

稻米色氨酸含量近红外反射光谱全局和局部飞速定标技术研究

张 斌, 张大鹏, 章文斌, 王克敏, 李 冠, 姚雅桢, 吴建国*, 石春海*

浙江大学农业与生物技术学院农学系, 浙江 杭州 310029

摘 要 色氨酸是人类一种必需氨基酸,也是稻米中一种重要的限制性氨基酸。从4年份1256份材料中择出272份有代表性的样品,采用碱水解-分光光度法测定了其色氨酸含量。比较不同定标方法的预测结果发现,运用改良的偏最小二乘法(modified partial least square, MPLS)的全局(Global)定标方法和局部(Local)飞速定标方法的预测效果较佳,基于精米粉光谱建立的方程的预测标准误差均为0.007%,外部验证决定系数分别为87.1%和87.4%,可用于定量分析;而基于糙米光谱建立的定标方程的预测效果略差,但仍具有良好的预测能力。研究结果表明,近红外光谱技术可作为水稻育种中间材料的快速筛选和食品工业中稻米原料的品质监控手段。

关键词 精米;色氨酸含量;近红外漫反射光谱;全局定标;局部飞速定标

中图分类号: S132 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2011)01-0073-04

引 言

水稻是世界上主要的粮食作物之一,也是我国最重要的粮食作物之一。氨基酸含量是衡量稻米营养品质的重要指标,尤其是必需氨基酸含量的高低会直接影响稻米的营养价值^[1]。色氨酸为人类的8种必需氨基酸之一,也是稻米中一种重要的限制性氨基酸^[2]。

近红外光谱(near infrared spectroscopy, NIRS)技术具有测定速度快,可同时分析多项指标,成本较低、无污染等特点,已广泛应用于稻米品质性状的快速分析,如直链淀粉含量^[3-5],糊化温度、凝胶稠度、碱消值^[6-8],矿质元素含量^[9],脂肪含量^[10, 11]和蛋白质含量^[6, 12]。Wu等^[13]建立了预测精米中17种氨基酸含量(不包括色氨酸)的校正方程,该方程具有较高的决定系数和较低的标准误差,可用于对除了胱氨酸、蛋氨酸、组氨酸之外的14种不同的氨基酸含量的测定。但至今尚未发现利用近红外光谱技术快速分析稻米色氨酸含量的报道。

本研究旨在开展稻米色氨酸含量的NIRS技术应用研究,运用全局和局部飞速定标技术建立近红外光谱校正方程快速测定精米色氨酸含量,提高种质资源以及水稻育种中间材料的选择效率;同时也可用于食品工业中对稻米原料的检测,使NIRS技术能够更好的为育种和食品生产实践服务。

1 材料与amp;方法

1.1 试验材料

研究材料包括两部分,第一部分1127份样品选自2005年和2006年种植于浙江大学试验农场的品种和高代品系;第二部分129份样品选自2007年和2008年种植于浙江大学试验农场的品质突变体群体材料,该群体遗传变异较大,蛋白质含量分布范围较宽,适合于氨基酸含量近红外定标。将这两部分合并起来共1256份样品作为定标原始群体。

将稻谷先用出糙机(型号:B-76,浙江黄岩)除去颖壳成糙米,再用精米机(型号:JNMJ3,浙江黄岩)磨成精米,最后用旋风磨(型号:3010-019, Fort Collins, Co., USA)磨成精米粉,过100目筛后装入小塑料袋自封袋中,敞口放置于干燥器中平衡水分3天,测定样品水分含量为9.82%。

1.2 近红外光谱扫描

运用近红外谷物分析仪(NIRSystems 5000, Silver Spring, MD, USA),将每个样品均以糙米和精米粉两种样品状态分别进行近红外光谱扫描,在波长为1100~2498 nm的范围内,每隔2 nm采集样品的反射强度。每份样品都进行2次装样,每次装样后旋转不同角度扫描2次,保存平均光谱待分析。

1.3 碱水解-分光光度法测定色氨酸含量

收稿日期:2010-05-10,修订日期:2010-08-20

基金项目:国家自然科学基金项目(30500319),浙江省科技厅科技攻关项目(2007C12902)和浙江省自然科学基金项目(Y3080217)资助

作者简介:张 斌,1984年生,浙江大学农学系硕士研究生 e-mail:zhangbinzju@sina.com

*通讯联系人 e-mail:jgwu@zju.edu.cn;chhshi@zju.edu.cn

采用碱水解-分光光度法(国标 GB/T 15400-94)测定精米粉样中色氨酸含量,所用仪器为可见-紫外分光光度计(Shimadzu UV-2550, Japan),标准样品购自 Sigma 公司。每个样品进行 3 次平衡试验,取平均值作为样品的化学参考值。

