

文章编号: 0253-2239(2008)12-2302-05

近红外光谱法快速检测绿茶中儿茶素的含量

赵杰文 郭志明 陈全胜* 吕强

(江苏大学食品与生物工程学院, 江苏 镇江 212013)

摘要 提出了一种应用傅里叶近红外漫反射光谱分析技术快速检测茶叶中主要儿茶素含量的新方法。首先获取茶叶在 $10000\sim 4000\text{ cm}^{-1}$ 范围的近红外漫反射光谱,然后以高效液相色谱分析值作参考值,采用偏最小二乘法建立茶叶中表没食子儿茶素没食子酸酯(EGCG)、表儿茶素没食子酸酯(ECG)和表没食子儿茶素(EGC)含量的定量分析模型,通过交互验证方法来优化模型的主成分数和所采用的光谱预处理方法。EGCG、ECG 和 EGC 三个模型预测值和参考值间的相关系数分别为 0.9800、0.9763 和 0.9853,预测均方根误差分别为 0.3509、0.1147 和 0.1365。研究表明,近红外光谱技术可成功地检测茶叶中 EGCG、ECG 和 EGC 的含量。

关键词 光谱学; 近红外光谱法; 儿茶素; 偏最小二乘法; 茶叶

中图分类号 O657.3 **文献标识码** A **doi**: 10.3788/AOS20082812.2302

Feasibility Study on Use of Near-Infrared Spectroscopy in Quantitative Analysis of Catechins in Green Tea

Zhao Jiewen Guo Zhiming Chen Quansheng Lü Qiang

(School of Food and Biological Engineering, Jiangsu University, Zhenjiang, Jiangsu 212013, China)

Abstract Near-infrared spectroscopy (NIR) combined with partial least squares (PLS) for the prediction of catechins compounds in the green tea was developed. The original spectra of tea in wavelength range of $10000\sim 4000\text{ cm}^{-1}$ was acquired. Reference measurement of epigallocatechin gallate (EGCG), epicatechin gallate (ECG) and epigallocatechin (EGC) were performed by reversed-phase high-performance liquid chromatography (HPLC). PLS is used to perform the calibration models to predict the contents of EGCG, ECG and EGC. To decide upon the number of PLS factors included in the PLS model and the method of spectra preprocessing, the model with the lowest root mean square error of cross-validation (RMSECV) for the training set is chosen. The correlation coefficient between the predicted and the reference results for the test set is used as an evaluation parameter for the models. The results show that the correlation coefficients of the prediction models are 0.9800 for the EGCG, 0.9763 for the ECG and 0.9853 for the EGC, with root mean squared error of prediction (RMSEP) 0.3509, 0.1147 and 0.1365 respectively. The study demonstrates that NIR spectroscopy with PLS can be used successfully to determine the content of EGCG, ECG and EGC in green tea.

Key words spectroscopy; near-infrared spectroscopy; catechins; partial least squares (PLS); tea

1 引 言

儿茶素已广泛应用于食品、制药、日用化妆品和农业等方面^[1-4],同时是构成绿茶滋味的物质基础,其品质指数已成为评判绿茶品质的重要指标。影响绿茶滋味的主要有表没食子儿茶素没食子酸酯(Epigallocatechin gallate, EGCG)、表儿茶素没食子酸酯(Epicatechin gallate, ECG)和表没食子儿茶素(Epigallocatechin, EGC)。现行的国际标准规定儿茶素的测定方法为高效液相色谱法(HPLC),该方

法虽然精度高,但样本的前处理过程繁琐、时间长、费用高,而且还污染工作环境,无法满足茶叶生产加工和贸易过程中内部有效成分快速检测的需要。

近红外光谱法(NIR)因其快速、简便、低成本、非破坏性和多组分同时测定等优点受到人们的重视,已广泛应用于农业、食品、石油、医药等领域^[5-7]。国内外学者先后利用近红外光谱法定性和定量分析了茶叶中总多酚、咖啡碱、类黄酮、游离氨基酸、茶多糖以及水分的含量^[8-12],但研究主要集

