

基于近红外光谱技术的道地山药快速无损分析

杜敏¹ 吴志生¹ 巩颖² 史新元¹ 华国栋² 乔延江¹

(1 北京中医药大学中药学院,北京,100102; 2 北京中医药大学东方医院药学部,北京,100102)

摘要 目的:采用近红外漫反射光谱技术,建立对道地山药的快速无损鉴别方法。方法:收集三个产地(河南、河北、广东)的山药饮片共 90 份,采集其近红外漫反射光谱。采用判别分析法建立定性鉴别模型,并考察不同数据预处理方法对模型预测结果的影响。结果:道地山药饮片与其他产地的山药饮片在原始光谱与主成分得分的空间分布中均表现出较大的差异性。所建定性判别模型中,经多元散射校正结合二阶导数及 SG 平滑处理的模型表现出最佳预测结果,其对校正集与验证集的判正率均为 100%。结论:近红外漫反射光谱技术在道地药材的质量分析中具有可行性,可实现道地山药饮片的快速无损鉴别。

关键词 道地山药;近红外漫反射光谱;判别分析法;鉴别

Rapid Non-destructive Analysis of Genuine Chinese Yam by Near Infrared Diffuse Reflectance Spectroscopy

Du Min¹, Wu Zhisheng¹, Gong Ying², Shi Xinyuan¹, Hua Guodong², Qiao Yanjiang¹

(1 Beijing University of Chinese Medicine, Beijing 100102, China; 2 Dongfang Hospital, Beijing University of Chinese Medicine, Beijing 100078, China)

Abstract To investigate the feasibility of NIR spectroscopy in the discrimination of genuine Chinese Yam 90 samples from three origins were collected and their near infrared diffuse reflectance spectra were acquired. Discriminant analysis was employed to create qualitative model and several data pre-processing methods were explored. The results showed that genuine Chinese Yam was greatly different from other samples according to raw spectra and score plot of PCA. The discriminant analysis model with raw spectra after pretreatment of multiplicative scatter correction 2nd derivative and Savitzky-Golay filter smoothing showed the best predicative ability (the percentage of correctly identified varieties were both 100% for the calibration sample set and validation sample set). This research indicated that NIR diffuse reflection spectroscopy could be used for the rapid analysis of genuine Chinese Yam.

Key Words Genuine Chinese Yam; Near infrared diffuse reflectance spectroscopy; Discriminant analysis; Identification

doi: 10.3969/j.issn.1673-7202.2013.11.004

引言

山药又名薯蓣,为薯蓣科植物薯蓣 *Dioscorea opposita* Thunb. 的干燥根茎,是一种药食同源的常用中药。山药在河南、河北、山东、山西和南方的广西、福建、广东、台湾都有广泛种植^[1],其中以古怀庆府(今河南省焦作市境内)所产山药最为地道,被称为怀山药^[2]。由于生长环境、栽培技术和炮制方法的不同,道地与非道地山药之间在化学性状、物理性状和生物性状等方面都有较大差异^[3-4]。这导致道地药材在疗效、口感、产量、贮藏等方面所体现出来的综合特性优于种内其他非道地药材。对道地山药的现有鉴别方法有显微鉴别、传统理化鉴别,以及热分析法^[5]、分子生物学方法^[6]、红外光谱法^[7]等。近期有学者采用基于元素和原子层面的 X 射线荧光和粉末 X 射线衍射法^[3]对不同产地山药进行了分析。然而相比于上述鉴别方法,近红外漫反射光谱法不仅速度快,而且具有全息性特点^[8],可以提取药材的化学性质、物理性质,甚至生物

学属性,这为道地性现代诠释及道地药材鉴别提供了思路和方法。目前,已有学者将近红外漫反射光谱技术应用于山药多糖^[9]及指标成分尿囊素^[10]的定量分析中,然而尚未有报道采用近红外漫反射光谱技术对山药进行定性鉴别。因此,我们在本文探讨采用近红外漫反射光谱法对道地山药进行快速无损鉴别。

1 实验部分

1.1 样品来源 实验用山药饮片分别为河南怀山药、河北山药及广东山药,每个产地各 30 份样品。其中河南山药饮片按照道地栽培与加工炮制技术制备,河北及广东山药按照当地栽培及加工方法进行制备。所有饮片由北京中医药大学东方医院刘志禄副主任药师鉴别为薯蓣科植物薯蓣 *Dioscorea opposita* Thunb. 的干燥根茎切制而成。

