

基于近红外光谱技术天麻的产地区分*

李飞雁¹,徐曼菲²,乔延江^{2**}

(1. 遵义医药高等专科学校 遵义 563000; 2. 北京中医药大学中药学院 北京 100102)

摘要:天麻在《神农本草经》中列为上品,其主要分布在我国西南地区,不同产地质量差别较大,野生与栽培品价格差别尤为悬殊。用传统性状鉴别方法难以客观准确区分天麻的质量,故本文采用近红外光谱技术结合模式识别方法,建立能够快速准确区分不同产地和不同生长、栽培方式的天麻药材的方法。结果表明,利用近红外光谱技术结合多类分类算法,可以实现天麻产地和栽培方式的快速判别,其交叉验证判别准确率分别达到 94.3% 和 96.4%,且所建多类分类模型中光谱数据未经预处理。因此,本研究所建立的天麻产地和质量的快速判定方法可以推广为现场应用。

关键词:天麻 产地判别 野生 近红外光谱 多类分类算法

doi :10.11842/wst.2015.07.013 中图分类号 :R282.5 文献标识码 :A

天麻为兰科植物天麻 *Gastrodia elata* Bl. 的干燥块茎,《神农本草经》中列为上品。天麻主要分布于贵州、四川、陕西、湖北、湖南、云南等地。不同产地的天麻质量有较大的差异^[1],野生与栽培天麻质量差异更为显著,导致市场价格悬殊,市面上常有以次充好,以栽培冒充野生,甚至以假乱真的现象,严重影响了天麻的质量和疗效。

目前对天麻产地及野生栽培的区分多用性状鉴别,如杨秀泽等^[2]基于天麻的基本性状和环节纹特征对野生品和不同代数的栽培品进行了鉴别,指出天麻野生品与栽培品的主要性状区别点是栽培品的质地不够厚实,形体瘦长单薄;环节纹特征具有明显差异,并且随着代数增加差异也随之增大。邓恒青等^[3]应用数理统计法对天麻野生品、家种品 1 代和家种品 3 代的性状特征进行了统计分析,环节数、节上点数和节间距均显示同样的规律:野生品 < 家种品 1 代 < 家种品 3 代。然而用传统性状鉴别方法难以客观准确区分天麻的质量,因此本文拟探寻一种准确快速的鉴别方法,以区分不同产地和不同

同生长、栽培方式的天麻。

近红外光谱区是介于可见光区与中红外光区之间,波长范围为 780–2 526 nm。近红外光谱可以反应样品中含氢基团的信息,结合模式识别的方法,可实现快速无损的鉴别。目前近红外光谱技术在农产品中已经得到广泛应用^[4,5]。在中药领域,刘沐华等^[6,7]用近红外漫反射光谱法和多类支持向量机等模式识别技术成功鉴别了白芷和丹参产地;汤彦丰等^[8]用人工神经网络法和光栅型近红外光谱仪对中草药大黄进行了鉴别,结果表明人工神经网络法优于相关系数法,可对正品和非正品大黄进行成功识别,准确率达到 96.15%;张录达等^[9]以支持向量机(Support Vector Machine, SVM)法结合傅里叶变换近红外光谱建立了大黄样品真伪识别模型,取得了较满意的结果。本文拟采用近红外光谱技术结合一些模式识别方法对来自不同产地的天麻药材进行区分,同时探讨近红外光谱法对区分野生与栽培天麻的适用性。

1 材料与方法

1.1 材料

天麻样品采自贵州、四川、陕西、安徽 4 个省,共

收稿日期 2015-03-23

修回日期 2015-04-14

* 科学技术部“国家重大新药创新”科技重大专项(2010zx09502-002)清开灵注射液安全性关键技术研究,负责人 乔延江。

** 通讯作者:乔延江,本刊编委,教授,主要研究方向:中药信息学与中药质量控制。

140 个,经北京中医药大学张贵君教授鉴定为兰科植物天麻(*Gastrodiae Rhizoma*)的栽培品及野生品种(表 1)。

1.2 仪器设备

Antaris I 傅里叶变换近红外光谱仪(美国 Thermo Nicolet 公司),配备 RESULT3.0 光谱采集软件。分析软件采用 TQ Analyst 7.8 和 WEKA3.6.6 软件进行数据预处理及建模。