1.4 近红外校正方程研究方法

采用 WinISI II (v1.5, FOSS NIRSystems, Silver Springs, MD, USA)软件采集近红外光谱并建立近红外校正方程。光谱预处理主要有导数和平滑处理、标准正态变量变换和去趋势算法(standard normal variance with de-trending, SNV+D)等。运用改良的偏最小二乘法(MPLS)、偏最小二乘法(partial least square, PLS)和主成分回归(principal component regression, PCR)方法建立近红外光谱全局(Global)校正模型^[14]。通过交叉检验可判断所创建校正方程的效果,其中校正决定系数(coefficient of determination in calibration, RSQ)、校正标准误差(standard error of calibration, SEC)、交叉检验标准误差(standard error of cross-validation, SECV)和交叉检验决定系数(1 minus the ratio of unexplained variance to total variance, 1-VR)是衡量校正方程优劣的重要参数。局部(Local)飞速定标技术则运用 Win ISI II 软件,根据每个待测样品的光谱信息从定标数据库的样品中挑选出光谱相似的目标样品子集,进而建立定标模型来预测未知样品的氨基酸含量。外部验证能够更有效地检验方程的预测效果

和可靠性,分析中常以预测标准误差(SEP)、偏差(Bias)、外部验证决定系数(coefficient of determination in external validation, R^2)和 SD/SEP 值等参数来衡量方程的预测效果。

2 结果与分析

2.1 精米样品选择及色氨酸含量测定结果

运用 WinISI II 中 CENTER 程序计算原始群体光谱的马氏距离(GH),剔除了 11 个 $GH > 3.0$ 的异常样品,再根据邻居距离($NH < 1.1$)从剩余的 1 245 份样品中共选出 272 份代表性样品。对所选样品进行色氨酸化学测定,用于发展近红外校正方程。随机选出 40 个样品用于校正方程的外部验证,剩余 232 份样品用于建立近红外校正方程。表 1 结果表明,272 份校正样品集中的色氨酸含量变异范围较广,分布于 0.086%与 0.203%之间,外部验证样品集分布于 0.094%和 0.201%之间,两者分布范围相近。此外,校正集与验证集样品色氨酸含量的均值和标准差相近,平均值分别为 0.133%和 0.132%,标准差分别为 0.021%和 0.020%,验证集中各个氨基酸含量的分布范围都在校正集的范围,说明以上两个样品集在整个群体中具有很好的代表性,用此验证集来检验校正方程是有效的,适合于近红外光谱模型的构建和应用。

Table 1 Statistic results of tryptophan content in milled rice for calibration and external validation sets

Samples	Sample No	Minimum/%	Maximum/%	Mean/%	SD/%
Calibration set	232	0.086	0.203	0.133	0.021
External validation set	40	0.094	0.201	0.132	0.020

2.2 色氨酸含量化学分析值与光谱相关性分析

一阶导数(1, 4, 4, 1)和 SNV+D 处理精米粉近红外光谱后,对其特征波长点吸光度与色氨酸含量化学值进行了相关性分析(图 1)。结果表明,在 1 196, 1 670, 1 760, 1 962, 2 062, 2 128, 2 304 nm 等波长点上,光谱吸光度值与化学值之间的具有较高的相关性。表 2 中数据显示了产生这些特征吸收谱带所对应的色氨酸中的官能团结构及基团的振动方式。其中 1 670, 1 962, 2 062 和 2 128 nm 为氨基酸的特征吸收波长点,这些波长点的特征吸收对解释精米中色氨酸含量的变异具有较大的贡献。而精米粉在 2 062 和 1 962 nm 的吸收谱峰分别来自色氨酸中 CONH_2 和 CONH 结构中 N-H

和酰胺基团的吸收信息。同时,在 1 670 nm 处的吸收谱峰包含了色氨酸中 RCH=CHR 结构与苯环结构共同吸收的信息,为两者的重叠峰。此外,在 2 128 nm 处的谱峰为氨基酸结构的特征吸收谱峰,主要来自 N-H 和 C=O 基团的伸缩振动吸收信息。光谱相关性分析结果表明,相关性较高的波长点与色氨酸结构中各基团的特征吸收波长具有很高的一致性,而这些特征波长点的变异能很好地解释色氨酸含量的变异。因此,充分利用近红外光谱特征波长区的变异特点及所含信息,能有效地开展精米色氨酸含量的近红外定标工作。

Table 2 Functional group and modes of vibration in tryptophan at important wavelength plots with high correlation

