

黍稷籽粒淀粉含量的近红外光谱分析

赵婷婷¹, 张丽珍², 郑淑华², 杨良良², 冯耐红³, 刘根科⁴

(1. 山西大学 生物工程学院, 山西 太原 030006; 2. 山西大学 生命科学院, 山西 太原 030006;

3. 山西省农业科学院 经济作物研究所, 山西 汾阳 032200; 4. 山西省农业科学院 右玉农业试验站, 山西 右玉 037200)

摘要: 利用 FOSS1241 型近红外谷物分析仪对 100 份黍稷籽粒原样进行光谱扫描, 分别采用偏最小二乘法和改进偏最小二乘法两种定标方法, 并对原始光谱分别进行不同的预处理, 建立了黍稷籽粒淀粉定标模型。试验结果显示: 对黍稷籽粒原样扫描的光谱采用偏最小二乘法进行定标, 在标准正常化 + 趋势变化散射处理、采用一阶导数、每隔 4 点求导、每隔 4 点作平滑处理的光谱预处理下建立的模型较好。近红外透射光谱分析技术为定量检测黍稷籽粒淀粉总量提供了一种新方法。

关键词: 近红外光谱; 黍稷籽粒; 总淀粉; 定量分析模型

中图分类号: S123 文献标识码: A 文章编号: 1000-7091(2011)01-0234-05

Analysis of Starch Content of *Panicum miliaceum* L. Seed by Near Infrared Transmittance Spectroscopy

ZHAO Ting-ting¹, ZHANG Li-zhen², ZHENG Shu-hua², YANG Liang-liang²,
FENG Nai-hong³, LIU Gen-ke⁴

(1. Institute of Biological Engineering, Shanxi University, Taiyuan 030006, China;

2. College of Life Science, Shanxi University, Taiyuan 030006, China;

3. Institute of Economic Crop, Shanxi Academy of Agricultural Sciences, Fenyang 032200, China;

4. Youyu Experimental Station, Shanxi Academy of Agricultural Sciences, Youyu 037200, China)

Abstract: Calibration models for starch content of *Panicum miliaceum* L. Seed was built using near-infrared transmittance spectroscopy (NITS) with partial least squares (PLS) and modified partial least squares (PLS). One hundred parts seed sample were scanned by the FOSS 1241 NIR Grain Analyzer. Prior to model establishing, the original spectra were processed with different pretreatment methods. The result showed that the best calibration model was to adopt partial least squares (PLS) with SNV + Detrend and 1.4.4.1 for mathematics method. Near infrared spectroscopy could be used to achieve the quantitative detection of starch of *panicum miliaceum* L. Seed.

Key words: Near-infrared transmittance spectroscopy; *Panicum miliaceum* L. seed; Total starch; Calibration model

黍稷(*Panicum miliaceum* L.) 属禾本科黍属, 一年生栽培谷物, 是北方年糕、蒙族炒米的主要原料, 食药同源, 保健性好。按籽粒的黏与不黏划分为糯性的黍和粳性的稷(糜)。淀粉含量是影响糯性和粳性的一个重要指标, 其直链淀粉含量越高, 糯性越小。稷(糜)淀粉颗粒形状呈双分布, 即小球形和大多角形^[1], 这种淀粉具有抗性淀粉的性质, 适用于

糖尿病和心血管病患者食用^[2]。现有的淀粉分析往往采用碘显色法, 必须破坏样品籽粒, 具有操作时间长, 处理步骤繁琐等缺陷^[3], 因此迫切需要建立一种快速无损测定黍稷中淀粉含量的方法来满足其品种选育与加工开发研究。

近红外光谱技术是利用有机物质在近红外谱区的振动吸收而快速测定样品中多种化学成分含量的一项技术, 其中透射谱区的近红外光在样品的穿透

收稿日期: 2010-09-15

基金项目: “十一五”国家科技支撑计划项目(2006BAD02B07)

