

顾玉琦,刘瑞婷,寿国忠,等.应用近红外光谱技术快速鉴别铁皮石斛的产地[J].江苏农业科学,2016,44(5):365-368.

doi:10.15889/j.issn.1002-1302.2016.05.107

# 应用近红外光谱技术快速鉴别铁皮石斛的产地

顾玉琦<sup>1,2</sup>, 刘瑞婷<sup>2</sup>, 寿国忠<sup>1</sup>, 张雯雅<sup>2</sup>, 赵旭<sup>2</sup>

(1. 浙江农林大学浙江省林业智能监测与信息技术研究重点实验室, 浙江临安 311300;

2. 浙江农林大学工程学院, 浙江临安 311300)

**摘要:**采用近红外漫反射光谱技术结合簇类独立软模式方法(SIMCA)建立了预测模型,并对 5 个产地的铁皮石斛进行了快速鉴别研究。光谱数据经过 Savitzky-Golay(S-G)平滑、1 阶 Savitzky-Golay(S-G)求导和均值中心化组合光谱预处理和波段选择优化后,所建的 SIMCA 模型为最佳模型。所建模型的识别率和拒绝率均为 100%,而且 5 个产地铁皮石斛的模型也能拒绝其他 3 种石斛金钗石斛、紫皮石斛和金石斛。研究表明,模型合理,能够成功实现铁皮石斛产地快速鉴别,为不同产地的铁皮石斛的快速鉴别提供了一种新的方法。

**关键词:**铁皮石斛;近红外漫反射;产地鉴别

**中图分类号:** O657.33

**文献标志码:** A

**文章编号:** 1002-1302(2016)05-0365-04

铁皮石斛(*Dendrobium officinale* Kimura et Migo)是兰科石斛属多年生附生草本植物,是传统名贵珍稀中药材,具有益胃生津、滋阴清热等功效<sup>[1]</sup>。铁皮石斛的主要化学成分包含多糖、氨基酸、生物碱、矿物元素、菲类、联苯类和菊酮化合物等<sup>[2]</sup>。

不同产地的铁皮石斛价格和药用价值差异很大。2013 年 12 月,位于浙江省临安市的天目山铁皮石斛和位于云南省的广南铁皮石斛已经通过国家质检总局批准实施地理标志产品保护<sup>[3]</sup>;2014 年 5 月,浙江省乐清市和广西省容县均被授予“中国铁皮石斛之乡”的称号。目前,在市场上有许多不法商家和企业,用非地理标志产品和非“中国铁皮石斛之乡”的铁皮石斛以次充好,以谋取高额利益。面对如此混乱的市场,非常需要研究 1 种快速鉴定铁皮石斛产地的技术。

在以前不同产地铁皮石斛的研究中,多以理化试验、紫外分光光度法、高效液相色谱等技术进行试验研究<sup>[4-7]</sup>,这些方法技术耗时、耗力、成本高、须要专业技术人员操作。与以上技术相比,近红外光谱分析技术具有分析速度快、成本低、光谱测量方便、样品前处理简单,且能够最大限度地保留不同产地样本间的微小差异等优点<sup>[8]</sup>,目前该技术已被用于麦冬、当归、人参、白芷、黄芪等药材的产地鉴别,效果较好<sup>[9-13]</sup>。近年来,国内外相关研究利用近红外光谱技术结合化学计量学对不同种类的石斛进行快速鉴别,均获得较好的结果<sup>[14-17]</sup>,但是该方法用于同一品种石斛不同产地的鉴别研究尚未见报道。主要原因是石斛产地之间的差异主要是受到地理位置、气候条件、栽培技术等多因素影响,这种差异较石斛

品种间差异小,因此产地鉴别比品种鉴别困难。本研究利用近红外漫反射光谱法结合簇类独立软模式(soft independent modeling of class analogy, SIMCA)法对 5 个产地的铁皮石斛进行鉴别分析,旨在为石斛产地的快速鉴别提供新的方法。

