

# 近红外光谱在虫草菌丝体腺苷和多糖含量检测中的应用

姜福佳\*

(长春职业技术学院, 长春 130033)

**摘要:** **目的** 利用近红外光谱技术建立蝙蝠拟青霉菌丝体中腺苷和多糖含量的定量分析模型。**方法** 采用常规方法对蝙蝠拟青霉菌丝体中腺苷和多糖含量进行测定, 利用近红外光谱技术建立测定菌丝体中腺苷和多糖含量的相关模型, 并通过蒙特卡罗偏最小二乘法(Monte Carlo Partial Least Square, MCPLS)和可移动窗口偏最小二乘法(Moving Window Partial Least Square, MWPLS)对模型进行优化。**结果** 该模型校正集预测值和真实值间的相关系数( $R_c$ )分别为 0.9400 和 0.8781, 预测均方根误差(Root Mean Square Error of Prediction Set,  $RMSEP$ )分别为 0.5949 和 1.6617, 校正均方根误差(Root Mean Square Error of Calibration Set,  $RMSEC$ )分别为 0.5844 和 1.5572。**结论** 该模型的稳健性、拟合度和预测能力均能达到令人满意的程度, 该方法可以推广应用到其他发酵产品的检测。

**关键词:** 近红外光谱; 蝙蝠拟青霉; 蒙特卡罗偏最小二乘法; 定量分析

## The application of near-infrared spectroscopy in detection of adenosine and polysaccharide of cordyceps mycelium

JIANG Fu-Jia\*

(Changchun Vocational Institute of Technology, Changchun 130033, China)

**ABSTRACT: Objective** Near-infrared spectroscopy was used to quantify the adenosine and polysaccharide content in *Paecilomyces hepiali* mycelia. **Methods** Quantification of adenosine and polysaccharide content were achieved by the conventional method. Models were established by near-infrared spectroscopy for detection of the contents of adenosine and polysaccharide. Monte carlo partial least square (MCPLS) and moving window partial least square (MWPLS) were applied to optimize the models. **Results** The related coefficient between actual and predictive values of calibration samples ( $R_c$ ) were 0.9400 and 0.8781, the root mean square error of prediction set ( $RMSEP$ ) were 0.5949 and 1.6617, and the root mean square error of calibration set ( $RMSEC$ ) were 0.5844 and 1.5572, respectively. **Conclusion** The robustness, fitness and predictive capability of these models were satisfied, and it can be applied to other fermentation processes.

**KEY WORDS:** Near-infrared spectroscopy; *Paecilomyces hepiali* mycelia; moving window partial least square; quantitative analysis

蝙蝠拟青霉(*Paecilomyces hepiali*)是天然冬虫夏草中普遍存在的一种无性型虫草<sup>[1,2]</sup>。其发酵的菌

丝体可以作为天然冬虫夏草的替代品<sup>[3,4]</sup>。冬虫夏草属于我国特有的名贵药用真菌。其主要功效是“补肺

\*通讯作者: 姜福佳, 硕士, 助教, 生物技术及应用专业带头人, 研究领域: 微生物与生化药学。E-mail: ai82180076@sohu.com

益肾, 止血, 化痰; 用于久咳虚喘, 劳嗽咯血, 阳痿遗精, 腰膝酸痛”。但是由于人们对天然虫草的过度采集, 天然虫草资源已经濒临枯竭<sup>[5,6]</sup>。目前很多研究结果都证实通过液体发酵产生的虫草菌丝体的有效成分和疗效与天然冬虫夏草相似, 因此作为可替代品的蝙蝠蛾拟青霉显得尤为重要, 近年来对该菌株的研究开发愈加深, 相应食品药物及保健品的开发及检测工艺急需建立。

相对其它分析技术来说, 近红外光谱(Near-infrared spectroscopy, NIR)技术具有方便、快捷、无损、可多组分同时测定、灵活等众多优点, 非常适合蝙蝠蛾拟青霉菌丝体多种活性成分含量的同时检测<sup>[7,8]</sup>。然而, 近红外光谱是分子振动光谱, 其光谱信息弱且复杂, 其有效信息的提取相当困难, 因此近红外光谱技术需结合化学计量学方法进行应用。同时近红外光谱是一种间接测量技术, 建立稳健性和泛化能力好的校正模型是近红外光谱技术推广应用的关键<sup>[9,10]</sup>。

