

FTIR 结合化学计量学对三七地下部位鉴别及皂苷含量预测

李 运^{1,2,3}, 张 霖^{1,2}, 金 航^{1,2}, 王元忠^{1,2*}, 张金渝^{1,2,3*}

1. 云南省农业科学院药用植物研究所, 云南 昆明 650200
2. 云南省省级中药原料质量监测技术服务中心, 云南 昆明 650200
3. 云南中医学院中药学院, 云南 昆明 650500

摘 要 当今中药市场上掺假现象屡见不鲜, 不良商贩利用三七须根粉末假冒主根和剪口粉末, 严重影响三七的质量与药效。通过傅里叶变换红外光谱(FTIR)结合化学计量学建立三七主根、剪口和须根粉末鉴别及四种皂苷含量快速预测模型, 为快速三七质量控制提供基础。采集三七主根、剪口和须根红外光谱, 超高效液相色谱(UPLC)测量样品中三七皂苷 R_1 、人参皂苷 R_{g_1} 、人参皂苷 R_{b_1} 和人参皂苷 R_d 含量。采用纵坐标归一化及二阶导数对原始红外光谱进行预处理; Kennard-stone 算法将 60 个样本分为 2/3 训练集与 1/3 预测集。训练集数据结合支持向量机(SVM)判别建立三七主根、剪口和须根粉末鉴别模型, 最佳核函数 c 和 g 采用交叉验证进行网格搜索, 预测集数据用于对判别模型进行外部验证。正交信号校正偏最小二乘回归(OSC-PLSR)建立三七中四种皂苷含量预测模型, 红外光谱采用一阶、二阶导数及 Savitsky-Golay 平滑 5 点、7 点、9 点、11 点预处理。60 个样本分为 2/3 训练集与 1/3 预测集, 训练集数据建立 OSC-PLSR 模型, 预测集数据对 OSC-PLSR 模型的预测结果进行外部验证。结果显示: (1)二阶导数可有效的分离原始谱图的叠合隐蔽谱峰, 并提高谱图的分辨率; (2)交叉验证网格搜索计算出最佳核函数 $c=2.828\ 43$, $g=4.882\ 81 \times 10^{-4}$, 此时训练集判别正确率为 100%; (3)SVM 判别模型核函数设置为最佳核函数, 预测集数据外部验证正确率为 100%, 所有样本均被正确鉴别; (4)三七皂苷 R_1 、人参皂苷 R_{g_1} 、人参皂苷 R_{b_1} 和人参皂苷 R_d 最优含量预测模型预测值与 UPLC 检测值接近, 预测效果良好。FTIR 结合 SVM 判别能对三七主根、剪口和须根粉末快速鉴别, 结合 OSC-PLSR 能对四种皂苷含量进行准确预测。该方法准确可靠, 可为中药材三七提供快速有效的质量控制。

关键词 傅里叶变换红外光谱; 三七; 地下部位鉴别; 皂苷含量预测; 化学计量学

中图分类号: O433 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)01-0103-06

引 言

中药悠久的药用历史中, 人们发现其不同部位存在药效差异。现代分析化学研究表明, 中药不同部位化学成分种类及含量有较大差异^[1-3]。当今中药市场上, 不法商贩将非药用或低含量部分混入中药中以谋取不良利益。例如: 将酸枣果壳掺入种子中粉碎作酸枣仁出售; 麻黄中混入大量地上茎等。临床用药时, 掺假中药会降低治疗效果, 甚至导致毒副作用的产生。这些不良现象严重影响了中药的质量及药用安全, 是中药质量标准化亟需解决的问题之一。

五加科 (Araliaceae) 人参属 (*Panax*) 植物三七 (*Panax*

notoginseng (Burk) F. H. Chen) 为我国特有名贵药材, 药用历史悠久^[4-5]。现代药理学研究表明, 三七总皂苷(PNS)是三七药效的主要物质基础, 具有抗糖尿病^[6-7]、抗癌^[8-9]、抗动脉粥样硬化^[10-11]、脑神经保护^[12]、降血压^[13]等诸多药理活性。2015 版中国药典记载中药三七为植物三七的干燥根和根茎; 质量控制按干燥品计算, 含人参皂苷 R_{g_1} 、人参皂苷 R_{b_1} 、三七皂苷 R_1 的总量不得少于 5%^[14]。现有皂苷类成分含量检测方法多为高效液相色谱法^[15], 样品的前处理繁琐费时, 检测过程漫长, 不利于快速检测分析, 寻找一种三七皂苷含量快速检测分析方法具有重要的实用意义。