## 1 材料与方法

### 1.1 材料

试验所用的材料为来自安徽霍山、浙江乐清、浙江天目山、广西容县、云南龙陵的铁皮枫斗,分别收集 20、30、30、30、30 个批次。所有样品来源于中国石斛博物馆。样品用中药粉碎机粉碎,分别过 40、60、80 目筛,取约 2 g 样品在上述条件下进行测量。随机把 140 份石斛样本划分为 84 份校正集和 56 份验证集,其中安徽霍山、浙江乐清、浙江天目山、广西容县、云南龙陵分别为 12、18、18、18、18 份作为校正集样品,剩余的样品作为验证集。

使用仪器:微型近红外光谱仪 SmartEye1700(尼迺光电科技有限公司),DFT-100 大德不锈钢中药粉碎机(温岭市林大机械有限公司),凯丰实验室电子天平称重精密电子秤(KF-0040)(中国凯丰基团有限公司),不锈钢筛盘(顶历医疗器械有限公司)。

### 1.2 光谱采集

试验所用微型近红外光谱仪,光谱扫描范围为 1 000 ~ 1 650 nm,分辨率为 8 cm<sup>-1</sup>,扫描次数为 50 次,由于样品为不透明的粉末状颗粒,所以试验采用积分球的漫反射采样方式。为了保证仪器能量的稳定,每次扫描光谱前,仪器需预热 30 min。在试验过程中,湿度基本保持不变,室内温度保持在 25 ℃左右。将样品置于采样窗口区域,以仪器内置背景为参比,100% Spectralon™标准白板为参考。每个样本在不同时间、不同位置分别采集 3 次,取其平均光谱作为该样品的原始光谱。

### 1.3 光谱数据预处理

在建立 SIMCA 模型之前,为消除基线漂移、高频随机噪声和光散射等因素对样品光谱的影响,需要对光谱进行预处

收稿日期:2015-04-29

基金项目:浙江省林业智能监测与信息技术研究重点实验室开放基金(编号:100151403);浙江农林大学人才启动项目(编号:2013FR013、2012FR070)。

作者简介:顾玉琦,博士,讲师,主要从事农林业监测研究。E-mail: guyuqi@zju.edu.cn。

通信作者:寿国忠,教授,硕士生导师,主要从事农业信息化研究。

E-mail: Shou8168@gmail.com。

理以获得有用信息。样本粒径的大小和均匀度都会对光的漫反射有一定影响;同时样本的密实度也会影响光的传播,因此需要对样本的原始光谱数据进行预处理。为消除光谱数据在采集时首端与末端产生的部分噪音,截取 1 100 ~ 1 630 nm 波段的光谱数据进行分析。本试验分别采用无处理、多元散射校正 (multiplicative signal correction, MSC)、标准正态变量 (standard normal variate, SNV)、S - G 平滑 + 1 阶 S - G 求导、S - G 平滑 + 1 阶 S - G 导数加均值中心化等组合预处理方法对样本的原始光谱进行预处理,最后选出最佳的光谱预处理方法为 S - G 平滑 + 1 阶 S - G 导数加均值中心化。

1.4 数据分析方法

主成分分析 (principal component analysis, PCA) 是把多个指标转换为几个综合指标的 1 种统计方法,沿着协方差最大的方向由高维数据空间向低维数据空间投影,所得的主成分向量相互正交,它可以将样本在高维空间的分布通过低维空间来展现<sup>[18]</sup>。

近红外光谱吸收弱,谱带较宽且重叠严重,因此采用化学计量学中的多变量数据分析方法进行分析。本研究使用 SIMCA 法结合近红外光谱技术对铁皮石斛的 5 个产地进行鉴别。

SIMCA 法在光谱色谱的定性分析中得到了广泛的应用。在本研究中,首先针对某一产地样本的光谱数据矩阵进行主成分分析,建立这个产地的主成分回归模型,然后依据该模型对未知样品进行识别,即分别试探将该未知样本与各样本的模型进行拟合,以确定未知样本的产地<sup>[19]</sup>。模型质量的评价使用识别率、拒绝率等指标。