作者简介: 赵婷婷(1984-), 女, 山西洪洞人, 在读硕士, 主要从事植物种质资源利用的研究。

通讯作者: 张丽珍(1977-), 女, 山西五台人, 副教授, 博士, 主要从事生物资源利用研究。

能力可达 30 mm, 适合整粒或原样样品分析^[4]。国内外学者利用近红外透射光谱技术先后对包附有坚硬外壳的籽粒, 如玉米^[5-6]、向日葵^[7]、油菜籽^[8-10]、水稻种子^[11]、大豆^[12-14]、甘草种子^[15] 等不同作物的单粒种子进行了化学品质、活力、真实性、颜色等方面的研究。鉴于黍稷籽粒外层覆盖物为光滑坚硬的稃片结构和黄红灰褐白棕黑及复色的多样化籽粒颜色, 本试验拟研究近红外透射光谱技术是否适用于黍稷籽粒淀粉含量的分析, 并尝试利用近红外光谱法建立黍稷籽粒中淀粉含量的无损测定方法, 促进该技术在黍稷育种中的应用。

1 材料和方法

1.1 材料

供试的 100 份黍稷材料由山西省农业科学院右玉农业试验站提供。这些材料的 11 个农艺性状中, 包括花序色、穗型、粒色、米色、抗倒伏性、株高、主穗长、主茎节数、单株粒质量、千粒质量、生育期, 均有较大的差异, 黍稷籽粒具有黄红灰褐白棕黑等颜色, 因此具有较好的代表性。

1.2 仪器

FOSS 公司生产的 FOSS1241 近红外谷物分析仪, 上海光谱 SP-756 紫外可见分光光度计。

1.3 方法

1.3.1 光谱测定 为了获得最佳的模型建立及预测效果, 需要先对用于扫描黍稷样品的扫描光程进行优化和选择。在选定的扫描光程下, 于 570 ~ 1 850 nm 的扫描范围内对上述样品进行近红外光谱扫描采集。

1.3.2 化学测定 采用酸水解法, 根据双波长比色原理, 从碘-直链淀粉和碘-支链淀粉复合物的吸收光谱中, 选择直链淀粉的测定波长 600 nm, 参比波长 503 nm; 支链淀粉的测定波长 580 nm, 参比波长 700 nm, 依据回归方程求出直链和支链淀粉的含量。支链淀粉与直链淀粉含量的和即为总淀粉含量。

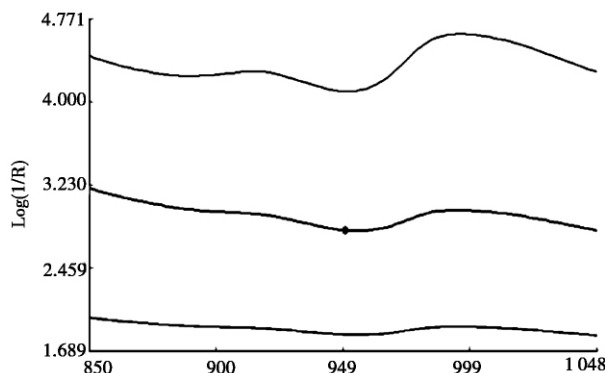
1.4 数据分析

在使用近红外光谱进行品质检测时, 对光谱信号进行消除噪音等预处理十分必要。本试验利用 WINISI III 专业建模软件系统对采集的光谱数据进行标准正常化处理 + 去散射处理, 导数光谱分别采取一阶导数和二阶导数处理, 每隔 4 点求导, 每隔 4 点作平滑处理的光谱处理方式对原始光谱图进行数学的处理后, 利用主成分分析(PCA) 技术自动剔除超常样品, 并选择有代表性的样品, 用偏最小二乘法(PLS) 和改进的偏最小二乘法(MPLS) 进行定标建模。

2 结果与分析

2.1 扫描光程的选择

图 1 是分别在光程 6, 10, 18 mm 下扫描一份黍稷样品得到的光谱图, 在保证近红外光谱仪检测器的吸光度线性满足的条件下, 尽可能选择最大的扫描光程。本研究中选定的试验测试用光程为 10 mm, 在此光程下可保证光谱的不失真, 又保证了各个组分的最大吸光度吸收要求。



图中中间的光谱为 10 mm 光程下扫描。

The intermediate spectra was transmitted under 10 mm distance.

