学分析的适应范围。

1 材料与方法

1.1 近红外检测模型建立

采用4个建模样品集,由权威检测中心(以M2表示)依照GB2905-1982和GB2906-1982获得化学值。近红外检测采用德国Bruker公司的Matrix-I傅立叶近红外光谱分析仪,取适量大豆籽粒置于样品杯中,用透射方式扫描获得大豆籽粒样品的近红外吸收光谱图,扫描谱区4 000~12 600 cm⁻¹,分辨率16 cm⁻¹,扫描速度10 KHz,扫描次数60次,每个样

品重复装样扫描3次。利用Bruker公司的OPUS4.2软件的Quant 2方法,以偏最小二乘法(PLS)建立大豆粉末、籽粒近红外检测粗蛋白及粗脂肪含量的模型,使用交叉检验法对模型做出评价。

1.2 近红外检测模型可靠性比较

1.2.1 材料 选用8个大豆品种(系),粗蛋白含量35%~45.5%,粗脂肪含量在16.2%~22.5%,其中,1、2、3、5、7、8号材料是黄种皮材料,4号是青种皮材料,6号是黑种皮材料。

1.2.2 检测方法 将每个参试品种种子充分混匀分成3等份(即设重复3次),随机编为24个样品号。24份样品分别用研究中建立的3种近红外籽粒检测模型(以M5、M6、M7表示)和近红外粉末检测模型(以M4表示)检测粗蛋白、粗脂肪含量;同时送检包括农业部 and 河北省指定的权威检测部门用化学法检测(以M1、M2、M3表示)。

1.2.3 统计方法 通过方差分析比较各结果间变异系数CV的差异以评价稳定性。以方差分析、相关分析衡量不同检测结果趋势的一致性。相关分析以平均数为准。方差分析、相关分析由SAS6.12软件完成。

2 结果与分析

2.1 近红外模型建立

近红外检测模型建模样品集样品数量分别为30、77、415和161个(表1)。其中,粉末模型样品覆盖范围相对较窄,粗蛋白决定系数0.9788,粗脂肪决定系数0.9601,在4个模型中最大,交叉验证均方根分别为粗蛋白0.439,粗脂肪0.287,在4个模型中最小。籽粒模型M6的样品集中包括415份样品,且粗蛋白和粗脂肪含量的样品范围在4个模型中最宽。它的粗蛋白和粗脂肪决定系数在3个籽粒模型中最大,交叉验证均方根在3个籽粒模型中最小。

表1 四种近红外模型参数比较
Table1 Statistical parameters of 4 models

模型 Model	样品份数 No. of sample	样品范围 Range of sample		决定系数 R ²		交叉验证均方根 RMSECV	
		粗蛋白 Protein	粗脂肪 Oil	粗蛋白 Protein	粗脂肪 Oil	粗蛋白 Protein	粗脂肪 Oil
M4 粉末模型,PM	30	33.34-46.51	16.63-22.91	0.9788	0.9601	0.439	0.287
M5 籽粒模型,SM	77	33.05-48.40	16.25-23.78	0.9376	0.8168	0.942	0.691
M6 籽粒模型,SM	415	33.05-49.80	16.10-24.20	0.9665	0.9559	0.649	0.397
M7 籽粒模型,SM	161	33.05-48.40	16.25-23.78	0.9287	0.8404	1.080	0.668

R²:coefficient of determination;RMSECV:root mean square error of cross validation;PM:powder model;SM:seed model.

2.2 不同检测结果的稳定性分析

以重复观察值的变异系数 CV 作为检测结果稳定性的衡量指标(表2)。粗蛋白含量3种籽粒模型检测结果的 CV 平均0.008,粉末模型 CV 为0.019,3个化学分析结果的 CV 平均0.014;粗脂肪含量3种籽粒模型检测结果的 CV 平均0.008,粉末模型 CV 为0.024,3个化学分析 CV 平均0.013。结果表明,在检测粗蛋白、粗脂肪含量时,3种籽粒模型检测结果均有较高的稳定性。

表2 不同粗蛋白、粗脂肪含量检测结果精确性比较
Table 2 Comparison of CV among different results by ANOVA

方法 Method	变异系数平均值 Mean of CV	
	蛋白变 Protein	油份 Oil
M1	0.021 A	0.013 AB
M2	0.006 B	0.011 AB
M3	0.014 AB	0.014 AB
M4	0.019 AB	0.024 A
M5	0.009 AB	0.009 AB
M6	0.007 B	0.008 B
M7	0.008 B	0.007 B

CV :coefficient of variation;ANOVA;analysis of variance

2.3 不同检测结果的一致性分析

2.3.1 方差分析 从8个参试品种的总体结果看(表3),3个化学检测中心粗蛋白含量检测结果存

在显著差异,粗脂肪含量检测结果无显著差异。近红外籽粒模型 M5、M6 和 M7 与其建模数据来源化学分析 M2 的粗蛋白含量和粗脂肪含量检测结果在0.01 极显著水平下均无差异。