识别率 =  $\frac{\text{识别自身类样本个数}}{\text{该类样本的总个数}} \times 100\%$ ; (1)

拒绝率 =  $\frac{\text{拒绝其它类样本个数}}{\text{其它样本的总个数}} \times 100\%$ 。 (2)

原始试验数据通过 Imethod 2000 软件 (仪器自带软件) 获取,再通过 Matlab (R2008a) 程序进行数据处理与建模。

2 结果与分析

2.1 样品粒度对近红外光谱的影响

以浙江乐清产地的样品为例,研究样品粒度对近红外光谱的影响。取浙江乐清产地的铁皮石斛为样品,采用中药粉碎机粉碎样品,将粉碎的样品通过不同目数的筛网,形成一系列粒径大小不同的样品:样品粒径 40、60、80 目。研究发现,近红外吸收强度与样品粒径大小有关 (图 1)。随着样品粒径的增加,样品反射光谱  $[\lg(1/R)]$  越大,吸收强度增大,近红外吸收强度与样品粒径大小正相关;样品越细,光谱稳定性越好<sup>[20]</sup>。

2.2 不同样品粒度的 SIMCA 模型的建立和评价

以浙江乐清产地的样品为例,研究不同样品粒度 SIMCA 模型的建立和评价。对 40、60、80 目 3 种不同样品粒度的近红外光谱,使用校正集样品建立 SIMCA 模型,用验证集检验模型的可靠性结,结果发现 80 目的光谱所建模型的效果优于 40、60 目;大于 60 目目筛的样品均匀性和光谱的重现性好 (表 1)。因此,80 目的光谱建立的模型效果最好,本试验采用 80 目筛的铁皮石斛样品进行测定。

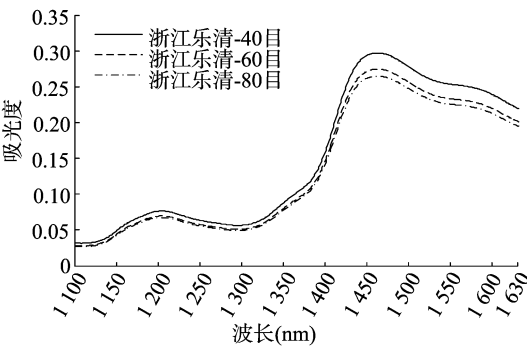


图1 不同粒径的近红外光谱

表 1 不同样品粒度下建立 SIMCA 模型的校正和验证结果

筛网 目数	主成 分数	校正集		验证集	
		识别率 (%)	拒绝率 (%)	识别率 (%)	拒绝率 (%)
40	1	96	88	94	100
60	3	96	100	95	100
80	2	100	100	100	100

2.3 近红外漫反射光谱

为消除光谱数据在采集时首端与末端产生的部分噪音,记录了 1 100 ~ 1 630 nm 波段的光谱数据进行分析。近红外光谱反映的是 C - H、O - H 等含氢基团的倍频和合频吸收,铁皮石斛中含有多糖、生物碱、氨基酸等大量化学成分,这些成分的含氢基团都能在近红外区域产生倍频与合频吸收 (图 2)。此外,所有样品在 1 140 ~ 1 240 nm 和 1 360 ~ 1 500 nm 附近有明显的吸收峰,这 2 个区域分别为 C - H 第 3 组合频和 C - H 第 2 组合频区域 (图 2)<sup>[21]</sup>。不同产地铁皮石斛的光谱曲线趋势相似,在部分区域有明显区别,具有一定的特征性和指纹性,这一差异为铁皮石斛不同产地的鉴别奠定了数据基础。

