

乌拉尔甘草单粒种子硬实特性的近红外光谱分析

孙群¹, 李军会², 王建华¹, 孙宝启^{1*}

1. 中国农业大学农学与生物技术学院植物遗传育种系, 农业部基因组学与遗传改良重点实验室, 北京市作物遗传改良重点实验室, 北京 100193
2. 中国农业大学信息与电气工程学院, 北京 100193

摘要 以乌拉尔甘草种子为材料, 采用近红外光谱结合定性偏最小二乘法对 244 粒种子(硬实种子和非硬实种子比例为 1:1)的硬实性进行了鉴别研究, 并特制一样品杯用于单粒种子的光谱采集, 以降低人为误差。研究表明, 4 次重复平均光谱所建模型鉴别率显著高于单次光谱所建模型, 光谱范围采用 4 000~8 000 cm^{-1} 时模型效果较好, 校正集、检验集、预测集样本的鉴别率分别为 95.53%, 95.94% 和 94.53%, 采用不同建模样品所建模型其预测准确率均在 90% 以上, 硬实种子和非硬实种子的预测准确率分别为 92.50% 和 96.56%。种子大小和颜色均会影响模型的鉴别率, 种子颜色的影响相对更大。

关键词 近红外光谱; 硬实; 单粒种子; 乌拉尔甘草

中图分类号: S123 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2009)10-2669-04

引言

乌拉尔甘草(*Glycyrrhiza uralensis* Fisch.) 属于豆科蝶形花亚科, 是我国的一种传统中药材, 种子具有硬实特性。硬实现象广泛存在于豆科植物种子中, 具有重要的生态学和遗传学意义。报道表明硬实种子的活力和耐贮藏能力均要显著高于非硬实种子^[1, 2]。迄今为止所发现的长寿命种子均为硬实种子, 这归因于其坚硬的种皮起到了重要的保护作用。如在室温下保存的大豆硬实种子保存 4 年发芽率仍在 90% 以上, 而非硬实种子 2 年左右的时间即失去生活力^[3]; 古莲硬实种子寿命可达几百年甚至上千年, 而一般作物种子在同样的条件下贮藏不出数年便已不能发芽^[4]。因此对硬实种子种皮物理结构和化学成分进行分析, 促进其在种质贮藏上的利用具有重要的实践意义。但由于硬实种子和非硬实种子在外部形态(如色泽、大小)上没有差异, 迄今为止尚无有效的无损判定方法。目前判断种子是否为硬实时只能采用浸种法^[5], 即浸泡 24 h 后根据种子是否吸胀来判断种子的硬实性, 但浸泡后非硬实种子的种皮结构和化学成分均发生了不同程度的改变, 无法对硬实种子和非硬实种子在原始状态下进行比较, 目前对硬实机制进行研究时一般采用硬实率不同的品种进行品种间的比较, 具有一定的局限性。

近红外光谱(NIR)分析具有非破坏性、速度快、成本低等诸多优点。国内外学者利用近红外技术先后对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^[6-9]。如 Tigabu(2004)利用近红外透射光谱法将无活力空瘪种子和虫害种子从种子批中剔除, 其准确率高达 90%^[10], Tigabu(2003)利用近红外光谱法对 *Pinus patula* 的种子进行逐粒分析, 所建立的偏最小二乘法模型识别好种子的准确率达到 96%, 识别空壳种子的准确率达到 88%^[11]。Wang 用近红外对单粒小麦的颜色进行分类^[12]。利用近红外光谱技术对紫苏子及其伪品进行了直接鉴定, 其红外光谱的吸收位置和强度均有明显差异。

近红外光谱的应用非常广泛, 它为豆科作物种子硬实特性的无损判定提供了一种可能性。本研究拟利用近红外光谱法尝试建立乌拉尔甘草种子硬实特性的无损鉴定方法, 以促进豆科作物种子硬实机制的研究。

1 材料与方法

1.1 实验材料

实验所用的材料为 2007 年收获于宁夏的乌拉尔甘草种子。种子直径为 2.5~3.0 mm、颜色为棕绿色, 由于非硬实种子的比例较低(约 20%左右), 为保证获得足量的非硬实种