1.3 光谱采集

光谱采集前,将天麻药材切片,40℃干燥 8 h,即得光谱分析样本。光谱采集方式为漫反射,分辨率为 8 cm⁻¹,扫描次数 64 次,扫描范围为 10 000–3 500 cm⁻¹。为减小实验误差,每个样本重复测定 3 次,取平均光谱作为该样本的原始光谱。光谱采集时控制温度 25℃左右,湿度保持稳定。

1.4 方法

本实验应用 Weka3.6.6 软件,采用逻辑回归(Logistic)、简单逻辑回归(Simple Logistic)、多类分类算法(Multiclass Classifier, MC)、SVM 和 Logistic Model Tree(LMT)等方法建立天麻野生、栽培鉴别和产地鉴别定性判别模型,参数设置为默认值。模型的优劣用判别正确率进行评价。

2 结果与分析

2.1 不同产地天麻的原始光谱图

不同产地的栽培和野生天麻光谱如图 1 所示。横坐标为波数,范围是 4 000–10 000 cm⁻¹,纵坐标为光谱吸光度值。可以看出不同产地天麻的光谱曲线并无明显差别,难以直观鉴别。

2.2 野生和栽培天麻的鉴别结果

贵州产区天麻中共有栽培天麻 40 个,野生天麻 20 个。随机选取栽培天麻 20 个、野生天麻 15 个,共 35 个样本组成训练集。剩余 20 个栽培天麻、5 个野生天麻,共 25 个样本组成测试集。为了客观

评价建模方法的性能,采用了 5 折交叉验证来选择和评估模型,结果见表 2,其中 MC 分类法,交叉验证的判正率为 94.3%,校正判正率均为 100%,验证的判正率为 91.7%。

2.3 不同产地天麻的鉴别结果

采用贵州、安徽、陕西、四川不同产地的天麻共 140 个样本,其中随机选用 84 个样本作为训练集,剩余的 54 个样本作为测试集。同样采用 5 折交叉验证的方法,对所建模型进行验证。采用了 Logistic、SVM、LMT、MC 方法建模,计算结果如下表。可见 Logistic 和 MC 的方法对天麻的产地区分有较好的

表 1 实验样本

样品	类别	产地	样本数 / 个
天麻	栽培	贵州大方	20
天麻	栽培	贵州遵义	20
天麻	野生	贵州遵义	20
天麻	栽培	四川成都	20
天麻	栽培	四川南川	20
天麻	栽培	陕西	20
天麻	栽培	安徽	20

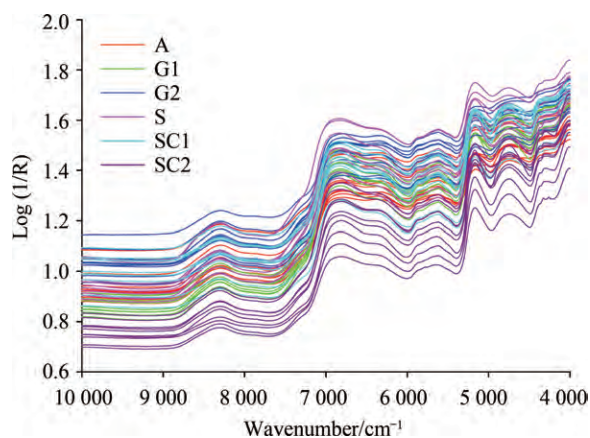


图 1 天麻原始近红外光谱图

注:A 为安徽产地,G1 为贵州大方产地,G2 为贵州遵义产地,S 为陕西产地,SC1 为四川成都产地,SC2 四川南川产地。

表 2 天麻野生栽培区分计算结果

方法	校正结果		验证结果		5fold CV	
	判错数	判正率/%	判错数	判正率/%	判错数	判正率/%
逻辑回归	0	100	3	87.5	3	94.3
简单逻辑回归	4	88.6	6	75.0	3	91.4
LMT	2	94.3	2	91.7	4	88.6
SVM	0	100	5	79.2	3	91.4
MC	0	100	2	91.7	2	94.3

表3 天麻产地区分计算结果

方法	校正结果		验证结果		5fold CV	
	判错数	判正率/%	判错数	判正率/%	判错数	判正率/%
逻辑回归	0	100	0	100	8	90.5
简单逻辑回归	0	100	12	75.9	18	78.6
LMT	1	99.0	12	75.9	18	78.6
SVM	16	83.2	13	72.2	18	78.6
MC	0	100	1	98.1	3	96.4