研究表明三七主根、根茎和须根总皂苷含量存在较大差异, 其中根茎总皂苷含量最高, 须根含量最低^[16]。商品三七

收稿日期: 2016-04-03, 修订日期: 2016-10-19

基金项目: 国家自然科学基金项目(31760360), 云南省科技入滇专项(2017IB038), 云南省重大科技专项(2016ZF001-2)资助

作者简介: 李 运, 1993 年生, 云南中医学院中药学院硕士研究生 e-mail: liyunsci@163.com

* 通讯联系人 e-mail: boletus@126.com; jy Zhang2008@126.com

分为头子(主根)、剪口(根茎)和须根;其中主根价格最高,须根价格最低。三七主根、剪口和须根形态分辨较为简单,但粉末状态极难分辨^[17]。李建蕊等采用红外光谱对三七剪口、主根、筋条和绒根粉末进行分析,发现不同部位红外光谱具有差异;但是仅通过个别吸收峰差异进行鉴别,主观性与专业性较强,不利于推广使用。一些不良商贩利用三七须根粉末假冒主根和剪口粉末,严重影响了三七的质量与药效。因此,鉴别三七主根、根茎和须根粉末并以皂苷含量为指标对其进行质量控制,有利于保证三七药材的药效以及该产业的规范化发展。

红外光谱法是一种通过被测物中化合物振动时吸收特定波长红外光,从而得到化合物官能团、化学键的分析方法^[18],具有快速、无损及高灵敏度的特点。结合化学计量学方法,红外光谱广泛应用于中药种类与产地鉴别^[19-20],中药与食品标志性成分含量快速预测研究^[21-22]。本实验采用傅里叶变换红外光谱(Fourier transform infrared spectroscopy, FTIR)结合化学计量学方法对三七主根、剪口和须根粉末进行鉴别,同时对三七皂苷 R₁、人参皂苷 R_{G1}、人参皂苷 R_{B1} 和人参皂苷 R_d 含量进行预测,以期对三七的快速质量控制提供基础。

1 实验部分

1.1 仪器与试剂

Frontier 型傅里叶变换红外光谱仪(配备 DTGS 检测器,扫描范围:4 000~400 cm⁻¹,扫描信号累加 16 次,分辨率:4 cm⁻¹,Perkin Elmer 公司);LC-30AD 超高效液相色谱仪(配备二极管阵列紫外检测器、二元梯度泵、自动进样器,岛津公司);Shim-pack XR-ODS III 色谱柱(75 mm×2.0 mm,1.6 μm,100 MPa);Milli-Q 超纯水系统(Millipore 公司);超声仪;粉碎机;压片机;80 目筛盘。KBr(分析纯);甲醇、乙腈(色谱纯,赛默飞世尔公司);标准品:三七皂苷 R₁、人参皂苷 R_{G1}、人参皂苷 R_{B1} 和人参皂苷 R_d(中国药品生物制品检定所)。

1.2 样品采集

样品采集于云南省昆明市寻甸县甸沙乡麦地新村,海拔 2 090 m,经度:103°07′25.02″,纬度:25°45′33.61″。所有样品经由云南省农业科学院药用植物研究所张金渝研究员鉴定为人参属植物三七的全株。新鲜样品清洗之后自然阴干,主根、剪口和须根分离后密封保存。实验前将样品粉碎后过 80 目筛,烘箱内 50 °C 干燥恒重备用。

1.3 红外光谱采集

准确称取样品粉末(1.2±0.2) mg 及 KBr 粉末(150.0±1.0) mg,于玛瑙研钵中混合均匀并充分研磨,研磨后粉末倒入模具中压制成均匀薄片。同一样品扫描两次,样品扫描前使用空白 KBr 压片扫描背景以去除二氧化碳和水的干扰。

1.4 超高效液相色谱(UPLC)检测

流动相为超纯水(A)-乙腈(B),流速:0.35 mL·min⁻¹,柱温:45 °C,进样量 1 μL,参照中国药典三七含量测定部分设置检测波长为 203 nm。梯度洗脱程序:0~3 min,20%~