收稿日期: 2008-02-29; 收到修改稿日期: 2008-04-24

基金项目: 国家科技支撑计划(2006BAD11A12)和江苏省自然科学基金重大项目(BK2006707-1)资助课题。

作者简介: 赵杰文(1945—),男,教授,博士生导师,主要从事食品与农产品无损检测等方面的研究。

* 通信联系人。E-mail: chenjiang0518@yahoo.com.cn

中在组分总量的预测分析,而利用近红外光谱法在定量分析茶叶中单一有效成分上的研究还较少。本文提出了一种可靠、准确、快速地检测茶叶中儿茶素(包括 EGCG、ECG 和 EGC)含量的方法。

2 实验部分

2.1 仪器与材料

Antaris II 傅里叶变换近红外光谱仪(FT-NIR)(美国 Thermo Scientific 公司),InGaAs 检测器,设备配备小样品杯。LC-20A 高效液相色谱仪(日本,SHIMADZU),C18 色谱柱(VP-ODS, 250 mm × 4.6 mm, 5 μm);紫外可见检测器(Prominence SPD-20A),梯度系统(LC-20AT)。对照品 EGCG、ECG、EGC 购于 Sigma 公司,乙腈为色谱纯(国药集团化学试剂有限公司),水为 Mili-Q 超纯水(美国 Millipore 公司),其他试剂均为分析纯。茶叶样品从市场上购买,等级均为一级,分别为黄山毛尖、黄山绿茶、信阳毛尖、信阳云雾绿茶、福建毛尖、云南毛尖、滇绿、西湖龙井、茅山青峰、碧螺春(吴江)和炒青绿茶(镇江),生产日期为 2007 年 3 月至 2007 年 5 月。

2.2 实验方法

试验前先将茶叶样品分别粉碎并过 40 目筛,随机称取 1g 作为一个样本。试验选用 11 种茶叶,每种茶叶选择 10 个样本,共 110 个茶叶样本,选择其中 75 个茶叶样本建立偏最小二乘法(PLS)校正模

型,其余的 35 个作为预测集来验证模型的稳健性。首先对茶叶样本进行近红外光谱采集,然后利用高效液相测定样本中 EGCG、ECG 和 EGC 的含量。

采用 FT-NIR 测定茶叶样本的光谱。光谱采集条件:以仪器内置背景为参比,积分球漫反射,扫描范围为 10000~4000 cm⁻¹,扫描次数 32 次,分辨率 4 cm⁻¹,每个样本不同时间测定四次,取其平均光谱作为该样本的原始光谱。

利用 PLS 对茶叶中 EGCG、ECG 和 EGC 的含量进行分析,用交互验证法确定主成分数。以含量的实测值与预测值的相关系数 *R*、交互验证均方根误差(Root mean square error of cross-validation, RMSECV)及预测均方根误差(Root mean squared error of prediction, RMSEP)作为评价各种方法的有效指标。数据分析在 Matlab V7.0 (Mathworks, Natick, USA) 的软件平台上完成。

3 结果与讨论

3.1 EGCG、ECG 和 EGC 含量的测定结果

采集光谱后,将每个茶叶样本经浸泡、超声、离心、过 0.45 μm 微孔滤膜等一系列预处理后,所得样本用高效液相色谱测定 EGCG、ECG、EGC 的含量,色谱条件参考 ISO-14502-2006。通过 HPLC 测得的 EGCG、ECG 和 EGC 结果如表 1。

表 1 HPLC 测定的样本中 EGCG、ECG 和 EGC 的质量含量

Table 1 Mass contents of EGCG, ECG and EGC in tea samples by HPLC

Component	Calibration set (75 samples)			Prediction set (35 samples)		
	Range	Mean	Standard deviation	Range	Mean	Standard deviation
EGCG / %	7.340~14.304	11.236	1.824	7.651~14.088	11.184	1.764
ECG / %	1.764~3.784	2.595	0.552	1.845~3.743	2.695	0.544
EGC / %	2.126~5.428	3.873	0.809	2.336~5.392	3.779	0.799