表3 不同粗蛋白、粗脂肪含量检测结果方差分析
Table 3 ANOVA analysis among different results

方法 Method	粗蛋白 Protein	粗脂肪 Oil
M1	39.37 C	19.77 BC
M2	40.76 AB	19.93 ABC
M3	38.13 D	19.53 C
M4	39.91B C	20.24 AB
M5	41.42 A	20.44 AB
M6	40.89 A	20.17 ABC
M7	41.14 A	20.50 A

同一列不同大写字母表示差异达极显著水平。下同。
Values followed by different capital letters are significantly different. Same as below.

将7种检测方法对8个参试品种的粗蛋白、粗脂肪检测结果进行方差分析。分析结果(表4、表5)显示,3个化学检测结果存在差异;M5和M7测得参试黑种皮品种6粗蛋白含量偏高,3个近红外籽粒模型对其他7个参试品种的检测结果无显著差异,且都与它的建模样品数据来源的化学分析M2结果一致。

表4 粗蛋白检测结果全品种(系)比较
Table 4 Comparison of protein content among all samples

方法 Method	品种(系) Variety(line)							
	1	2	3	4	5	6	7	8
M1	41.42A	37.12AB	37.83A	42.49AB	37.05B	39.29E	40.29B	39.45B
M2	42.38A	38.36A	39.20A	43.15A	38.05A	41.63C	42.09A	41.22A
M3	39.59B	35.96AB	37.37A	40.20C	35.33C	39.18E	39.21C	38.21C
M4	41.40A	38.24A	39.51A	40.93BC	37.01B	40.37DE	42.06A	39.80B
M5	42.42A	38.61A	39.66A	43.62A	38.31A	45.04A	42.85A	40.84A
M6	42.52A	38.78A	39.43A	43.13A	38.79A	40.85CD	42.54A	41.09A
M7	42.98A	38.17A	38.85A	43.71A	38.71A	43.22B	42.20A	41.27A

表5 粗脂肪检测结果全品种(系)比较
Table 5 Comparison of oil content among all samples

方法 Method	品种(系) Variety(line)							
	1	2	3	4	5	6	7	8
M1	18.16DE	21.36CD	20.91B	18.10AB	21.67C	20.24B	19.26C	18.42C
M2	18.32CD	21.53BCD	21.94AB	18.18AB	21.78BC	19.91BC	19.43BC	18.37C
M3	17.71E	21.20D	20.89B	17.91B	21.50C	19.57CD	19.28C	18.18C
M4	18.38CD	21.78ABC	21.58AB	17.64B	21.87BC	21.29A	20.53A	18.84BC
M5	18.98AB	22.01AB	22.18A	18.65AB	22.36A	19.66BC	19.73BC	19.92A
M6	18.83BC	21.54BCD	21.39AB	18.73AB	21.56C	20.11BC	19.54BC	19.68AB
M7	19.52A	22.15A	21.76AB	19.45A	22.18AB	19.02D	19.92B	20.04A

2.3.2 相关分析 各检测方法对不同品种检测结果大小趋势的一致性研究者更为关心的问题。为了直观地观察不同检测结果的趋势是否相同,以 M2 对 8 个参试品种的粗蛋白、粗脂肪含量结果为横坐标, M5、M6、M7 对 8 个参试品种的粗蛋白、粗脂肪检测结果为纵坐标,分别绘图(图 1)。从中看出, M5、M6、M7 对 8 个参试品种的粗蛋白检测结果较粗脂肪检测结果与化学分析结果 M2 存在更明显的

线性关系。M6 与 M2 的线性关系强于 M5、M7。在 M5 和 M7 的结果中,明显看到对品种 6 的粗蛋白检测结果偏离线性关系,且近红外结果偏高。相关分析表明,3 个籽粒模型的粗蛋白检测结果与 M2 检测结果呈极显著正相关,其中 M6 与 M2 化学分析的相关系数最高。从粗脂肪含量检测结果看,3 个籽粒模型与 M2 达极显著正相关,同样以 M6 与 M2 化学分析的相关系数最高。