24% B;3~6 min,24%~38% B;6~12 min,38%~56% B;12~13 min,56%~20% B;13~14 min,20% B。准确称取样品粉末(40.0±2.0) mg 置于洁净试管中,加入 2.0 mL 甲醇,密封超声提取 40 min。提取结束后补充损失甲醇,使用 0.22 μm 有机微孔滤膜过滤提取液。精密称取三七皂苷 R₁ 0.7 mg,人参皂苷 R_{G1} 3.0 mg,人参皂苷 R_{B1} 3.2 mg,人参皂苷 R_d 0.8 mg,加入 1.6 mL 甲醇配置成标准品母液,将标准品母液按 50% 等梯度稀释 6 次,得到 7 个浓度梯度的标准品溶液。自动进样器吸取母液 2 与 1 μL,其余梯度吸取 1 μL 注入超高效液相色谱仪,记录不同浓度梯度标准品峰面积,并计算峰面积与标准品浓度的线性回归方程。吸取 1 μL 样品提取液注入超高效液相色谱仪,按照梯度洗脱程序进行检测,并记录 203 nm 色谱图。

1.5 数据处理

使用 EXCEL 建立标准品溶液浓度与峰面积的线性回归方程,结合样品色谱图中峰面积计算出对应皂苷含量。参与支持向量机(support vector machines, SVM)判别分析的原始红外光谱使用 OMNIC 软件进行平均光谱、纵坐标归一化、二阶导数等预处理,参与正交信号校正偏最小二乘回归模型建立的原始红外光谱进行平均光谱、纵坐标归一化、平滑、一阶导数及二阶导数等预处理。MATLAB 2014a 软件对数据进行训练集与预测集的筛选并建立 SVM 判别模型;SIMCA-P+ 11.5 软件建立正交信号校正偏最小二乘回归(orthogonal single collection partial least squares regression, OSC-PLSR)模型;ORIGIN 8.0 软件作图。

2 结果与讨论

2.1 三七主根、剪口、须根红外光谱分析

三七主根、剪口和须根的原始红外光谱见图 1(a)。由图可知,三七不同部位的原始红外光谱相似,共有峰主要集中在 2 000~400 cm⁻¹ 波段。由于原始红外光谱的谱峰堆积,分辨率较低,难以对三七主根、剪口和须根进行分类鉴别,因此将原始红外光谱进行二阶导数处理。

通过对原始红外图谱进行二阶导数处理,可以分离隐蔽的叠合谱峰,有效提高谱图的分辨率^[23-24]。三七主根、剪口和须根的二阶导数红外光谱见图 1(b)。由图可知,须根的二阶导数光谱在 3 800~3 300 cm⁻¹ 波段内有明显吸收峰,主根和剪口在此波段吸收峰不明显;2 000~400 cm⁻¹ 波段,三七主根、剪口和须根吸收峰强度有较大差别。3 800~3 300 cm⁻¹ 波段主要归属于羟基伸缩振动吸收,2 000~400 cm⁻¹ 波段主要为 C=C, C=O, C=N 伸缩振动,C-H 弯曲振动以及 C-O 伸缩振动等。红外光谱记录了丰富的样品物理化学信息,为三七主根、剪口和须根粉末鉴别以及皂苷含量快速预测提供了理论依据。

2.2 UPLC 结果分析

对不同浓度梯度的标准品溶液进行检测,得到标准品三七皂苷 R₁、人参皂苷 R_{G1}、人参皂苷 R_{B1} 和人参皂苷 R_d 峰面积与浓度的回归系数分别为 0.998 8, 0.999 9, 0.998 7 和 0.999 9,回归系数均大于 0.998 0,线性关系良好。精密度