在 1 142、1 410、1 480 nm 处有显著的吸收峰 (图 3),其中 1 142 nm 光谱特征吸收峰是官能团芳烃 (ArCH) 的 C - H 伸缩振动基频的二级倍频谱带;1 410 nm 光谱特征吸收峰是官能团亚甲基的 C - H 亚甲基的 C - H 组合频 (伸缩振动基频的一级倍频谱带和弯曲振动基频吸收谱带);1 480 nm 处光谱特征吸收峰是官能团聚酰胺 11 非键合 N - H 伸缩振动基频的一级倍频谱<sup>[21]</sup>。

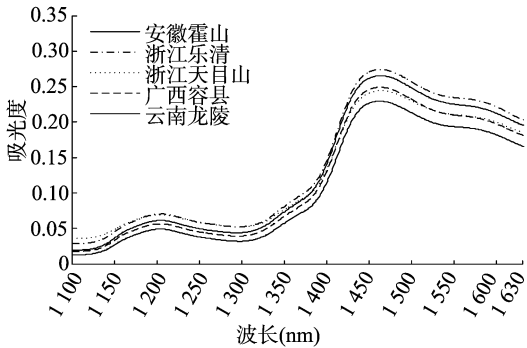


图2 不同产地样品的平均光谱

2.4 PCA 分析

本研究通过比较不同光谱预处理方法、波长范围和主因子数对建模精度的影响,对 5 个产地的铁皮石斛建立不同的模型。5 个产地铁皮石斛的最优预处理方法为 S - G 平滑、1

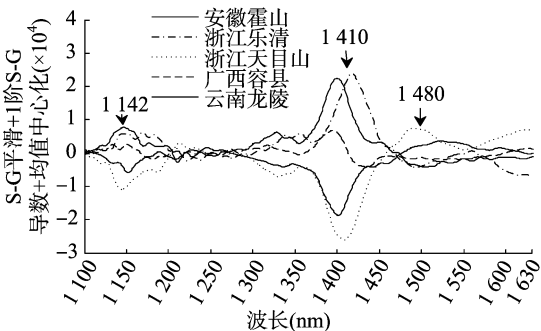


图3 每个产地的预处理后的平均光谱

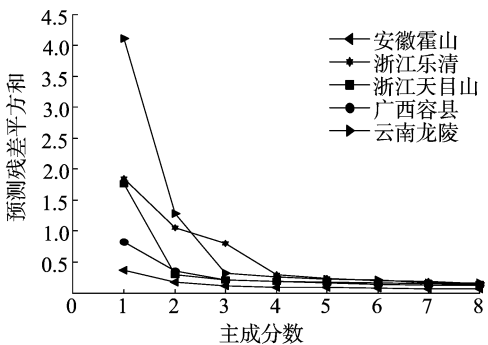


图5 PRESS 值与模型主成分数之间的关系

阶 S-G 导数和均值中心化,5 个产地铁皮石斛的最优波长范围为 1 100 ~ 1 630 nm。用主成分分析法对校正集样本进行分析,基于样本的主成分得分图能够更直观地揭示样本的内部特征和聚类信息。 $X$  轴表示样本的第一主成分得分 (PC1), $Y$  轴表示样本第二主成分得分 (PC2), $Z$  轴表示样本第二主成分得分 (PC3),5 个产地的铁皮石斛样本在该示意图中有明显的分类趋势 (图 4),再次说明利用近红外光谱判别铁皮石斛的产地是可行的。