本文采用近红外漫反射光谱技术结合蒙特卡罗偏最小二乘法和可移动窗口偏最小二乘法提取各组分的 NIR 有效信息并建立非线性校正模型, 构造一种同时测定蝙蝠蛾拟青霉菌丝体中腺苷和多糖含量的新方法。这对于丰富我国蝙蝠蛾拟青霉菌丝体有效成分定量分析的快速检测手段, 提高相应检测水平, 以及对新型食品和保健品的研发都具有重要的学术意义。

## 1 材料与方 法

### 1.1 材料和仪器

蝙蝠蛾拟青霉, 购于安徽农业大学。

葡萄糖、磷酸二氢钾、硫酸镁和高碘酸钠等均为分析纯; 蛋白胨和酵母浸粉等均为化学试剂; 甲醇为色谱纯。

UV3150 紫外可见漫反射近红外分光光度计(日本岛津公司); ISR-3100 积分球(日本岛津公司); 752 紫外分光光度计(上海分析仪器厂); SHIMADZU 高压液相色谱仪(日本岛津公司)。

### 1.2 微生物发酵及样品采集

蝙蝠蛾拟青霉在不同的发酵条件下进行发酵培养, 其发酵条件为: 发酵培养基初始 pH 为 6~8, 发酵温度为 22~26 °C, 摇床转速 200~300 r/min, 接种量为 2%~4%, 种龄为 3~5 d。发酵过程结束后, 收集菌丝体样本, 冻干并研磨成粉保存于塑封袋中。

### 1.3 测定虫草酸和腺苷含量

采用常规方法测定蝙蝠蛾拟青霉菌丝体中腺苷<sup>[11]</sup>和多糖含量<sup>[12]</sup>。

### 1.4 收集样本近红外光谱

对各个批次的蝙蝠蛾拟青霉菌丝体样本进行近红外光谱扫描, 设置光谱范围为 800~2500 nm, 光谱带宽为 12 nm, 光谱间隔为 1 nm, 每个样本扫描 3 次取平均值作为该样本的近红外光谱。

### 1.5 校正模型的建立及优化

应用 Origin 7.5 程序对采集的近红外原始光谱依次进行卷积平滑(Savitzky-Golay Smoothing)、快速傅里叶变换(Fast Fourier transform, FFT)、一阶导数(First order derivative)和二阶导数(Second order derivative)方法的预处理, 应用 Matlab 2008Ra 程序对光谱进行标准正态分布(SNV)法的预处理; 应用 Matlab 2008Ra 程序结合蒙特卡罗偏最小二乘法识别样本中的异常样本并考察校正集和预测集的配比, 根据光谱第一主成分得分(PC<sub>1</sub>)和第二主成分得分(PC<sub>2</sub>)的散点图选择校正集样本和预测集样本; 结合可移动窗口偏最小二乘法考察不同预处理方法窗口大小、波长变量和隐变量数对模型的影响, 以各个模型的逼近度(Da)值、校正均方根误差(RMSEC)和预测均方根误差(RMSEP)为评价指标选择最佳模型。