模型预测结果的好坏不仅取决于样品评价参数的化学检测值的精度,还取决于评价参数覆盖的范围。所以在实验过程中,采用分析精度高的高效液相色谱法,并选择不同产地不同品种的茶叶样本,有利于所建模型的适应性和稳定性。

3.2 光谱波段的选择

控制光谱采集在相同的条件下进行,样品的装样厚度、装样的紧密性和颗粒均匀度等都力求一致,以免影响实验结果^[13]。图 1 为预测集样本的原始光谱图,*R* 为样本光谱的反射值,可见,在 10000~9000 cm⁻¹ 区间内光谱曲线平滑,吸光度变化不大,在 5000~4000 cm⁻¹ 区间存在较强的末端吸收。研

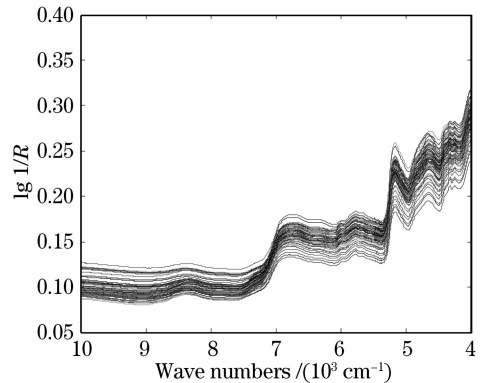


图 1 预测集样本的原始光谱图

Fig. 1 Original spectra of tea prediction sets

究选取各波段进行对比分析,结果显示在 9000 ~ 5000 cm^{-1} 范围内的光谱数据取得了较好的试验结果,并且避开了水分的影响。

3.3 光谱预处理方法的选择

近红外光谱采集时,有许多高频随机噪声、基线漂移、样本颗粒大小和光散射等噪声^[14],这将干扰近红外光谱与样品内有效成分含量间的关系,并直接影响建立模型的可靠性和稳定性。文献[8~10, 14]表明均一化(Mean centering, MC)、多元散射校正(Multiplicative scatter correction, MSC)、标准归一化(Standard normal variate transformation, SNVT)和极小/极大归一(min/max normalization, Min/Max)等四种方法可有效校正近红外原始光

谱,消除夹入的噪声信息。本文对比分析这四种光谱预处理方法,优选最佳的方法建立 PLS 模型并进行检验。

3.4 模型与预测

在建立 PLS 模型的过程中,光谱预处理方法和选择的最佳主成分数对模型有很大影响。试验中设置置信度为 95%,以交互验证均方根误差最小来选择最佳的光谱预处理方法,并确定校正模型的主成分因子数。图 2 为 EGCG、ECG 和 EGC 分别采用不同预处理方法建立的 PLS 模型中交互验证均方根误差 E_{RMSECV} 与主成分因子数间的关系图。表 2 为 EGCG、ECG 和 EGC 分别采用不同预处理方法所建立的 PLS 模型的结果。