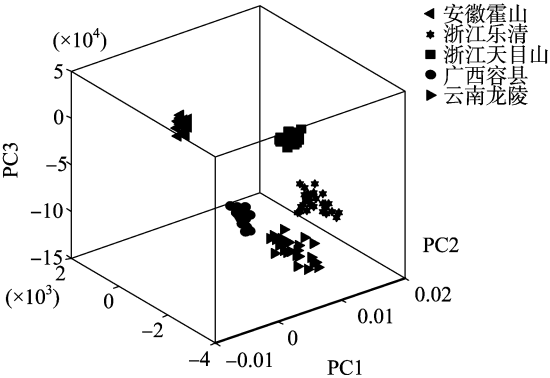


图4 所有样本的前 3 个主成分示意图

2.5 5 个产地铁皮石斛的 SIMCA 分类模型

通过交互验证 (cross-validation) 来确定上述铁皮石斛的 5 个产地模型的最佳主成分数,即在预测残差平方和 (prediction residual error sum of square, PRESS) 变化不大的情况下选取比较小的主成分数。在不同的主成分数下,5 个产地的模型的预测残差平方和见图 5。因此,不同模型主因子数的选取结果为:安徽霍山产地取主因子数为 4,提取 4 个主成分的贡献率已达全光谱贡献率的 97.89%;浙江乐清模型取主因子数为 2,提取 2 个主成分的贡献率已达全光谱贡献率的 97.92%;浙江天目山、广西容县、云南龙陵的模型取主因子数都为 3,提取 3 个主成分的贡献率分别已达全光谱贡献率的 97.38%、97.09%、99.29%。

在极显著性水平条件下来检验模型的可靠性,获得了所建立的预测模型性能最佳校正和验证的结果 (表 2)。研究发现,在对校正集样品进行预测的时候,5 个产地的模型对于本产地样品的识别率均为 100%;在拒绝非本产地样品时,5 个产地模型的拒绝率均为 100%。对验证集样品进行预测时,5 个产地的模型对样品的识别率均为 100%,拒绝率均为 100%。结果表明,可以用该 SIMCA 模型鉴别不同产地来源的铁皮石斛,所建立的产地识别模型能够达到正确识别 5 个

表 2 预测模型的校正和验证结果					
产地	主成分数	校正集 (%)		验证集 (%)	
		识别率	拒绝率	识别率	拒绝率
安徽霍山	4	100	100	100	100
浙江乐清	2	100	100	100	100
浙江天目山	3	100	100	100	100
广西容县	3	100	100	100	100
云南龙陵	3	100	100	100	100

产地铁皮石斛样品的目的。

2.6 对其他样品的鉴别

为进一步确认近红外判别分析模型的可靠性,利用建好的 5 个产地铁皮石斛的模型对其他 3 种石斛金钗石斛、紫皮石斛、金石斛的样品进行判别分析,发现所建模型对其他样品拒绝率都为 100% (表 3)。

表 3 SIMCA 模型对其他样品鉴别的预测结果					
模型	被测样品	产地	样本光谱总数 (个)	错误数 (个)	拒绝率 (%)
安徽霍山	金钗石斛	贵州赤水	30	0	100
	紫皮石斛	浙江乐清	30	0	100
	金石斛	云南龙陵	30	0	100
浙江乐清	金钗石斛	贵州赤水	30	0	100
	紫皮石斛	浙江乐清	30	0	100
	金石斛	云南龙陵	30	0	100
浙江天目山	金钗石斛	贵州赤水	30	0	100
	紫皮石斛	浙江乐清	30	0	100
	金石斛	云南龙陵	30	0	100
广西容县	金钗石斛	贵州赤水	30	0	100
	紫皮石斛	浙江乐清	30	0	100
	金石斛	云南龙陵	30	0	100
云南龙陵	金钗石斛	贵州赤水	30	0	100
	紫皮石斛	浙江乐清	30	0	100
	金石斛	云南龙陵	30	0	100

3 结论

本研究研究了不同样品粒径对近红外光谱的影响,建立了不同样品粒度的 SIMCA 模型,并对该模型进行了评价,发现随着样品粒径的增加,样品反射光谱和吸收强度逐渐增大,因此近红外吸收强度与样品粒径大小正相关。此外,80 目的光谱建立的模型效果优于 40、60 目光谱所建模型。