收稿日期: 2008-10-08, 修订日期: 2009-01-12

基金项目: 国家科技攻关课题项目(2001BA701A59)和国家科技支撑计划项目(2006BA106A15)资助

作者简介: 孙群, 女, 1971 年生, 中国农业大学农学与生物技术学院副教授 e-mail: sunqun@cau.edu.cn

* 通讯联系人 e-mail: sqcau@126.com

子, 一共选用 800 粒种子进行光谱采集。同时选取种子直径分别为 3.0~3.5 mm 和 2.0~2.5 mm 的棕绿色种子各 50 粒, 2.5~3.0 mm 的黑色种子 50 粒进行光谱采集, 以确定所建近红外模型的适应范围。

1.2 方法

1.2.1 采集光谱前的准备工作

采用手持光纤探头进行光谱采集时, 重复性较差, 为了消除每次手持位置的不确定性, 我们将光纤探头固定, 探头向下, 距离桌面 20 cm 左右。为了提高种子位置的稳定性, 我们根据光纤探头的直径和甘草种子的直径特制了一个不锈钢样品杯(结构见图 1、图 2), 样品杯内壁直径为 10.0 mm, 杯内小凹槽为浅球形, 深度为 2.0 mm, 直径为 3.5 mm, 略大于甘草种子直径, 以方便种子的取放。样品杯内部为黑色,

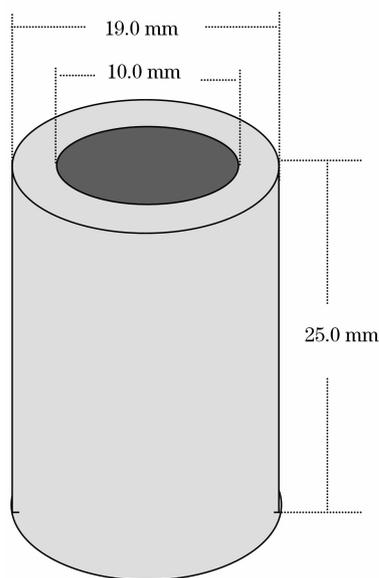


Fig. 1 External sketch of the sample cup

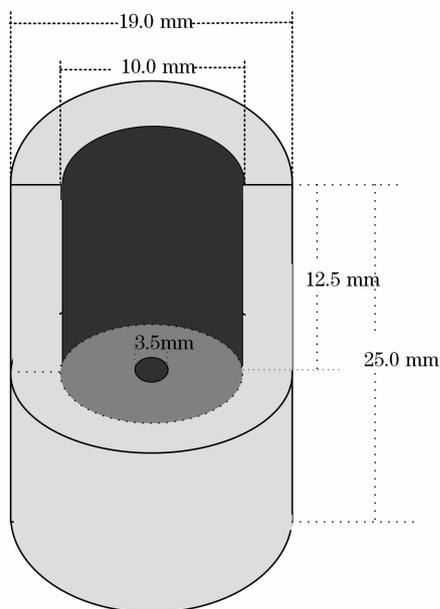


Fig. 2 Internal sketch of the sample cup

以降低杯壁对光的反射。测定时将种子平放入样品杯内的小凹槽内, 再将样品杯放到小升降台的固定位置上, 通过升降台将样品杯上升, 直至种子与探头紧密接触。由于光纤探头、样品杯的位置都相对固定, 在很大程度上避免了人为的操作误差, 提高了光谱采集的稳定性。

1.2.2 光谱采集

采用德国布鲁克光谱仪器公司生产的 MPA 傅里叶变换近红外光谱仪, 带有固体漫反射光纤探头。分辨率设为 4 cm^{-1} , 扫描范围 $4\ 000\sim 12\ 500\text{ cm}^{-1}$, 扫描次数 32 次。实验时将种子用镊子小心地放入特制的样品杯内, 在每粒种子正面的不同位置采集光谱 2 次, 在背面的不同位置采集 2 次, 共 4 次, 4 次重复的平均光谱作为每粒种子的近红外光谱。每粒种子采集完光谱后单独存放, 并标好序号。