1.2 山药饮片近红外漫反射光谱的采集 所用仪器为 Antaris I 傅里叶变换近红外漫反射光谱仪(美国 Thermo Nicolet 公司),配备 RESULT 3.0 光谱采集软

基金项目:国家自然科学基金项目(编号:81303218);省部级中药基础与新药研究重点实验室开放课题

第一作者:杜敏,研究方向:中药材质量的快速评价方法研究

通信作者:乔延江,教授,TEL: (010) 84738621, E-mail: yjqiao@263.net; 华国栋,副主任药师,TEL: (010) 67638124, E-mail: zhjhg@tom.com

件。光谱范围: 10001.03 ~ 3999.64 cm^{-1} , 共 1557 个变量。具体采集条件为: 采集模式为积分球漫反射; 累积扫描次数为 32 次; 分辨率为 8 cm^{-1} 。

由于山药饮片质地均匀、性状较规则, 因此未对山药饮片进行粉碎处理, 只采集饮片中间部位的近红外漫反射光谱, 并平行采集 3 次, 计算平均光谱用于分析。

1.3 数据处理与分析 由于不同山药质地间的差异与外界环境的影响, 近红外漫反射光谱会存在一定的噪音和基线漂移。因此为提高光谱质量, 需要对原始光谱进行数据预处理。本文主要考察了多元散射校正 (Multiplicative Signal Correction, MSC)、标准正则变换 (Standard Normal Variate, SNV)、一阶导数、二阶导数及 SG (Savitzky - Golay) 平滑。

为实现道地山药的鉴别, 本文采用 TQ Analyst 8.4 中的判别分析法 (Discriminant Analysis, DA) 建立定性判别模型。判别分析法是一种距离判别法, 校正过程首先对光谱数据进行主成分分析, 得到每个样品的主成分得分; 然后利用主成分得分建立基于马氏距离的判别准则。对未知样品进行预测时, 将未知样品与校正集中每个总体的马氏距离进行比较, 并将其判定为马氏距离最小的那个总体。本文在建立模型时从每个产地中随机选取三分之二, 共 60 个样本作为校正集, 剩余 30 个样本作为验证集。

数据预处理与建模采用 TQ Analyst 8.4 (美国 Thermo Nicolet 公司)。

2 结果与讨论

2.1 原始光谱 三个产地山药饮片的原始光谱图见图 1。由图 1 可看出, 河南山药饮片的光谱吸收值明显高于河北与广东山药, 而河北与广东山药的光谱严重重叠。这主要是因为不同产地山药由于质地的差异而光程有所不同, 且化学性质存在一定的差异。

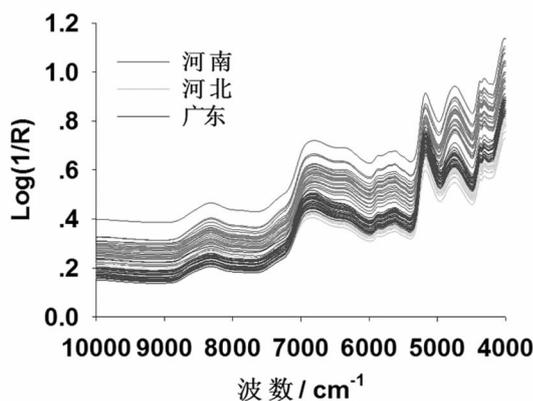


图 1 山药饮片原始光谱

2.2 主成分分析 首先对原始光谱进行主成分分析,

观察各类样本在主成分空间的分布情况。由图 2 可看出当主成分数为 3 时, 累积贡献率达到 99.48%, 即前三个主成分可表征原变量 99.48% 的信息。由前三个主成分得分图 (图 3) 可以看出, 只有怀山药样本很好地单独聚为一类, 而河北山药与广东山药有严重的交叉, 结果表明了道地山药的独特性。