图 1 不同光程下的黍稷样品的扫描光谱

Fig. 1 Sample transmittance spectra under different spectra distance

2.2 光谱扫描及光谱数据分析与处理

在选定的 10 mm 光程条件下, 扫描采集的 100 份黍稷样品, 100 份样品的原始光谱图如图 2 所示。从图 2 可以看出, 收集到的近红外样品光谱在一个很宽的背景下有明显的吸收但特征峰却很小。为消除光谱中的噪声信息, 提高近红外回归分析方程的可靠性和准确性, 光谱分析采用了 WinISI 软件具备的标准正常化处理 + 散射处理数学方法, 通过对原始光谱数据标准正常化处理, 扣除样品间因散射而引起的光谱误差; 通过去散射处理, 消除由光散射产生的线性基线和背景的影响。由于对光谱求导数, 可以使峰变陡, 平滑可以提高分析信号的信噪比。进一步选择参数 1.4.4.1(指求一阶导数, 每隔 4 点求导, 每隔 4 点作平滑处理) 和 2.4.4.1(指求二阶导数, 每隔 4 点求导, 每隔 4 点作平滑处理) 对光谱进行优化, 剔除异常光谱。一阶导数和二阶导数光谱图如图 3 和图 4 所示。

2.3 定标模型的建立

运用偏最小二乘法(PLS) 和改进的偏最小二乘法(MPLS) 两种算法, 结合不同的光谱处理分别建立黍稷中总淀粉、直链淀粉和支链淀粉几种成分的定标模型, 采用交互验证方法来测试模型的预测精确性。表 1 列出了构建定标模型样的定标相关系数