效果。其中 MC 的分类方法,交叉验证的判正率为 96.4%,校正判正率为 100%,验证的判正率为 98.1%。

3 结论

由于地理生态因子等的不同,不同产地的天麻,其活性成分天麻素含量差异很大,疗效的差异也比较显著。同时,现在野生天麻资源现已十分匮乏,野生、栽培品市场价格悬殊。本实验采集了不同产地的天麻野生和栽培品的近红外光谱,未对原始光谱进行预处理,直接采用 Logistic、SVM、LMT 等方法建模对天麻进行鉴别,比较了几种分类方法。不同的判别模型分类结果相差较大,因此分类方法的

选择很重要,恰当的建模方法可以准确的鉴别中药材产地,以及对野生和栽培的鉴别。其中 MC 分类方法无论是在天麻的产地区分及野生和栽培的鉴别中,正确率都在 90% 以上,能获得很好的分类效果,说明利用近红外光谱技术结合适当的模型可以实现对天麻快速无损的产地鉴别及野生和栽培的区分。本研究所采用的近红外光谱技术,结合化学计量学的方法,可实现对天麻进行快速无损的产地鉴别与野生和栽培的区分,近红外光谱技术可以作为一种有效鉴别中药材产地及真伪的新方法、新技术用于中药材的鉴别,具有较高的经济价值和临床应用价值。

参考文献

- 梁纪军,田大丰,王中彦,等.HPLC 法测定不同产地天麻中天麻素的含量.沈阳药科大学学报,2006,23(1):26-28.
- 杨秀泽,周汉华,童红.天麻野生品与栽培品的鉴别比较.中国药房,2011,22(19):1804-1805.
- 邓恒青,周汉华,童红,等.天麻环节纹的观察比较及数理统计分析.中国民族民间医药,2009,18(11):22-23.
- 徐惠荣,应义斌.近红外图像和光谱分析技术在农产品品质无损检测中的应用与展望.浙江大学学报(农业与生命科学版),2002,28(4):109-113.
- 籍保平.近红外光谱技术在农产品加工中的应用.粮油加工与食品机械,2000(6):31-33.
- 刘沐华,张学工,孙素琴.中药材产地的近红外光谱自动鉴别和特征谱段选择.科学通报,2005,50(4):393-398.
- 刘沐华,张学工,周群,等.近红外漫反射光谱法和模式识别技术鉴别中药材产地.光谱学与光谱分析,2006,26(4):629-632.
- 汤彦丰,张卓勇,范国强.中草药大黄的近红外光谱和神经网络鉴别研究.光谱学与光谱分析,2004,24(11):1348-1351.
- 张录达,苏时光,王来生,等.支持向量机(SVM)在傅里叶变换近红外光谱分析中的应用研究.光谱学与光谱分析,2005,25(1):33-35.

Determination of Geographical Location of *Gastrodia Elata* Using NIR

Li Feiyan¹, Xu Manfei², Qiao Yanjiang²

(1. Zunyi Medical and Pharmaceutical College, Zunyi 563000, China;

2. School of Pharmacy, Beijing University of Chinese Medicine, Beijing 100102, China)

Abstract: *Gastrodia elata* is graded as top medication in the *Shen Nong's* Herbal Classic. It was mainly distributed in

{ World Science and Technology/Modernization of Traditional Chinese Medicine and Materia Medica } 1407

southwest China. Its quality varied with geographical location. And the quality difference between wild and cultivated sample was extreme. Identifications using traditional methods were unable to accurately distinguish the quality of *G. elata*. Therefore, near-infrared (NIR) spectroscopy combined with pattern recognition method was used to distinguish the quality of *G. elata* from different geographical locations as well as cultivated or wild. The results demonstrated that using NIR spectroscopy combined with multiclass classification algorithm, the geographical location of *G. elata* can be accurately distinguished. The prediction accuracy can reach as high as 94.3% and 96.4% for both applications. Besides, the classification model was built without preprocessing; hence, it can be extended to be applied on-site.

Keywords: *Gastrodia elata*, geographical location, wild, near-infrared spectroscopy, multiclass classifier

(责任编辑:李沙沙 张志华,责任译审:王 晶)