Wavelengths	1 196	1 760	2 062	2 304	1 670	1 962	2 128
Correlation	0.621	0.784	0.718	0.682	-0.767	-0.652	-0.769
Mode of vibration	3×C—H Str.	2×C—H Str.	N—H Ben.	C—H Str+Ben.	2×C—H Str. N—H	Str. +amide	N—H Str. + C=O Str
Functional group	CH	CH ₂	CONH ₂	CH ₂	RCH=CHR Ar.	CONH	Amino acid

Str. 和 Ben. 分别为伸缩振动和变形振动; 2×和 3×分别为一级倍频和二级倍频; Ar. 为芳香族官能团

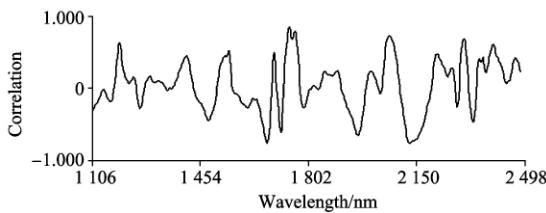


Fig 1 Correlation plot for tryptophan reference value vs wavelength absorbance by using mathematic treatment (1, 4, 4, 1) and SNV+D

2.3 精米色氨酸含量近红外定标方法比较研究

对精米粉光谱采用导数处理(1, 4, 4, 1)和 SNV+D 散射校正的光谱预处理方法, 分别采用 MPLS、PLS 和 PCR 三种回归分析方法进行近红外漫反射光谱全局(Global)定标研

究(表 3)。采用 MPLS 方法具有最佳的效果, 所建立的方程标准误差(SEC: 0.007%; SECV: 0.008%; SEP: 0.007%)均低于其他两种方法, 决定系数(RSQ: 88.4%; 1-VR: 85.0%; R^2 : 87.1%)则高于后两者, SD/SEP 值为 2.857, 也大于其他两种方法。而后两者中, 尽管预测偏差(Bias)、预测标准误差(SEP)以及 SD/SEP 值均一样, 但 PLS 发展的校正方程的 SEC, SECV 和 Slope 均小于 PCR 法, 同时前者的各项决定系数(RSQ、1-VR、 R^2)要大于后者, 说明 PLS 的方法略优于 PCR。

运用局部(Local)飞速定标技术预测精米色氨酸含量也取得了良好的预测结果, 预测标准误差为 0.007%, 决定系数达到了 87.4%, SD/SEP 值为 2.857。与以上三种回归方法的全局定标模型的预测效果相比, 运用局部飞速定标技术的预测结果与 MPLS 的方法相接近, 两者的预测标准误差较小, 外部验证决定系数较高(表 3)。

Table 3 Statistic results of the calibration for tryptophan content in milled rice with different methods based on the NIR spectroscopy of milled rice flour

Methods	Calibration				External validation				
	SEC	RSQ	SECV	1-VR	Bias	SEP	Slope	R^2	SD/SEP
MPLS	0.007	0.884	0.008	0.850	-0.001	0.007	1.009	0.871	2.857
PLS	0.007	0.881	0.008	0.846	-0.001	0.008	1.027	0.843	2.500
PCR	0.009	0.810	0.009	0.803	-0.001	0.008	1.059	0.825	2.500
Local-PLS	—	—	—	—	-0.001	0.007	1.040	0.874	2.857

对糙米光谱也采用导数处理(1, 4, 4, 1)和 SNV+D 散射校正的光谱预处理方法, 运用全局定标方法和局部飞速定标方法分别进行了色氨酸含量近红外定标研究(表 4)。分析结果表明: 其中运用 MPLS 的全局定标方法和局部飞速定标方法的定标效果较好, 预测标准误差均为 0.010%, 外部验证决定系数分别为 74.1% 和 74.7%, SD/SEP 比值均为

2.00, 具有较强的预测能力。而运用 PLS 的全局定标方法的效果次之, 运用 PCR 的全局定标效果最差。同基于精米粉光谱的定标相比, 糙米光谱的色氨酸含量定标略差些(表 3, 表 4), 其 MPLS 的全局和局部飞速定标的 SEP 后者略大, SD/SEP 后者略小。

Table 4 Statistic results of the calibration for tryptophan content in milled rice with different methods based on the NIR spectroscopy of brown rice grain

Methods	Calibration				External validation				
	SEC	RSQ	SECV	1-VR	Bias	SEP	Slope	R^2	SD/SEP
MPLS	0.010	0.762	0.011 6	0.684	-0.002	0.010	1.079	0.741	2.00
PLS	0.011	0.730	0.011 8	0.674	-0.001	0.012	1.008	0.657	1.67
PCR	0.013	0.590	0.013 5	0.563	-0.001	0.014	1.029	0.486	1.43
Local-PLS	—	—	—	—	-0.001	0.010	1.135	0.747	2.00