1.2.3 种子硬实性的测定

将采集光谱后的种子单粒浸泡, 24 h 后吸胀的种子记为非硬实种子, 其余未吸胀的种子记为硬实种子。

1.2.4 数据分析与处理

采用中国农业大学开发的 CAUNIRS 近红外光谱分析软件, 采用定性偏最小二乘法建立乌拉尔甘草单粒种子硬实特性的定性分析模型。

2 结果与分析

2.1 样品硬实性的判断

通过浸泡鉴定, 800 粒种子中有 122 粒种子为非硬实种子, 我们又随机选择了 122 粒硬实种子共 244 粒种子进行分析, 种子分为校正集、检验集和预测集三组, 校正集 120 粒, 检验集 60 粒, 预测集 64 粒。校正集、检验集和预测集中硬实种子和非硬实的比例均为 1:1。

2.2 模型的建立

(1) 重复次数的比较

图 3 为乌拉尔甘草硬实和非硬实种子的近红外光漫反射光谱图, 乌拉尔甘草非硬实种子近红外光谱的吸收峰与硬实种子的完全一样, 仅通过吸收峰是无法区分硬实种子和非硬实种子的。我们分别利用乌拉尔甘草种子的一次光谱和 4 次光谱平均值进行建模, 光谱范围选择 $4\ 000\sim 8\ 000\text{ cm}^{-1}$, 结果表明: 4 次光谱平均值所建模型, 其校正集、检验集和预测集的鉴别率分别为 95.83%, 95.00%, 96.88%; 均显著高

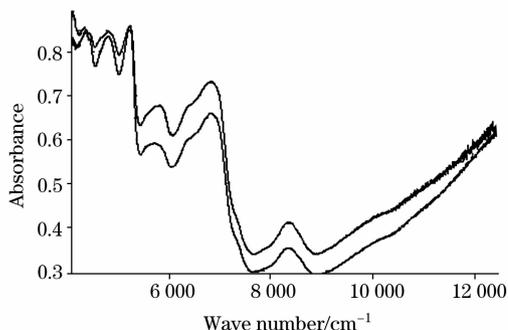


Fig. 3 Near infrared spectra of hard seed and soft seed of licorice

于单次重复所建模型的鉴别率(93.33%, 91.67%和82.81%)。

(2) 光谱范围的选择

从图 2 可以看出 10 000 cm^{-1} 以后的光谱噪声很大, 我们在光谱选择时尽量避免使用这一段的信息。根据对光谱的低频、合频、高频区的划分, 我们分别采用 4 000~5 000 cm^{-1} , 5 000~9 000 cm^{-1} , 8 000~10 000 cm^{-1} , 4 000~8 000 cm^{-1} , 4 000~9 000 cm^{-1} , 4 000~10 000 cm^{-1} 这几个谱区进行建模, 通过内部交叉和外部验证, 发现光谱范围为 4 000~8 000 cm^{-1} , 4 000~9 000 cm^{-1} , 4 000~10 000 cm^{-1} 时, 其鉴别率相差不大, 但是要显著高于其他谱区。8 000~10 000 cm^{-1} 为高频谱区, 噪声相对较大, 增加这一段谱区并不能增加鉴别率。我们最终选择 4 000~8 000 cm^{-1} 为适宜的建模光谱范围, 其校正集、检验集、预测集的鉴别率分别为 95.83%, 95.00%, 96.88%(见表 1), 与其他光谱范围比较, 鉴别效果较好。

由于不同的建模样品所建的模型其预测性能存在一定的差异, 为了进一步鉴定近红外光谱方法鉴别甘草种子硬实特

性的可行性及稳定性。我们分 10 次随机选择不同的建模样品, 进行建模和验证, 结果表明采用不同的样品进行建模, 虽然模型的鉴别率略有差异, 但均在 90% 以上(表 2), 预测准确率的平均值为 94.53%。说明利用近红外光谱法建立乌拉尔甘草种子硬实特性的无损鉴定模型较为稳定, 预测集中硬实种子和非硬实种子的平均鉴别率分别为 92.50% 和 96.56%。