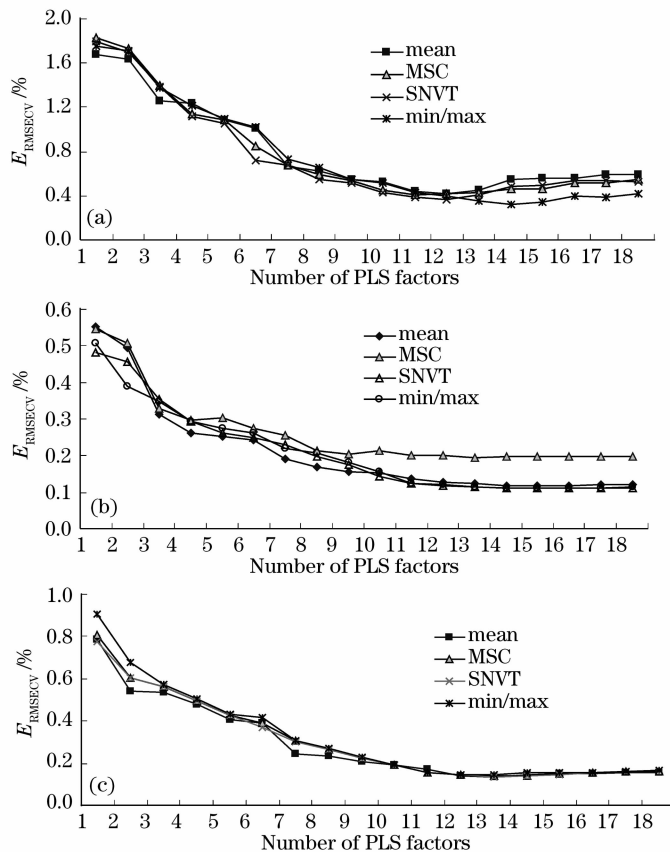


图 2 EGCG (a), ECG (b)和 EGC (c)的均方根误差与主成分因子数间的关系

Fig. 2 Number of PLS factors versus root-mean-square error for EGCG (a), ECG (b) and EGC (c)

对 EGCG 含量的预测模型,由图 2(a)可见,当采用 min/max 预处理方法可以得到最佳的 PLS 模型,模型的最佳主成分数为 14。由表 2 可见,其模型的校正集相关系数 R_{train} 、预测集相关系数 R_{test} 、 E_{RMSECV} 和预测均方根误差 E_{RMSEP} 分别为 0.9846, 0.9800, 0.3183 和 0.3509。图 3(a)所示为经 min/max 预处理后 PLS 模型的茶叶中 EGCG 百分质量含量的实测值与预测值的散点图,可知 EGCG 的测量值与预测值的

相关性较高。

对 ECG 含量的预测模型,由图 2(b)可见,当采用 SNVT 预处理方法可得到最佳的 PLS 模型,模型的最佳主成分数为 14。SNVT 处理可达到消除固体颗粒大小、表面散射及光程变化对 NIR 光谱的影响。由表 2 可见,ECG 含量的 PLS 模型的 R_{train} 、 R_{test} 、 E_{RMSECV} 和 E_{RMSEP} 分别为 0.9794, 0.9763, 0.1106 和 0.1176。图 3(b)所示为经 SNVT 预处理后 PLS 模型

的茶叶中 ECG 百分含量的实测值与预测值的散点图,可知 ECG 的测量值与预测值有较高相关性。

由图 2(c)可见,对 EGC 含量的预测模型,当采用 MC 预处理方法可得到最佳的 PLS 模型,模型的最佳主成分数为 13。归一化处理可消除获取光谱时因测

量带来的比例误差。EGC 含量的 PLS 模型的 R_{train} 、 R_{test} 、 E_{RMSECV} 和 E_{RMSEP} 分别为 0.9859,0.9853,0.1352 和 0.1365。图 3(c)所示为经 MC 预处理后 PLS 模型的茶叶中 EGC 百分质量含量的实测值与预测值的散点图,可知 EGC 的测量值与预测值也有较高的相关性。

表 2 每种预处理方法 EGCG、ECG 和 EGC 预测模型的最佳结果

Table 2 Best results for different preprocessing methods for prediction model of EGCG, ECG and EGC

Components	Preprocessing method	PLS factors	E_{RMSECV}	E_{RMSEP}	R_{train}	R_{test}
EGCG	MC	12	0.4234	0.4047	0.9727	0.9733
	MSC	11	0.4098	0.3800	0.9744	0.9765
	SNVT	12	0.3694	0.3839	0.9763	0.9760
	min/max	14	0.3183	0.3509	0.9846	0.9800
ECG	MC	15	0.1183	0.1177	0.9764	0.9764
	MSC	13	0.1937	0.12042	0.9355	0.9753
	SNVT	14	0.1106	0.1176	0.9794	0.9763
	min/max	15	0.1109	0.11470	0.9793	0.9775
EGC	MC	13	0.1352	0.1365	0.9859	0.9853
	MSC	13	0.1382	0.1444	0.9853	0.9835
	SNVT	13	0.1375	0.1424	0.9854	0.9840
	min/max	12	0.1462	0.1639	0.9836	0.9787