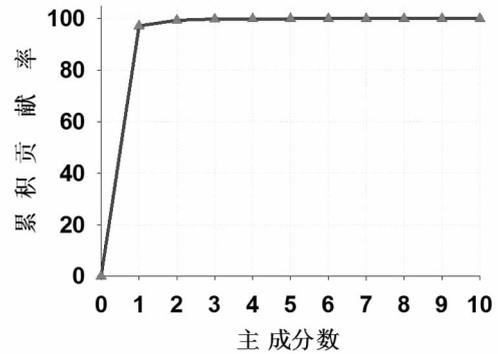


图 2 累计贡献率随主成分数变化图

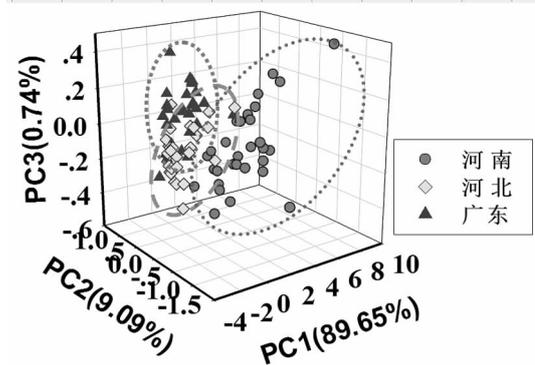


图 3 主成分分析得分图

2.3 判别分析结果 在校正过程, 首先考察了在不同预处理条件下的模型预测性能。以校正集和验证集的判正率为评价指标, 总判正率等于被正确判别的样品数除以样本集总数。山药饮片某一产地的判正率等于该组中被正确判别的样品数除以该组样本总数。不同预处理方式下模型的判正率见表 1。

表 1 不同预处理方式下模型的判正率

预处理方法	主成分数	校正结果 (%)			验证结果 (%)				
		总	河南	河北	广东	总	河南	河北	广东
无处理	3	80	100	80	60	76.7	100	80	50
多元散射校正	3	95	100	95	90	90	100	90	80
标准正则变换	3	95	100	95	90	90	100	90	80
多元散射校正 + 一阶导数 + SG(9 2) 平滑	3	96.7	100	100	90	96.7	100	100	80
多元散射校正 + 二阶导数 + SG(9 2) 平滑	3	100	100	100	100	100	100	100	100

注: SG (data points, polynomial order), 表中 SG 平滑参数为优化结果。

结果表明经数据预处理后, 模型的预测性能均有较大提高, 对校正集和验证集的判正率均大于 90%,

其中多元散射校正结合二阶导数及 SG 平滑的处理结果最好(见图 4。)其对校正集和验证集的判正率达到 100%。同时对比所建模型对不同产地山药的鉴别结果发现,所有模型对河南山药的判正率都达到 100%,而河北与广东山药一定程度上出现混淆。以上结果表明近红外漫反射光谱技术在道地山药的鉴别中具有较大的可行性。

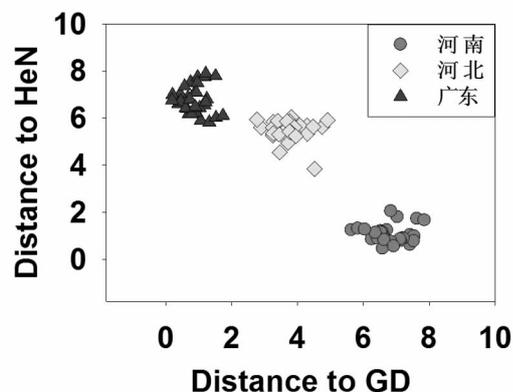


图 4 判别分析结果

注:横坐标为样本离广东组的马氏距离,纵坐标为样本距离河南组的马氏距离

3 结论

道地山药与非道地山药在化学、物理和生物学属性上的综合差异能一定程度上反映在近红外漫反射光谱上,通过主成分分析,这种差异更为明显。对原始数

据进行多元散射校正、二阶导数及 SG 平滑处理后,利用判别分析法建立定性模型,可很好的实现道地山药的鉴别,其判正率为 100%。上述结果表明利用近红外漫反射光谱技术在道地药材的质量分析中具有可行性,可用于道地山药饮片的快速无损鉴别。