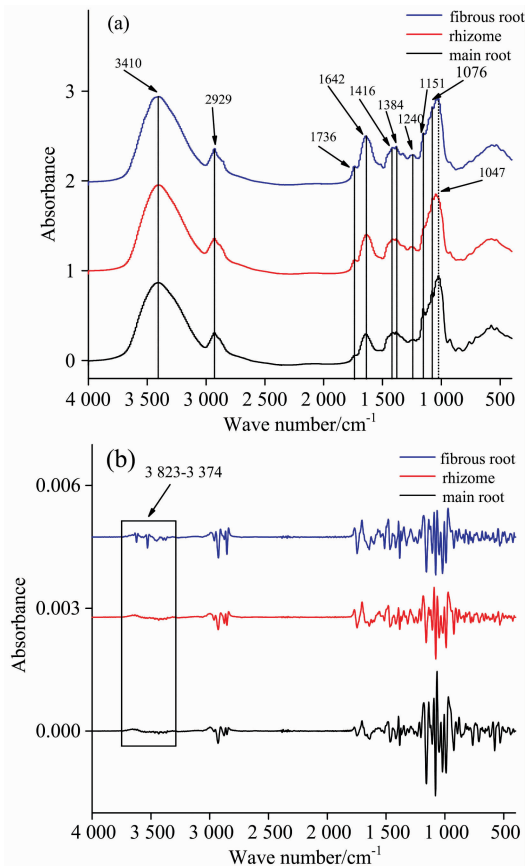


图 1 三七主根、剪口和须根的红外光谱图

(a): 原始红外光谱; (b): 二阶导数红外光谱

Fig. 1 Infrared spectra of main root, rhizome and fibrous root of *Panax notoginseng*

(a): Original infrared spectra;

(b): Second derivative infrared spectra

考察显示 ($n=6$), 样品三七皂苷 R_1 、人参皂苷 R_{g_1} 、人参皂苷 R_{b_1} 和人参皂苷 R_d 峰面积 RSD 分别为 1.09%, 0.95%, 0.96% 和 1.83%, 表明该方法精密度良好。稳定性考察显示 ($t=1$ h, $n=6$), 以上指标峰面积 RSD 分别为 1.44%, 1.98%, 1.51% 和 2.05%, 表明样品溶液在 6 h 内稳定。样品测量所得各皂苷面积结合标准品线性回归方程计算出样品中各皂苷含量。每个样品平行测定两次并取平均值, 皂苷含量结果见表 1。由表 1 可知, 剪口样品中四种皂苷以及总皂苷含量均最高, 总皂苷含量为主根的 1.8 倍, 须根的 4.6 倍。按照国家药典对三七的质量控制标准计算, 每组 20 个样品中, 剪口与主根质量合格率为 100%, 须根质量合格率仅为 20%。

2.3 SVM 判别模型建立及结果分析

采用台湾大学工业工程研究所林智仁等^[25]开发的 libsvm-3.20 工具箱建立 SVM 判别模型, 对三七主根、剪口和须根粉末进行鉴别。原始红外光谱二阶导数处理后的 1 869 个数据点作为 X 变量, 分类标签作为 Y 变量, 判别模型建立前, 将 X 变量标准化到 $[-1, 1]$ 区间。60 个样本采用 Kennard-stone 算法筛选训练集 (2/3) 与预测集 (1/3), 训练

集数据建立判别模型, 预测集数据对模型精确度进行外部验证。

表 1 三七样品皂苷含量信息

Table 1 Saponins content information of *Panax notoginseng* samples

皂苷	主根/ ($\text{mg} \cdot \text{g}^{-1}$)	剪口/ ($\text{mg} \cdot \text{g}^{-1}$)	须根/ ($\text{mg} \cdot \text{g}^{-1}$)
	Mean \pm SD	Mean \pm SD	Mean \pm SD
三七皂苷 R_1	10.02 \pm 3.30	18.69 \pm 5.90	4.53 \pm 2.00
人参皂苷 R_{g_1}	45.70 \pm 12.52	86.67 \pm 18.49	19.13 \pm 8.08
人参皂苷 R_{b_1}	40.68 \pm 12.19	63.08 \pm 19.09	14.97 \pm 6.55
人参皂苷 R_d	9.90 \pm 3.70	22.48 \pm 9.35	2.72 \pm 1.59
总皂苷	106.30 \pm 21.82	190.92 \pm 28.77	41.35 \pm 15.23