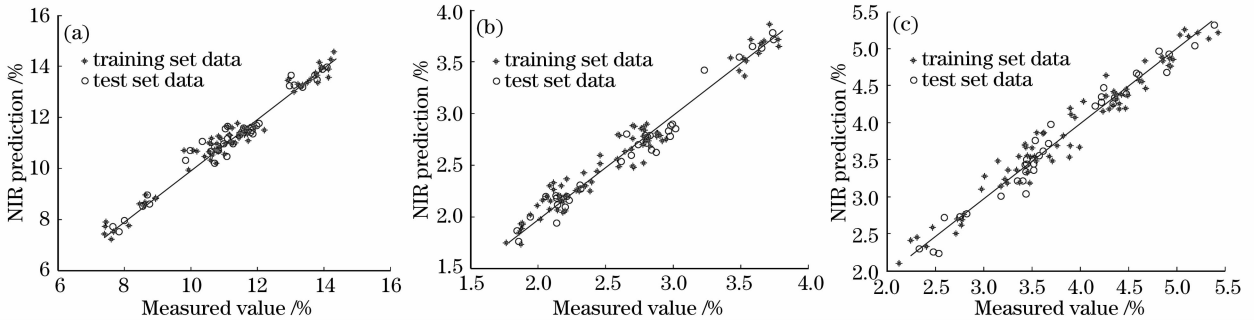


图 3 EGCG (a)、ECG (b)和 EGC (c)模型的测定值与预测值之间的散点图

Fig. 3 Measured value versus NIR prediction for EGCG (a), ECG (b) and EGC (c)

4 结 论

试验采用偏最小二乘回归法建立了茶叶中 EGCG、ECG 和 EGC 的近红外光谱定量分析模型。研究表明,茶叶中 EGCG、ECG 和 EGC 的含量用近红外光谱法和高效液相色谱分析的结果之间有良好的相关性,校正集相关系数分别达到了 0.9846、0.9794 和 0.9859,预测集相关系数分别达到 0.9800、0.9763 和 0.9853,可达到国际标准中测定儿茶素所要求的精度(测定同一样本的差异不超过 5%),这说明利用近红外光谱技术可以作为一种可靠、准确、快速的检测方法检测茶叶中 EGCG、ECG 和 EGC 的含量。

参 考 文 献

1 Yoshinori Masukawa, Yuji Matsui, Namii Shimizu *et al.*.

Determination of green tea catechins in human plasma using liquid chromatography-electrospray ionization mass spectrometry [J]. *J. Chromatography B*, 2006, **834**(1): 26~34

2 Jae Min Song, Ki Duk Park, Kwang Hee Lee *et al.*. Biological evaluation of anti-influenza viral activity of semi-synthetic catechin derivatives[J]. *Antiviral Research*, 2007, **76**(2): 178~185

3 Liu Chao, Chen Ruoyun. Advance of chemistry and bioactivities of catechin and its analogues [J]. *China J. Chinese Materia Medica*, 2004, **29**(10): 1017~1021

刘超,陈若芸. 儿茶素及其类似物的化学和生物活性研究进展[J]. *中国中药杂志*, 2004, **29**(10): 1017~1021

4 Li Hua, Liu Yuming. Study of the relationship between structures and antioxidative activities of catechin [J]. *Chemical Research and Application*, 2003, **15**(4): 524~526

李华,刘玉明. 儿茶素抗氧化活性的构效关系[J]. *化学研究与应用*, 2003, **15**(4): 524~526

5 Wang Duojia, Zhou Xiangyang, Jin Tongming *et al.*. Application of near-infrared spectroscopy to agriculture and food analysis [J]. *Spectroscopy and Spectral Analysis*, 2004, **24**(4): 447~450