Table 1 Influence of spectral range on NIR prediction results

光谱谱区 / cm^{-1}	主成分 分数	校正集 鉴别率/%	检验集 鉴别率/%	预测集 鉴别率/%
4 000~10 000	8	95.83	93.33	96.88
4 000~9 000	11	96.97	95.00	98.44
4 000~8 000	10	95.83	95.00	96.88
4 000~5 000	5	94.17	93.33	93.75
5 000~9 000	3	92.50	93.33	89.06
8 000~10 000	9	92.50	93.33	93.75

Table 2 NIR prediction results of models from different samples

随机选择 建模样品	主成分 分数	校正集 鉴别率/%	检验集 鉴别率/%	预测集 鉴别率/%	预测集硬实 种子鉴别率/%	预测集非硬实 种子鉴别率/%
第 1 次	10	95.83	95.00	96.88	93.75	100
第 2 次	11	97.50	91.67	95.31	96.88	93.75
第 3 次	8	94.17	98.23	96.88	93.75	100
第 4 次	11	94.17	96.67	95.31	93.75	96.88
第 5 次	8	97.04	91.11	95.31	93.75	96.88
第 6 次	5	94.81	97.78	90.63	84.38	96.88
第 7 次	5	94.07	95.56	92.19	87.50	96.88
第 8 次	10	96.30	95.56	96.88	100	93.75
第 9 次	11	95.56	97.78	93.75	90.63	96.88
第 10 次	10	95.83	100.00	92.19	90.63	93.75
平均	9	95.53	95.94	94.53	92.50	96.56

2.3 种子大小和种子颜色对于牛型识别准确率的影响

在一批甘草种子中, 还存在一定比例的大种子(3.0~3.5 mm)、小种子(2.0~2.5 mm)以及黑色的种子, 各约占 10% 左右。我们利用所建模型分别分析了棕绿色的大种子和小种子的硬实特性, 其预测准确率分别为 84.00% 和 82.00%, 预测准确率有所下降。而对于黑种子, 预测准确率下降极为显著, 只有 60.00%, 说明相对于种子大小而言, 种子颜色对于模型预测准确率的影响更大, 在利用该模型进行硬实特性分析时应予以注意。

3 结论与讨论

近年来, 关于近红外光谱技术在样品定性分析方面的研究报道很多, 但尚未见到用于豆科作物种子硬实判断的报道, 而该技术应用于种子硬实特性的无损判定具有很高的可能性。本文尝试利用近红外光谱分析技术对乌拉尔甘草种子的硬实特性进行了定性鉴别分析, 当采用 4 000~8 000 cm^{-1}

光谱范围建模, 鉴别率可达到 95% 以上。不同的建模样品对于模型的鉴别率具有一定的影响, 但是影响不显著, 其鉴别率均在 90% 以上。我们在建模过程中, 曾利用多种预处理方法, 其建模效果与当前模型效果相当, 这可能是因为种子的近红外光谱采用的是 4 次重复的平均光谱, 在很大程度上已经消除了部分噪声和基线漂移的影响。在比较乌拉尔甘草硬实种子与非硬实种子的结构和成分差异时, 需要事先了解种子的硬实特性, 预测集中硬实种子和非硬实种子的平均鉴别率分别为 92.50% 和 96.56%。我们利用该模型对不同大小和颜色的种子进行硬实性判定, 发现种子大小对该模型准确率的影响相对较小。种子颜色的影响相对较大, 该模型对于黑色种子硬实性的判断准确率只有 60%。

由于硬实的形成主要跟种皮的结构和化学成分有关, 因此利用近红外漫反射光谱可以较好地地区分硬实种子和非硬实种子。近红外光谱技术用于乌拉尔甘草种子硬实特性的定性分析具有简单快速、无需样品预处理的优点且具有较理想的鉴定结果, 适用于乌拉尔甘草种子硬实特性的快速鉴别分

析, 打破了甘草硬实种子无损识别的困局, 必将促进硬实机理的进一步研究。

致谢: 本论文得到中国农业大学理学院张录达教授的悉

心指点和布鲁克公司赵丽丽博士的热情帮助, 在此谨表谢意!