本研究利用近红外光谱技术结合 SIMCA 模式识别方法对 5 个产地的铁皮石斛建立了鉴别模型。经过对比不同光谱

预处理和光谱波段对模型的影响,最后选择 S-G 平滑 + 1 阶 S-G 导数 + 均值中心化组合为最佳预处理方法和 1 100 ~ 1 630 nm 为建模光谱波段,所建模型为最优的模型。模型能正确识别 5 个产地铁皮石斛,对其他产地识别时,拒绝率达到 100%,而且 5 个产地铁皮石斛的模型也能拒绝其他 3 种石斛金钗石斛、紫皮石斛和金石斛。结果表明,近红外光谱结合 SIMCA 模式识别方法在铁皮石斛产地溯源中具有可行性,为解决铁皮石斛的质量安全问题<sup>[22-24]</sup>提供了一种新的思路。与理化试验、紫外分光光度法和高效液相色谱等技术方法相比,本研究所提出的铁皮石斛产地识别方法具有识别率高、识别方法可扩充性好、适用范围广等优点。

#### 参考文献:

- [1] 戴燕萍. 铁皮石斛生产质量管理规范研究[D]. 杭州:浙江农林大学,2012.
- [2] 华允芬. 铁皮石斛多糖成分研究[D]. 杭州:浙江大学,2005.
- [3] 刘文杰. 铁皮石斛的红外光谱定性定量研究[D]. 北京:北京中医药大学,2014.
- [4] 徐程,詹忠根,廖苏梅. 8 种不同地域铁皮石斛农艺性状及多糖和纤维素分析[J]. 浙江大学学报:理学版,2008,35(5):576-579.
- [5] 龚庆芳,周浩,王新桂,等. 不同产地铁皮石斛的品质比较研究[J]. 北方园艺,2013(8):162-165.
- [6] 戚辉,陈健,易燕群,等. 不同产地铁皮石斛形态及有效部位成分含量比较[J]. 广州中医药大学学报,2013,30(4):558-561.
- [7] 刘文杰,孙志蓉,杜远,等. 不同产地铁皮石斛主要化学成分及指纹图谱研究[J]. 北京中医药大学学报,2013,36(2):117-120.
- [8] 赵海燕,郭波莉,魏益民,等. 近红外光谱对小麦产地来源的判别分析[J]. 中国农业科学,2011,44(7):1451-1456.
- [9] 余国梅,王革,张延莹,等. 近红外光谱技术鉴别麦冬药材产地[J]. 解放军药科学学报,2012,28(5):451-453.
- [10] 李波霞,魏玉辉,席莉莉,等. 近红外光谱和化学计量学对不同产地不同产期当归的定性研究[J]. 光谱实验室,2011,28(4):2128-2134.
- [11] Woo Y, Cho C, Kim H, et al. Classification of cultivation area of ginseng by near infrared spectroscopy and ICP - AES [J]. Microchemical Journal, 2002, 73(3):299-306.
- [12] Liu S, Zhang X, Sun S. Discrimination and feature selection of geographic origins of traditional Chinese medicine herbs with NIR spectroscopy [J]. Chinese Science Bulletin, 2005, 50(2):179-184.
- [13] 史春香, 杨悦武, 郭治昕, 等. 近红外技术鉴别黄芪产地[J]. 天津药学, 2006, 18(1):19-21.
- [14] Wang C, Xiang B, Zhang W. Application of two - dimensional near - infrared ( 2D - NIR ) correlation spectroscopy to the discrimination of three species of *Dendrobium* [J]. Journal of Chemometrics, 2009, 23(9):463-470.
- [15] 汤青. 近红外光谱分析技术快速鉴别霍山石斛的建模研究[D]. 山东大学, 2013.
- [16] 丁长春, 方向京, 赵艳丽, 等. 近红外漫反射光谱法快速鉴别石斛属植物[J]. 光谱学与光谱分析, 2014, 34(1):82-86.
- [17] Hu T, Tang Q, Zhang H, et al. Absolutely nondestructive discrimination of *Huoshan Dendrobium nobile* species with miniature near - infrared ( NIR ) spectrometer engine [J]. Spectroscopy and Spectral Analysis, 2014, 34(10):2808-2814.