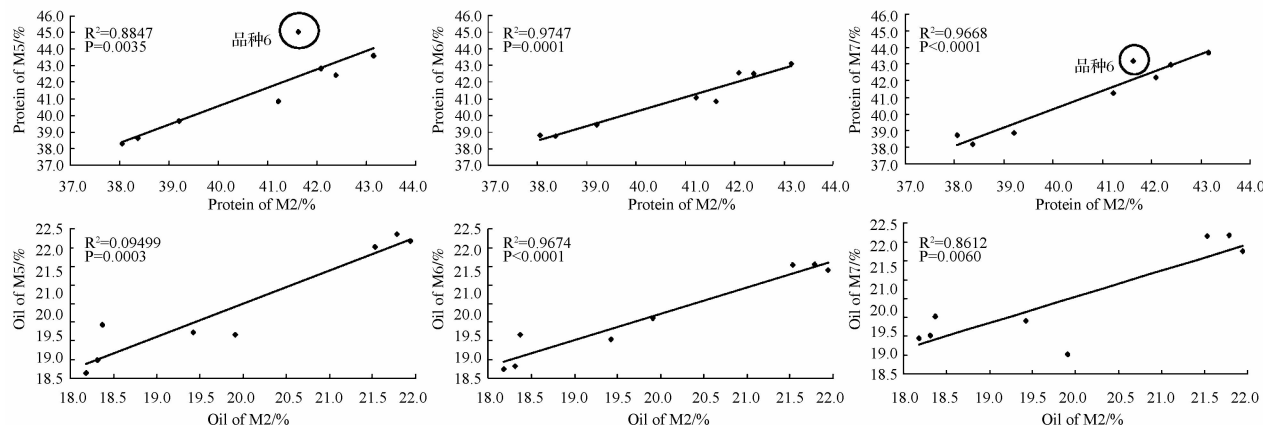


图 1 近红外籽粒模型与 M2 粗蛋白、粗脂肪含量检测结果比较

Fig. 1 The trends of protein and oil content among NIS seed model and M2 methods

3 结论与讨论

利用近红外技术检测大豆粗蛋白、粗脂肪含量的前提条件是尽可能利用样品数量多、覆盖范围广、样品类型丰富的样品集构建近红外检测模型。所建的模型中以建模样品集样品量最大的 M6 可靠性最高。这与吴金红等^[7]建立的水稻蛋白质含量近红外检测模型的结果一致。所以,在建立近红外检测模型时,应尽可能的利用覆盖范围宽的众多样品作为样品集。

所建模型决定系数最小为 0.8168,最高达 0.9788,平均为 0.9231。吴金红等^[11]建立的 3 个水稻蛋白质含量近红外检测模型决定系数最小为 0.805,最高达 0.909,平均为 0.8714。方彦等^[12]建立的玉米籽粒含油量近红外检测模型决定系数达到 0.9176。尽管近红外模型决定系数较高,但通常人们更乐于接受化学分析的结果,而对近红外检测结果持怀疑态度。分析结果表明,近红外籽粒模型的变异系数小于化学分析的变异系数,与姚鑫森等^[7]利用近红外检测技术分析大豆品质,得出近红外结果的标准偏差均小于化学分析结果的标准偏差,且经过 F 检验和 t 检验,两种方法之间无显著性差异

的结果一致。化学分析过程中人工操作的步骤远多于近红外籽粒模型检测时人工操作的步骤,是降低化学分析稳定性的可能原因。

从结果的一致性看,近红外结果与它的建模样品数据来源的化学分析结果一致,而且 3 个籽粒模型与 M2 对不同品种检测结果的一致性优于 3 个化学分析(M1、M2、M3)结果的一致性,特别是 M6 与 M2 的相关系数最高。表明利用样品数量多、覆盖范围广、样品类型丰富的样品集构建近红外检测模型检测大豆粗蛋白、粗脂肪含量,能得到与化学分析一致的检测结果,具有较好的稳定性和可靠性。在需要保存籽粒完好的育种材料低代选择、QTLs/基因定位、导入系筛选和突变体筛选等的品质研究中,近红外检测是替代化学分析检测的有效方法。

近红外光谱特征除与蛋白、脂肪相关外,籽粒大小、籽粒颜色等也是重要的影响因素。通过全部品种内的方差分析比较表明,近红外检测可能使黑种皮品种的粗蛋白检测结果偏高,原因可能是深色种皮对近红外光谱吸收有一定影响,而建立模型的样品集又以黄种皮材料为主。所以,在构建近红外模型时,应根据种皮色的不同构建专用模型。