(RSQ)、定标准偏差(SEC)、交叉验证标准误差(SEC_V)以及交叉验证的相关系数(1-VR)。对于同

一样品集所构建的近红外模型来说,各类决定系数越大,标准误差越小,表明模型的定标及预测效果越好。

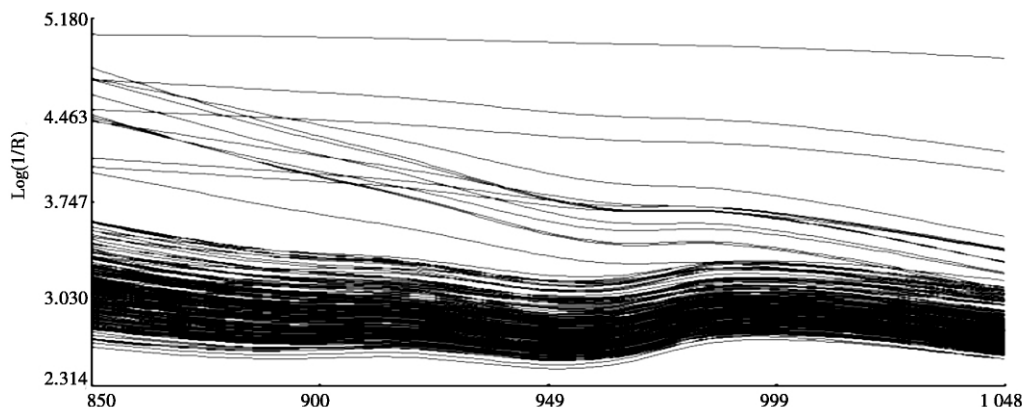


图2 100份黍稷的NIRS连续扫描光谱

Fig.2 Original transmittance spectra of calibration set

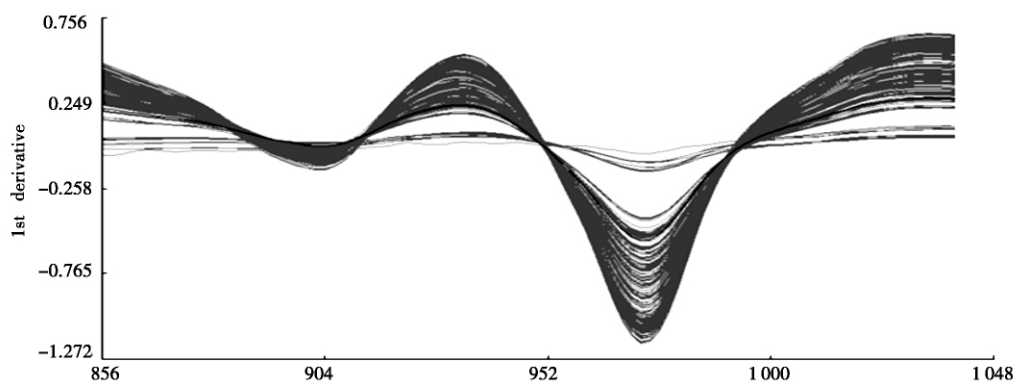


图3 一阶导数后的光谱

Fig.3 Spectrum treated by first derivate

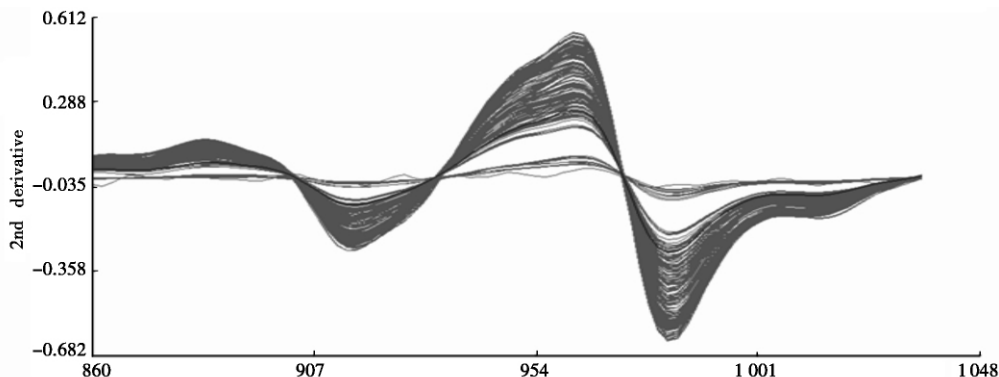


图4 二阶导数后的光谱

Fig.4 Spectrum treated by second derivate

从表1可以看出,光谱经过一阶导数处理后采用偏最小二乘法的总淀粉各项建模指标较好。近红外光谱分析方法适用于总淀粉含量的分析,但不适用于直链淀粉和支链淀粉的分析。为满足生产实践中的测量要求,仍需对不同籽粒颜色的黍稷样品分类建模,以进一步提高检测精度和减小预测误差。

近红外光谱分析法不适用于直链淀粉和支链淀粉分析的原因可能是:由于近红外分析方法预测效果的好坏直接取决于化学测量值的准确性。本试验

的化学测量值采用碘比色法,碘比色法是测定淀粉含量的经典方法。对于黍稷的直链淀粉和支链淀粉含量,没有确定的国家标准,本研究参照水稻直链淀粉含量国家标准(GB7648-87)进行。直链淀粉与碘生成深蓝色复合物,支链淀粉与碘生成棕红色复合物。在淀粉总量不变的条件下,将这两种淀粉分散液按不同比例混合,在一定波长和酸度条件下与碘作用,生成由紫红到深蓝一系列颜色,代表不同的直链淀粉与支链淀粉含量比例,根据朗伯-比尔定律,

用比色法测定直链淀粉与支链淀粉含量^[16]。不同标准品直接影响直链淀粉和支链淀粉含量的测定结果,本研究使用的标准样品是 Fluke 公司的支链标品和 Sigma 公司的直链标品,但黍稷籽粒淀粉的理化性状与马铃薯淀粉有较大区别^[2]。没有确定的

国家标准和缺乏对应的标准样品带来了黍稷籽粒直链淀粉和支链淀粉的化学检测数据的误差,进而影响到近红外光谱分析的适用性。Barton 等^[17]用近红外光谱(NIR) 技术对大米直链淀粉含量的进行测定,效果较差。