## 2 结果与讨论

### 2.1 蝙蝠蛾拟青霉样本全光谱

对收集的 159 个蝙蝠蛾拟青霉菌丝体样本进行近红外光谱扫描, 样本的原始光谱如图 1 所示。

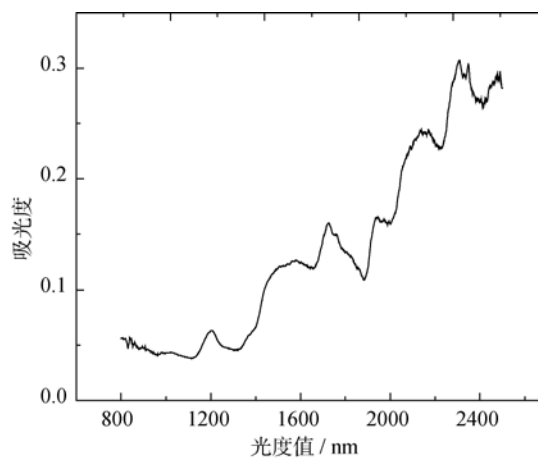


图 1 原始近红外光谱

Fig. 1 The average of NIR reflectance spectra

## 2.2 异常样本的识别及剔除

应用蒙特卡罗偏最小二乘法(MCPLS)识别用于建立模型的异常样本, 计算得到的预测残差平均值(Mean Predictive Residual Error, MPRE)和预测残差间标准偏差(Standard Deviation Predictive Residual Error, SDPRE)散点图如图2所示, 由图2可知, 对于159个样本, 用于建立测定腺苷和多糖含量的模型样本中分别存在7和6个异常样本, 予以剔除。

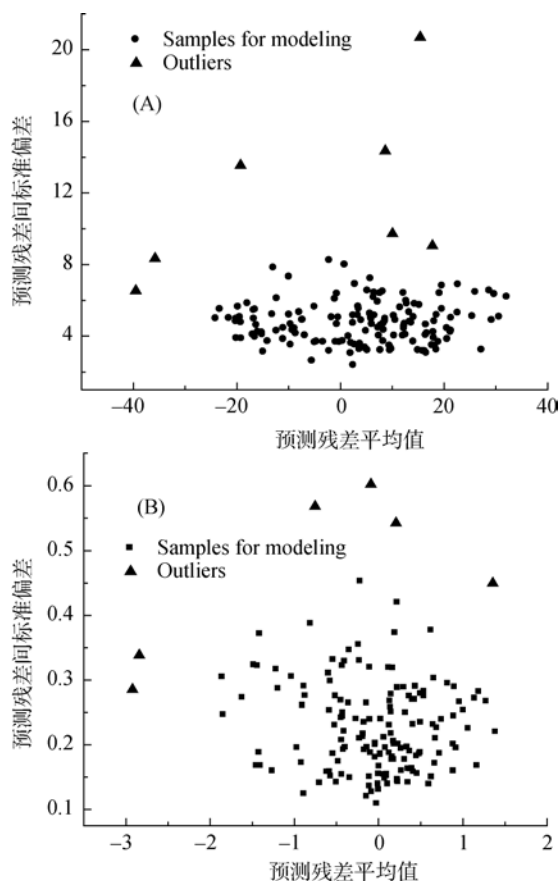


图2 SDPRE对MPRE的散点图(A: 腺苷; B: 多糖)

Fig. 2 Scatter diagram of SDPRE and MPRE (A: adenosine; B: polysaccharide)

## 2.3 样本的分集

采用MCPLS法选择合适的校正集样本数, 其中校正集样本数所占总样本数的比率对平均拟合度(the mean degree of fitting, MDf)的影响如图3所示, 当校正集样本所占比率分别达到70%和60%时, 建立测定腺苷和多糖模型的MDf值达到最高, 并且依据各自模型光谱的第一主成分得分(PC<sub>1</sub>)和第二主成分得分(PC<sub>2</sub>)的散点图分别选择出校正集样本和预测集样品。校正集与预测集样本中腺苷和多糖含量统计结果

如表1所示。

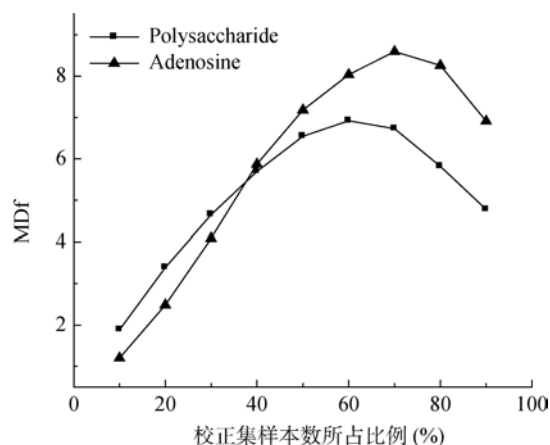


图3 校正集样本所占比率对MDf值的影响(A: 腺苷; B: 多糖)

Fig. 3 The ratio of calibration set vs MDf (A: adenosine, B: polysaccharide)