参考文献

- [1] 邱丽娟,常汝镇.大豆种质资源描述规范和数据标准[M].北京:中国农业出版社,2006. (Qiu L J, Chang R Z. Descriptors and data standard for soybean (*Glycine* spp.) [M]. Beijing: Agricultural Press, 2006.)
 - [2] 李卫东,张孟臣.黄淮海夏大豆及品种参数[M].北京:中国农业科学技术出版社,2006. (Li W D, Zhang M C. Summer soybean variation in Huanghuaihai distinct [M]. Beijing: China Agricultural Science Press, 2006.)
 - [3] Xu Y B, Crouch J H. Marker-assisted selection in plant breeding: from publications to practice [J]. Crop Science, 2008, 48: 391-407.
 - [4] Li Z K, Fu B Y, Gao Y M, et al. Genome-wide introgression lines and their use in genetic and molecular dissection of complex phenotypes in rice (*Oryza sativa* L.) [J]. Plant Molecular Biology, 2005, 59: 33-52.
 - [5] Blanco M, Villarroya I. NIR spectroscopy: a rapid-response analytical tool [J]. Trends in Analytical Chemistry, 2002, 21 (4): 240-250.
 - [6] 刘晓庚.近红外光谱技术在粮油储藏及其品质分析中的应用[J].粮食储藏,2007(1):29-35. (Liu X G. Application of the near infrared spectroscopy (NIRS) in food and oil storage and its quality analysis [J]. Food Storage, 2007(1): 29-35.)
 - [7] 姚鑫淼,张瑞英,李霞辉,等.近红外透射光谱法(NITS)分析大豆品质的研究[J].大豆科学,2006,25(4):417-424. (Yao X M, Zhang R Y, Li X H, et al. Study on analysis of soybean quality by near infrared transmittance spectroscopy [J]. Soybean Science, 2006, 25(4): 417-424.)
 - [8] 李君霞,张洪亮,严衍禄,等.水稻蛋白质近红外定量模型的创建及在育种中的应用[J].中国农业科学,2006,39:836-841. (Li J X, Zhang H L, Yan Y L, et al. Establishment of math models of NIRS analysis for protein contents in seed and its application in rice breeding [J]. Scientia Agricultura Sinica, 2006, 39: 836-841.)
 - [9] Delwiche S R. Protein content of single kernels of wheat by near-infrared reflectance spectroscopy [J]. Journal of Cereal Science, 1998, 27: 241-254.
 - [10] 魏良明,严衍禄,戴景瑞.近红外反射光谱测定玉米完整籽粒蛋白质和淀粉含量的研究[J].中国农业科学,2004,37:630-633. (Wei L M, Yan Y L, Dai J R. Determining protein and starch content of whole maize kernel by NIRS [J]. Scientia Agricultura Sinica, 2004, 37: 630-633.)
 - [11] 吴金红,张洪江,梅捍卫,等.水稻蛋白质含量 NIR 模型适配范围的研究[J].中国农业科学,2006,39(12):2435-2440. (Wu J H, Zhang H J, Mei H W, et al. Study on adaptability of NIR models of protein content in rice [J]. Scientia Agricultura Sinica, 2006, 39: 2435-2440.)
 - [12] 方彦,王汉宁.利用近红外光谱法测定玉米籽粒含油量的研究[J].西北农业学报,2007,16(1):111-113. (Fang Y, Wan H N. Studies on the determining of oil contents for maize kernels using near infrared spectroscopy [J]. Acta Agriculturae Boreali-occidentalis Sinica, 2007, 16(1): 111-113.)
-
- (上接第 832 页)
- [4] 胡晓娟,黄晓书,李卫华,等.高效液相色谱法测定保健食品中的大豆异黄酮[J].分析实验室,2008,27(1):103-106. (Hu X J, Huang X S, Li W H, et al. Determination of soybean isoflavone in health food by high performance liquid chromatography [J]. Chinese Journal of Analysis Laboratory, 2008, 27(1): 103-106.)
 - [5] Stephen P, Klump J. Determination of isoflavones in soy and selected foods containing soy by extraction, saponification, and liquid chromatography: collaborative study [J]. Journal of AOAC International, 2001, 84(6): 1865-1883.
 - [6] 李淑云,赵福芹,李月雄.高效液相色谱法测定大豆异黄酮粉中大豆苷和染料木苷的含量[J].中国现代应用药学杂志,2007,24(8):716-717. (Li S Y, Zhao F Q, Li Y X. Determination of daidzin and genistin in soybean isoflavone flour by HPLC [J]. The Chinese Journal of Modern Applied Pharmacy, 2007, 24 (8): 716-717.)
 - [7] 蒋子刚,顾雪梅.分析测试中的数理统计与质量保证[M].上海:华东化学学院出版社,1991. (Jiang Z G, Gu X M. Statistics and quality assurance in analytical testing [M]. Shanghai: East China University of Science and Technology Press, 1991.)
 - [8] 郝青南,马超,马兵钢.大豆异黄酮的生理功能及其分离检测方法研究进展[J].中国医药生物技术,2007,2(5):383-386. (Hao Q N, Ma C, Ma B G. Progress of physiological function as well as separation and testing method of soybean isoflavone [J]. Chinese Medicinal Biotechnology, 2007, 2(5): 383-386.)
 - [9] 潘志芬,潘开文.大豆及大豆制品中异黄酮的研究现状[J].四川农业大学学报,1999,17(2):217-223. (Pan Z F, Pan K W. Current researches on isoflavones in soybean and soybean products [J]. Journal of Sichuan Agricultural University, 1999, 17 (2): 217-223.)