参考文献

- [1]王彦博.我国山药的分布及其鉴别[J].吉林中医药,2009,29(12):1072-1073.
- [2]江苏新医学院.中药大辞典[M].上海:上海科学技术出版社,1999:166-168.
- [3]张景超,朱艳英,丁喜峰,等.不同产地山药的 XRF 和 PXRD 分析与表征[J].光谱学与光谱分析,2012,32(7):1972-1974.
- [4]张重义,谢彩侠,黄晓书,等.怀山药道地产区与非道地产区药材质量比较[J].现代中药研究与实践,2003,17(1):19-21.
- [5]宋爱新,张经纬,李明静,等.热分析方法对几种不同产地山药的鉴别[J].中草药,2003,34(2):169-171.
- [6]王东,白雁,陈志红.不同产地山药 rDNAITS 区序列的比较[J].时珍国医国药,2007,18(1):54-55.
- [7]徐永群,孙素琴,袁子民,等.红外光谱结合主成分分析鉴别道地山药[J].分析化学,2002,30(10):1231-1233.
- [8]郭兰萍,黄璐琦,Christian W Huck.近红外光谱技术及其在中药道地性研究中的应用[J].中国中药杂志,2009,34(14):1751-1757.
- [9]白雁,龚海燕,宋瑞丽.近红外漫反射光谱法快速测定山药药材中多糖的含量[J].中成药,2010,32(1):110-112.
- [10]白雁,龚海燕.近红外漫反射光谱法测定山药中尿囊素的含量[C].中药药效提高与中药饮片质量控制交流研讨会·北京,2009. (2013-10-30 收稿)

(上接第 1276 页)

- [27]LIU Jie-fang,GAO Pu-mei,HOU Zhen-yu et al. CWT-SVR model and its application in NIR analysis of corn[J]. Journal of Dongguan university of technology,2008,15(5):61-65.
- [28]侯振雨,汤长青,姚树文,等.离散小波变换-支持向量回归方法及其在谷物分析中的应用[J].河南农业科学,2006(8):40-42.
- [29]周洪成.基于奇异值的小波变换微弱信号消噪法[J].兵工自动化,2013(8):64-67.
- [30]丁永军,李民赞,郑立华,等.基于近红外光谱小波变换的温室番茄叶绿素含量预测[J].光谱学与光谱分析,2011,31(11):2936-2939.
- [31]Di Wu,Xiaojing Chen,Pinyan Shi,et al. Determination of α -linolenic acid and linoleic acid in edible oils using near-infrared spectroscopy improved by wavelet transform and uninformative variable elimination[J]. Analytica Chimica Acta,2009,634:166-171.
- [32]汤彦丰,侯占忠,王志宝,等.中草药大黄小波变换的近红外光谱的聚类分析[J].安徽农业科学,2012,30:14726-14727,14747.
- [33]Shao XG; Wang F; Chen D; et al. A method for near-infrared spectral calibration of complex plant samples with wavelet transform and elimination of uninformative variables[J]. Analytical and Bioanalytical Chemistry,2004,378(5):1382-1387.
- [34]Bart M. Nicolai,Karen I. Theron,Jeroen Lammertyn et al. Kernel PLS regression on wavelet transformed NIR spectra for prediction of sugar content of apple[J]. Chemometrics and Intelligent Laboratory System,2007,85(2):243-252.
- [35]田高友,褚小立,袁洪福,等.小波变换-分段直接校正法用于近红外光谱模型传递研究[J].分析化学,2006,34(7):927-932.
- [36]朱大洲,籍保平,史波林,等.基于小波变换的苹果汁多程近红外光谱信息提取研究[J].红外与毫米波学报,2009,28(5):371-375.
- [37]Zou Xiaobo,Zhao Jiewen,Malcolm J. W. Povey et al. Variables selection methods in near-infrared spectroscopy[J]. Analytica Chimica Acta,2010,667:14-32.
- [38]David Douglas Sousa Fernandes,Adriano A. Gomes,Gean Bezerra da Costa et al. Determination of biodiesel content in biodiesel/diesel blends using NIR and visible spectroscopy with variable selection[J]. Talanta,2011,87:30-34.
- [39]Roman M. Balabin,Sergey V. Smirnov. Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data[J]. Analytica Chimica Acta,2011,692:63-72.
- [40]张巧杰,张军.基于小波变换的大米直链淀粉波长选择方法[J].农业机械学报,2010,41(2):138-142.
- [41]Xiaojing Chen,Di Wu,Yong He,et al. An Integration of Modified Uninformative Variable Elimination and Wavelet Packet Transform for Variable Selection[J]. Spectroscopy,2011,26(4):42-47.
- [42]Ju Xiang,Wang Zhi Na,Xing Jun,Qu. NIR Model Transfer Based on Wavelet Transform Algorithms[J]. Spectroscopy,2013,28(6):36-41. (2013-11-18 收稿)