- [18] 赵杰文, 蒋培, 陈全胜. 雪莲花产地鉴别的近红外光谱分析方法[J]. 农业机械学报, 2010, 41(8):111-114.
- [19] 陆婉珍. 现代近红外光谱分析技术[M]. 2 版. 北京:中国石化出版社, 2010.
- [20] 杨丹, 刘新, 王川丕, 等. 绿茶样品粒度对近红外光谱图和模型的影响[J]. 食品科技, 2012(11):278-281.
- [21] 杰尔·沃克曼, 洛伊斯·文依, 褚小立. 近红外光谱解析实用指南[M]. 北京:化学工业出版社, 2009.
- [22] 胡玉伟, 戢太云, 管楚雄, 等. 人工种植的铁皮石斛主要有害生物及防治对策[J]. 江苏农业科学, 2014, 42(4):98-100.
- [23] 余乐, 兰芹英, 姜宗庆. 铁皮石斛离体快繁技术[J]. 江苏农业科学, 2014, 42(11):268-270.
- [24] 陈淑钦, 陈清西. 铁皮石斛栽培及光合特性研究进展[J]. 江苏农业科学, 2015, 43(7):262-264.
- [15] 杨杰, 田永超, 姚霞, 等. 水稻上部叶片叶绿素含量的高光谱估算模型[J]. 生态学报, 2009, 29(12):6561-6571.
- [16] 许大全. 叶绿素含量的测定及其应用中的几个问题[J]. 植物生理学通讯, 2009, 45(9):896-898.
- [17] 郑坚, 陈秋夏, 金川, 等. 不同 TTC 法测定枫香等阔叶树容器苗根系活力探讨[J]. 浙江农业科学, 2008, 1(1):39-42.
- [18] 周瑞金, 杜晓华, 刘遵春, 等. 金光杏树叶绿素含量变化规律研究[J]. 中国农学通报, 2010, 26(6):201-203.
- [19] 唐延林, 黄敬峰, 王人潮. 水稻不同发育时期高光谱与叶绿素和类胡萝卜素的变化规律[J]. 中国水稻科学, 2004, 18(1):59-66.
- [20] 陆鉴眉, 林金水, 陈金河. 7 种观叶小盆栽植物的光合特性与耐阴性研究[J]. 热带作物学报, 2013, 34(4):732-737.
- [21] 梁秀英, 魏长礼, 李沿宁, 等. 君子兰叶片光合特性的研究——叶绿素和酸含量变化[J]. 东北师大学报:自然科学版, 1993(4):72-76.
- [22] Poorter L. Growth responses of 15 rain - forest tree species to a light gradient; the relative importance of morphological and physiological traits [J]. Functional Ecology, 1999, 13(3):396-410.
- [23] Lusk C H, Contreras O. Foliage area and crown nitrogen turnover in temperate rain forest juvenile trees of differing shade tolerance [J]. Journal of Ecology, 1999, 87(6):973-983.
- [24] Malavasi U C, Malavasi M M. Leaf characteristics and chlorophyll concentration of *Schyzolobium parahybium* and *Hymenaea stilbocarpa* seedlings grown in different light regimes [J]. Tree Physiology, 2001, 21(10):701-703.
- [25] Azzaz N A, Hassan E A, El - Emarey F A. Physiological, anatomical, and biochemical studies on pot marigold ( *Calendula officinalis* L. ) plants [C]. African Crop Science Conference Proceedings. 2007, 8:1727-1738.
- [26] 陈燕玲, 吴柳泳, 柯沛强. 不同光照处理对绿萝生长发育的影响[J]. 四川林业科技, 2015, 36(1):102-103, 42.
- [27] 董守坤, 赵坤, 刘丽君, 等. 干旱胁迫对春大豆叶绿素含量和根系活力的影响[J]. 大豆科学, 2011, 30(6):949-953.

(上接第 289 页)