表 1 不同算法及光谱处理近红外定标模型结果

Tab.1 Results of different regression type and pretreatments on calibration						
性状 Traits	回归类型 Regression type	光谱处理 Spectral treatment	RSQ	SEC	SECV	1-VR
淀粉 Starch	偏最小二乘法	一阶导数	0.660 9	5.866 6	6.059 3	0.640 6
		二阶导数	0.615 7	5.784 8	6.008 1	0.587 2
	改进偏最小二乘法	一阶导数	0.655 9	5.543 6	5.727 9	0.635 7
		二阶导数	0.548 0	5.587 1	5.954 6	0.493 6
直链 Amylose	偏最小二乘法	一阶导数	0.116 8	1.996 5	2.039 6	0.079 2
		二阶导数	0.171 5	1.900 9	1.950 0	0.127 0
	改进偏最小二乘法	一阶导数	0.167 3	1.935 5	2.048 9	0.068 9
		二阶导数	0.076 3	2.198 4	2.244 1	0.035 5
支链 Amylopectin	偏最小二乘法	一阶导数	0.366 8	4.437 3	4.552 8	0.337 6
		二阶导数	0.193 3	4.412 9	4.617 3	0.126 0
	改进偏最小二乘法	一阶导数	0.376 4	4.394 4	4.555 9	0.331 9
		二阶导数	0.099 0	0.099 0	4.686 3	0.077 1

注: RSQ. 定标相关系数; SEC. 定标标准偏差; SECV. 交互验证标准偏差; 1-VR. 交互验证相关系数。
Note: RSQ. Calibration coefficient; SEC. Standard deviation of calibration; SECV. Standard deviation of cross validation; 1-VR. Correlation coefficient of cross validation.

2.4 定标模型的检验与评价

定标模型建立后,需要用预测集样品对新建立的定标模型进行检验。通过选择一批与建模样本集独立无关的样本,比较独立样本预测值与化学值的差异来判断模型的预测准确性。本研究中,在建立

好不同指标模型后,选取 30 份未参与定标的样品作为盲样,对各个指标的预测模型的预测性能予以评价,来进一步验证所建模型的可靠性。校正模型的外部检验结果如表 2 所示。

表 2 外部验证结果

Tab.2 External validation results					
性状 Traits	RSQ	Slope	Bias	SEP	SEPC
总淀粉 Starch	0.744	1.012	0.583	5.255	5.311
直链淀粉 Amylose	0.114	1.706	2.289	5.688	5.299
支链淀粉 Amylopectin	0.491	1.022	0.610	4.408	4.440

近红外分析结果与实验室常规分析结果的相关系数,总淀粉、直链淀粉、支链淀粉分别是 0.774, 0.114 0.491。由预测效果来看,总淀粉、直链淀粉、支链淀粉预测模型对于盲样的预测误差(SEP) 分别为 5.255 5.688 4.408,这个实际盲样预测效果和建立定标时的评估效果是一致的。近红外透射光谱分析技术为定量检测黍稷籽粒淀粉总量提供了一种新方法。

3 结论

在收集样品时,因希望样品能充分代表自然样品的随机性和多样性,没有对样品进行人工筛选,仅按常规风选除杂处理,籽粒颜色有红、黄、灰、褐、白、棕和黑色。在进行光谱评价的时候,主要利用自动筛选,仅删除了发生报警现象的黑色样品光谱信号。

总淀粉含量值分布广,最大值为 65.35%,最小值为 12.25%,具有较好的代表性。

利用近红外透射光谱技术无损检测具有硅质稗片且颜色多样的黍稷籽粒总淀粉含量是可行的。在建模过程中探讨数学导数、回归方法等优化处理对建模的影响,结果表明,一阶导数优于二阶导数,偏最小二乘法优于改进的偏最小二乘法。光谱经过一阶导数处理后采用偏最小二乘法所建立的模型中,预测相关系数为 0.640 6,预测误差为 6.059 3。对 30 份样品进行外部检验,预测值与实测值之间的检验相关系数为 0.744,定标模型的预测相关性较好,但精确性有待提高。在实际应用中,是否需要对各种颜色的样品给出各自的模型,确定每个模型具体的应用评价方案,尚有待进一步研究。