参 考 文 献

- [1] XU Ben-mei, SHI Xiao-hua(徐本美, 史晓华). Chinese Bulletin of Botany(植物学通报), 1997, 14(2): 45.
- [2] CAO Bang-hua, ZHAI Ming-pu, GUO Jun-jie(曹帮华, 翟明普, 郭俊杰). Scientia Silvae Sinicae(林业科学), 2005, 41(2): 42.
- [3] WANG Jin-long(王金龙). Soybean Science(大豆科学), 1999, 18(4): 351.
- [4] Shen-Miller J, Mudgett M B, Schopf J W, et al. American Journal of Botany, 1995, 82: 1367.
- [5] YANG Qi-he, YIN Xiao-juan, YE Wan-hui(杨期和, 尹小娟, 叶万辉). Chinese Bulletin of Botany(植物学通报), 2006, 23(1): 108.
- [6] Pasikatan M C, Dowell F E. Cereal Chemistry, 2004, 81(1): 145.
- [7] Lestander T A, Oden P C. Seed Science and Technology, 2002, 30: 383.
- [8] Soltani A, Lestander T, Tigabu M, et al. Journal of Near Infrared Spectroscopy, 2003, 11: 357.
- [9] Mulualem T, Joakin F, Per C O, et al. New Forests, 2007, 33: 155.
- [10] Tigabu M, Oden P C. New Forests, 2004, 27(1): 39.
- [11] Tigabu M, Oden P C. New Forests, 2003, 25(3): 163.
- [12] Wang D, Dowell F E, Lacey R E. Cereal Chemistry, 1999, 76(1): 30.

Identification of Hardness of Licorice Single Seed Using Near Infrared Spectroscopy

SUN Qun¹, LI Jun-hui², WANG Jian-hua¹, SUN Bao-qi^{1*}

1. Department of Plant Genetic and Breeding, College of Agriculture and Biotechnology, China Agricultural University/Key Laboratory of Crop Genomics and Genetic Improvement of Ministry of Agriculture/Beijing Key Laboratory of Crop Genetic Improvement, Beijing 100193, China
2. College of Information and Electrical Engineering, China Agricultural University, Beijing 100193, China

Abstract To break the dilemma on judging hard seeds and soft seeds of licorice and other legume families nondestructively, a distinguishing model for the hardness of licorice single seed was tried to be built by near infrared reflectance spectroscopy with distinguished partial least squares(DPLS). A total of 244 licorice seeds were divided into three groups: calibration set (120 samples), validation set (60 samples) and prediction set (64 samples), and each group has the same number of hard seeds and soft seeds. To eliminate the human error as far as possible, a specially made sample cup was designed for spectrum acquisition. Then the locations of the seed and the fiber-optic probe were fixed during each spectrum acquisition process. The influences of different replicate time, different spectral region and different calibration samples on the identification rate were compared. The result indicated that four replicates could increase the identification rate of the model significantly, the identification rates of the model of four replicates in calibration, validation and prediction set samples were 95.83%, 95.00% and 96.88% respectively, while that of one replicate were 93.33%, 91.67% and 82.81% respectively. The model of the spectral region between 4 000 and 80 000 cm^{-1} was better than that of other regions, and the identification rate in calibration, validation and prediction set samples were 95.53%, 95.94% and 94.53% respectively. Even with different samples, the predication rates were all more than 90%. The identification rates of hard seed and soft seed in prediction set samples were 92.50% and 96.56% respectively. The prediction for seeds with different size and different color showed that this model was not suitable for bigger and smaller seeds, especially not for black seeds. NIR offered a new way to distinguish the hardness of licorice single seed quickly, precisely and nondestructively, which will advance the study on the mechanism of hardness of crop seeds.

Keywords Near-infrared spectroscopy; Hardness; Single seed; Licorice

* Corresponding author

(Received Oct. 8, 2008; accepted Jan. 12, 2009)