王多加,周向阳,金同铭等. 近红外光谱检测技术在农业和食品

- 分析上的应用[J]. 光谱学与光谱分析, 2004, **24**(4): 447~450
- 6 Zou Xiaobo, Zhao Jiewen. Methods of characteristic wavelength region and wavelength selection based on genetic algorithm[J]. *Acta Optica Sinica*, 2007, **27**(7): 1316~1321
- 邹小波, 赵杰文. 用遗传算法快速提取近红外光谱特征区域和特征波长[J]. 光学学报, 2007, **27**(7): 1316~1321
- 7 Zhao Jiewen, Zhang Haidong, Liu Muhua. Preprocessing methods of near-infrared spectra for simplifying prediction model of sugar content of apples[J]. *Acta Optica Sinica*, 2006, **26**(1): 136~140
- 赵杰文, 张海东, 刘木华. 简化苹果糖度预测模型的近红外光谱预处理方法[J]. 光学学报, 2006, **26**(1): 136~140
- 8 Quansheng Chen, Jiewen Zhao, Haidong Zhang *et al.*. Feasibility study on qualitative and quantitative analysis in tea by near infrared spectroscopy with multivariate calibration [J]. *Analytica Chimica Acta*, 2006, **572**(1): 77~84
- 9 Quansheng Chen, Jiewen Zhao, Xingyi Huang *et al.*. Simultaneous determination of total polyphenols and caffeine contents of green tea by near-infrared reflectance spectroscopy [J]. *Microchemical J.*, 2006, **83**(1): 42~47
- 10 J. Luypaert, M. H. Zhang, D. L. Massart. Feasibility study for the use of near infrared spectroscopy in the qualitative and quantitative analysis of green tea, *Camellia sinensis* (L.) [J]. *Analytica Chimica Acta*, 2003, **478**(2): 303~312
- 11 H. Schulz, E. Joubert, W Schutze. Quantification of quality parameters for reliable evaluation of green rooibos (*Aspalathus linearis*) [J]. *Eur. Food Res. Technol.*, 2003, **216**(6): 539~543
- 12 Luo Yifan, Guo Zhenfei, Zhu Zhenyu *et al.*. Studies on ANN models of determination of tea polyphenol and amylose in tea by near-infrared spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2005, **25**(8): 1230~1233
- 罗一帆, 郭振飞, 朱振宇等. 近红外光谱测定茶叶中茶多酚和茶多糖的人工神经网络模型研究[J]. 光谱学与光谱分析, 2005, **25**(8): 1230~1233
- 13 Chen Quansheng, Zhao Jiewen, Zhang Haidong *et al.*. Identification of authenticity of tea with near infrared spectroscopy based on support vector machine [J]. *Acta Optica Sinica*, 2006, **26**(6): 933~937
- 陈全胜, 赵杰文, 张海东等. 基于支持向量机的近红外光谱鉴别茶叶的真伪[J]. 光学学报, 2006, **26**(6): 933~937
- 14 Lu Wanzhen. *Modern Near Infrared Spectroscopy Analytical Technology* [M]. second edition, Beijing: China Petrochemical Press, 2007. 35~37
- 陆婉珍. 现代近红外光谱分析技术[M]. 第二版, 北京: 中国石化出版社, 2007. 35~37

《中国激光》“全固态激光技术”专题征稿启事

全固态激光技术是目前我国在国际上为数不多的从材料源头到激光系统集成拥有整体优势的高技术领域之一,随着全固态激光器件与材料研究的迅速发展,对激光先进制造技术、激光显示技术和激光医疗等领域的发展产生了巨大的推动作用,已取得大量研究成果。《中国激光》计划于2009年6月正刊上推出“全固态激光技术”专题栏目,现特向国内外广大读者以及作者征集“全固态激光技术”方面原创性的研究论文和综述,旨在集中反映该方面最新的研究成果及研究进展。

征稿范围包括:

- 激光二极管抽运固体激光器
- 光纤激光器
- 半导体激光器及相关技术
- 晶体材料及相关技术
- 光学薄膜技术
- 其他

截稿日期:2009年3月31日

投稿方式以及格式:通过网上投稿系统(<http://www.opticsjournal.net/zgjg.htm>)直接上传稿件(主题标明“全固态激光技术”投稿),也可直接将稿件电子版发至邮箱:zhgjg@mail.shcnc.ac.cn(主题标明“全固态激光技术”投稿),详情请参见中国光学期刊网:www.opticsjournal.net。本专题投稿文体不限,中英文皆可,其电子版请使用MS-word格式,有任何问题请发邮件至zhgjg@mail.shcnc.ac.cn询问。

《中国激光》编辑部