本研究中利用近红外透射光谱技术建立的模型

不适合单独检测支链淀粉含量和直链淀粉含量。黍稷淀粉颗粒由直链淀粉和支链淀粉构成,直链淀粉与支链淀粉的比例因植物来源和品种的不同而有很大差异,其中直链淀粉含量为 0~23.2%,支链淀粉含量为 12.25%~49.2%,并且受黍稷生长过程中气候和土壤条件的影响。由于 NITS 分析测试完全依赖于传统的分析方法,化学分析测试数据的准确性将直接决定建立模型的可靠性与适用性。因此,为了获得准确的化学测量值,需在深入研究黍稷籽粒淀粉性质的基础上获得其标准样品,提高原始数据的准确度可能会提高 NITS 的测试准确度。

参考文献:

- [1] 姚亚平,田呈瑞,张国权,等.糜子淀粉理化性质的分析[J].中国粮油学报,2009,24(9):45-51.
- [2] Kumari S K,Thayumanavan B. Characterization of starches of proso,foxtail,barnyard,kodo,and little millets [J]. Plant Foods for Human Nutrition,1998,53:47-56.
- [3] 肖 昕,谢新华,陈 奕,等.应用近红外透射光谱法测定水稻种子直链淀粉含量的初步研究[J].中国农业科学,2004,37(11):1709-1721.
- [4] 刘建学.实用近红外光谱分析技术[M].北京:科学出版社,2008.
- [5] 孟兆芳,赵龙莲,程 奕,等.近红外光谱法测定玉米品质指标的研究[J].华北农学报,2008,3(2):147-150.
- [6] 卢宝红,张义荣,魏良明,等.近红外分析方法在玉米轮回选择中的应用[J].山西农业科学,2005,33(2):28-31.
- [7] Leonardo Velasco,Begoña Pérez-Vich,José M *et al.* Non-destructive screening for oleic and linoleic acid in single sunflower achenes by near-Infrared reflectance spectroscopy [J]. Crop Sci,1999,39:219-222.
- [8] 吴建国,石春海,张海珍,等.应用近红外反射光谱法整粒测定小样品油菜籽含油量的研究[J].作物学报,2002,28(3):421-425.
- [9] 高建芹,张洁夫,浦惠明,等.近红外光谱法在测定油菜籽含油量及脂肪酸组成中的应用[J].江苏农业学报,2007,23(3):189-195.
- [10] 吴建国,石春海,张海珍.构建整粒油菜籽脂肪酸成分近红外反射光谱分析模型的研究[J].光谱学与光谱分析,2006,26(2):259-262.
- [11] 肖 昕,陈 奕,罗文永,等.单粒活体稻谷种子直链淀粉含量的近红外透射光谱分析[J].中国水稻科学,2003,17(3):287-290.
- [12] 闫 龙,蒋春志,于向鸿,等.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[J].大豆科学,2008,27(5):833-837.
- [13] 韩 熏,严红兵,卢福洁,等.基于国产近红外谷物分析仪的大豆中粗蛋白与粗脂肪含量的检测研究[J].现代科学仪器,2009(2):119-121.
- [14] 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.
- [15] 孙 群,李军会,王建华,等.乌拉尔甘草单粒种子硬实特性的近红外光谱分析[J].光谱学与光谱分析,2009,29(10):2669-2672.
- [16] 张巧杰,王一鸣,吴静珠,等.基于比色原理的直链淀粉测定仪设计与试验[J].农业机械学报,2005,36(7):81-84.
- [17] Barton F E,Himmelsbach D S,McClung A M,*et al.* Rice quality by spectroscopy analysis: precision of three spectral regions [J]. Cereal Chem,2000,77(5):669-672.