## 2.4 模型光谱预处理方法、波长变量和隐变量数的选择

分别采用原始光谱和不同预处理方法建立测定蝙蝠蛾拟青霉菌丝体中腺苷和多糖含量的定量分析模型, 建立的每个模型经过可移动窗口偏最小二乘法(MWPLS)选择最适的波长变量和隐变量数, 其结果如表2所示。由表2可知, 在建立测定腺苷含量模型中仅一阶导数的预处理方法 $Da$ 值相较原始光谱的 $Da$ 值略低, 其它模型 $Da$ 值均高于原始光谱建立的模型。表明MWPLS法选择波长变量和隐变量数的效果很好, 同时MWPLS能够有效地筛选信息, 去除大量的随机噪音和无用波长变量等, 大大减小输入数据矩阵, 从而减少计算量和缩短计算时间。建立测定腺苷和多糖含量的模型, 最有效的预处理方法均为FFT、最适的窗口大小( $W$ )分别为120和24、最优的 $n_w$ 分别50和20、最适的隐变量数分别为7和13。

## 2.5 最优模型的建立

依照MWPLS法优化的最适预处理方法、波长变量和隐变量数建立测定白囊耙齿菌菌丝体中腺苷和多糖含量的最佳多元校正模型。该模型结果如图4所示, 最佳模型对腺苷和多糖含量预测的模型校正集预测值和真实值间的相关系数( $R_c$ )和模型预测集预测值与真实值间的相关系数( $R_p$ )均高于0.8500, 该模型的拟合度、稳健性和预测能力均能达到令人满意的程度。

表 1 腺苷和多糖含量统计结果

Table 1 Statistics results of the adenosine and polysaccharide

成分	组别	数量	最大值	最小值	平均值
腺苷(μg/g)	校正集	112	7.2549	0.2686	3.0284
	预测集	47	6.2467	0.3261	3.1302
多糖(g/g)	校正集	96	0.1444	0.0332	0.0617
	预测集	63	0.1398	0.0428	0.0636

表 2 光谱预处理方法、波长变量和隐变量数对模型的影响

Table 2 Influence of different preprocessing methods, wavelength and Implicit variable on capabilities of the model

成分	预处理方法	Windows	W	n <sub>w</sub>	n <sub>LV</sub> <sup>1</sup>	RMSEC	RMSEP	Da
腺苷	Original spectra	---	24	5	4	0.6120	0.6469	1.5158
	Savitzky-Golay Smoothing	5	120	5	6	0.5245	0.5877	1.6478
	FFT	13	24	20	13	0.5844	0.5949	1.6720
	First order derivative	7	60	5	5	0.6065	0.6247	1.5867
	Second order derivative	13	24	5	5	0.6397	0.6541	1.5186
	SNV	---	120	5	6	0.4629	0.5862	1.6037
	Original spectra	---	60	35	7	12.6955	16.8704	0.5387
多糖	Savitzky-Golay Smoothing	5	180	5	5	15.6144	16.6979	0.5835
	FFT	9	120	50	7	1.55724	1.66171	0.5868
	First order derivative	9	180	10	3	1.61347	1.74256	0.5576
	Second order derivative	11	60	5	7	1.88475	2.14833	0.4434
	SNV	---	180	5	5	1.58253	1.73355	0.5572