3 讨论

上述研究发现, 基于精米粉光谱, 运用 MPLS 的全局定标方法和局部飞速定标方法对色氨酸含量的定标效果均较为理想, 方程的标准误差小、决定系数较高。而基于糙米光谱的定标结果略差于精米粉光谱的定标结果, 但其 SD/SEP 达 2.0, 仍可用于定量分析。由于糙米经扫描分析后可用于播种繁殖, 对于少量育种材料的早世代筛选具有重要意义。此外, 在样品数目相对较小的情况下, 局部飞速定标方法也能

取得良好的结果。随着现代计算机性能的提高, 局部飞速定标技术将能克服分析速度慢、时间长的不足, 必将会在今后的研究中得到更多的应用。

本研究选用的样品包括了不同的年份和生长季节, 在遗传背景和农艺性状上具有较大的差异。定标群体的广泛变异性和化学分析结果的精确可靠使得定标结果良好, 方程具有良好的预测能力, 弥补了 Wu 等^[14]未发展色氨酸含量定标方程的不足, 对稻米营养品质的快速分析具有重要的应用价值。本研究所建立的定标方程可以直接用于稻米色氨酸含量的快速测定, 近红外光谱技术为水稻育种工作中氨基酸含量

的测定和食品工业中稻米品质的监控提供了有利的检测技术。

致谢: 浙江大学 985 农生环平台李源龙老师对本研究提供了帮助, 深表谢意。

References

- [1] Mitchell H H, Block R J. The Journal of Biological Chemistry, 1946, 163(3): 599.
- [2] Sotelo A, Hernandez M, Montalvo I, et al. Nutrition, 1994, 71: 605.
- [3] Bean M M, Miller R E, Juliano B O, et al. Cereal Food World, 1990, 35, 834.
- [4] Delwiche S R, Bean M M, Miller R E, et al. Cereal Chemistry, 1995, 72(2): 182.
- [5] WU J G, SHI C H. Field Crops Research, 2004, 87: 13.
- [6] Delwiche S R, Mckenzie K, Swebb B D. Cereal Chemistry, 1996, 73: 257.
- [7] BAO J S, CAI Y Z, Corke H. Journal of Food Science, 2001, 66: 936.
- [8] WU J G, SHI C H. Food Chemistry, 2007, 103: 1054.
- [9] Font R, Vélez D, Río-Celestino M D, Haro-Bailón A D, Montoro R. Microchim Acta, 2005, 151: 231.
- [10] LI W S, Shaw J T. Cereal Chemistry, 1997, 74 (5): 556.
- [11] TANG Shao-qing, SHI Chun-hai, JIAO Gui-ai, et al(唐绍清, 石春海, 焦桂爱, 等). Chinese Journal of Rice Science(中国水稻科学), 2004(6): 563.
- [12] XIE Xin-hua, XIAO Xin, LI Xiao-fang, et al(谢新华, 肖 昕, 李晓方, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2006, 37(8): 120.
- [13] Wu J G, Shi C H, Zhang X M. Field Crops Research, 2002, 75 (1): 1.
- [14] Barnes R J, Dhanoa M S, Lister S J. Applied Spectroscopy, 1989, 43: 772.

Study on the Global and Local Calibration Methods of Tryptophan Content in Rice by Near Infrared Reflectance Spectroscopy

ZHANG Bin, ZHANG Da-peng, ZHANG Wen-bin, WANG Ke-min, LI Guan, YAO Ya-zhen, WU Jian-guo*, SHI Chun-hai*

Department of Agronomy, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310029, China

Abstract Tryptophan is one of the essential amino acids in human, and also considered as an important limiting amino acid in rice. Alkali hydrolysis method was used to determine the tryptophan content of 272 representative samples selected from the 1 256 samples in four years. Based on the milled rice flour spectroscopy, by comparing the prediction abilities of the equations with global calibrations and local calibration, the global calibration method with MPLS and the local calibration method were better than others. Their calibration models could be introduced to application because of low standard errors of prediction for both methods (0.007%) and high coefficients of determination (87.1% and 87.4%, respectively). Furthermore, the prediction ability of the equation based on the brown rice spectroscopy could be still used for determination application though their effects were worse than that of the milled rice flour spectroscopy. It was concluded that the equation could be directly used to estimate the content of tryptophan in milled rice, and the technology of near infrared reflectance spectroscopy developed in this study could be applied in selecting breeding materials for higher tryptophan content and in quality control during rice processing.

Keywords Milled rice; Tryptophan content; Near infrared reflectance spectroscopy (NIRS); Global calibration; Local calibration

(Received May 10, 2010; accepted Aug. 20, 2010)

* Corresponding author