W: 选择的波长点数; n<sub>w</sub>: 选择的窗口数; Windows: 预处理的窗口大小

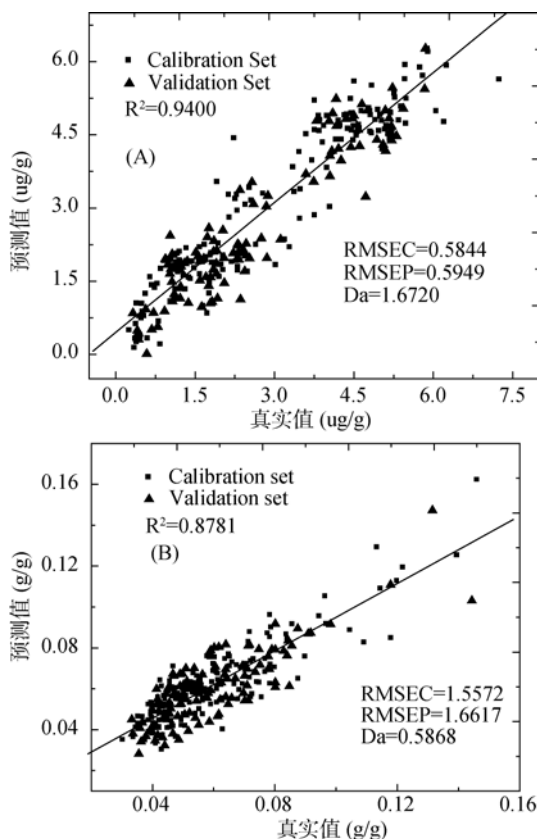


图 4 最佳模型的预测值与真实值的相关图(A 腺苷; B 多糖)

Fig. 4 Correlation between experimental values and predicted values (A: adenosine, B: polysaccharide)

### 3 总 结

本文采用近红外漫反射光谱技术结合蒙特卡罗偏最小二乘法和可移动窗口偏最小二乘法对蝙蝠蛾拟青霉菌丝体中腺苷和多糖含量进行建模分析, 得到最佳的定量分析模型。建立的两个分析模型稳健性、拟合度和预测能力均能达到令人满意的程度, 而且建立模型的样本为不同发酵条件下得到的蝙蝠蛾拟青霉菌丝体, 表明该模型具有很强的泛化能力, 可以充分地满足蝙蝠蛾拟青霉菌丝体品质的快速鉴定, 便于对其深度开发及产品的监测。

采用近红外光谱技术结合偏最小二乘法建模对物质进行定量定性分析的方法, 能够被广泛的应用于各种物质的检测当中, 同时该方法较传统方法更加的省时省力且对环境不会造成较大的损伤, 说明近红外光谱技术在食品药品及物质的检测领域具有极大的应用前景。

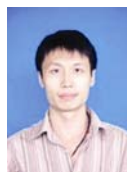
### 参考文献

- [1] 戴如琴, 兰江丽, 陈伟华, 等. 蝙蝠蛾拟青霉新种的研究[J]. 北京农业大学学报, 1989, 15(2): 221-225.
- [2] 杨金玲, 肖薇, 何慧霞, 等. 蝙蝠蛾拟青霉与冬虫夏草关系的分子系统性研究[J]. 药学学报, 2008, 43(4): 421-426.

- [3] 姜文侠, 孙武岳, 马琳, 等. 蝙蝠蛾拟青霉菌丝体发酵的研究[J]. 微生物学通报, 1999, 26(3): 192-194.
- [4] 金成, 邬晓敏, 陈高阳. 金水宝的临床应用[J]. 首都医药, 2005, 18: 42-43.
- [5] 国家药典委员会编. 中华人民共和国药典. 2010 版一部[S]. 北京: 中国医药科技出版社, 2010: 106.
- [6] 冯健, 王红旗, 张鸿军. 新引进冬虫夏草菌液体深层发酵实验[J]. 微生物学杂志, 2002, 22(3): 55-58.
- [7] Blanco M, Villarroya I. NIR spectroscopy: a rapid-response analytical tool [J]. Trac-Trend Anal Chem, 2002, 21: 240-250.
- [8] 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.
- [9] Du LN, Wu LH, Lu JH, *et al.*. Application of Near Infrared Diffuse Reflectance Spectroscopy with Radial Basis Function Neural Network to Determination of Rifampicin Isoniazide and Pyrazinamide Tablets [J]. Chem Res Chin Univ, 2007, 23(5): 518-523.
- [10] Pedro AMK, Ferreira MMC. Simultaneously calibration solids, sugar and acidity of tomato products using PLS2 and NIR spectroscopy[J]. Anal Chim Acta, 2007, 595: 221-227.
- [11] 2005 版《中国药典》第一部[M]. 2005(1), 75.
- [12] 王黎明, 夏文水. 蒽酮-硫酸法测定茶多糖含量的研究[J]. 食品科学, 2007, 26(7): 185-188.

(责任编辑: 孙媛媛)

### 作者简介



姜福佳, 硕士, 助教, 生物技术及应用专业带头人, 研究领域: 微生物与生化药学。  
E-mail: ai82180076@sohu.com