标准化的 40 个训练集数据代入 libsvm-3.20 工具箱建立 SVM 判别模型, 最佳核函数 c 和 g 采用交叉验证 (cross validation) 进行网格搜索。首先将 $\log_2 c$ 与 $\log_2 g$ 在 $[-20, 20]$ 区间内粗略搜索, 根据交叉验证正确率缩小范围, $\log_2 c$ 在 $[-5, 10]$ 区间及 $\log_2 g$ 在 $[-5, -20]$ 区间进行精细搜索, 计算出最佳核函数 $c=2.82843$, $g=4.88281 \times 10^{-4}$, 此时判别模型交叉验证正确率为 100%。将判别模型核函数设置为最佳核函数, 20 个预测集数据对训练好的模型判别精确度进行外部验证, 20 个样本全部被正确鉴别, 预测正确率为 100% (图 2)。结果显示, 建立的 SVM 判别模型外推能力较强, 能够对三七主根、剪口和须根粉末进行有效鉴别。

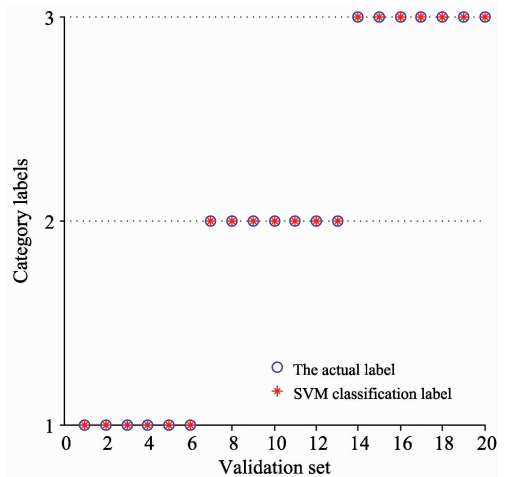


图 2 支持向量机分类标签与实际标签

标签类别 1: 主根, 2: 剪口, 3: 须根

Fig. 2 SVM classification label and the actual label

Category label 1: main root, 2: rhizome, 3: fibrous root

2.4 三七主根、剪口和须根粉末中四种皂苷含量的快速预测与结果分析

对原始红外光谱进行纵坐标归一化, 考察一阶 (1D)、二阶导数 (2D) 以及 Savitsky-Golay 平滑 5 点 (SG5)、7 点 (SG7)、9 点 (SG9)、11 点 (SG11) 预处理对模型拟合能力及预测能力的影响, 并根据预测能力筛选最优模型。预处理后的红外光谱数据作为 X 变量, 样品中三七皂苷 R_1 、人参皂苷

R_{g_1} 、人参皂苷 R_{b_1} 和人参皂苷 R_d 含量作为对应预测模型的 Y 变量, 结合正交信号校正去除 X 变量中与 Y 变量相关性较低的数据。采用 Kennard-stone 算法筛选不同预处理方法模型的训练集(2/3)与预测集(1/3); 训练集数据建立回归模型, 预测集数据对模型预测效果进行外部验证。

不同预处理 OSC-PLSR 模型采用 1/7 交叉验证(7-fold cross-validation)并以交叉验证均方根误差(root mean squares error of cross-validation, RMSECV)及相关系数 R_{cal}^2 为参数确

定最佳潜在变量数, 以最佳数量的潜在变量建立皂苷含量预测模型。预测集数据对模型进行外部验证, 以预测均方根误差(root mean square error of prediction, RMSEP)及相关系数 R_{pre}^2 为参数对模型预测能力进行评价。三七皂苷 R_1 、人参皂苷 R_{g_1} 、人参皂苷 R_{b_1} 和人参皂苷 R_d 的最佳含量预测模型及其参数见表 2, 样品皂苷含量预测值与 UPLC 检测值见图 3。结果显示, OSC-PLSR 模型对样品中四种皂苷含量预测效果较好, 预测值与检测值接近。

表 2 皂苷含量最佳预测模型参数

Table 2 The parameters of the best models of different saponins

皂苷种类	最佳模型	潜在变量数	训练集		预测集	
			R_{cal}^2	RMSECV	R_{pre}^2	RMSEP
三七皂苷 R_1	2D+SG5+OSC-PLSR	3	0.998 6	1.230 3	0.996 4	0.551 9
人参皂苷 R_{g_1}	2D+SG7+OSC-PLSR	4	0.999 7	2.149 1	0.999 2	0.955 0
人参皂苷 R_{b_1}	1D+SG5+OSC-PLSR	3	0.999 3	2.076 2	0.995 8	1.156 3
人参皂苷 R_d	1D+SG5+OSC-PLSR	3	0.999 0	0.902 6	0.996 6	0.509 5

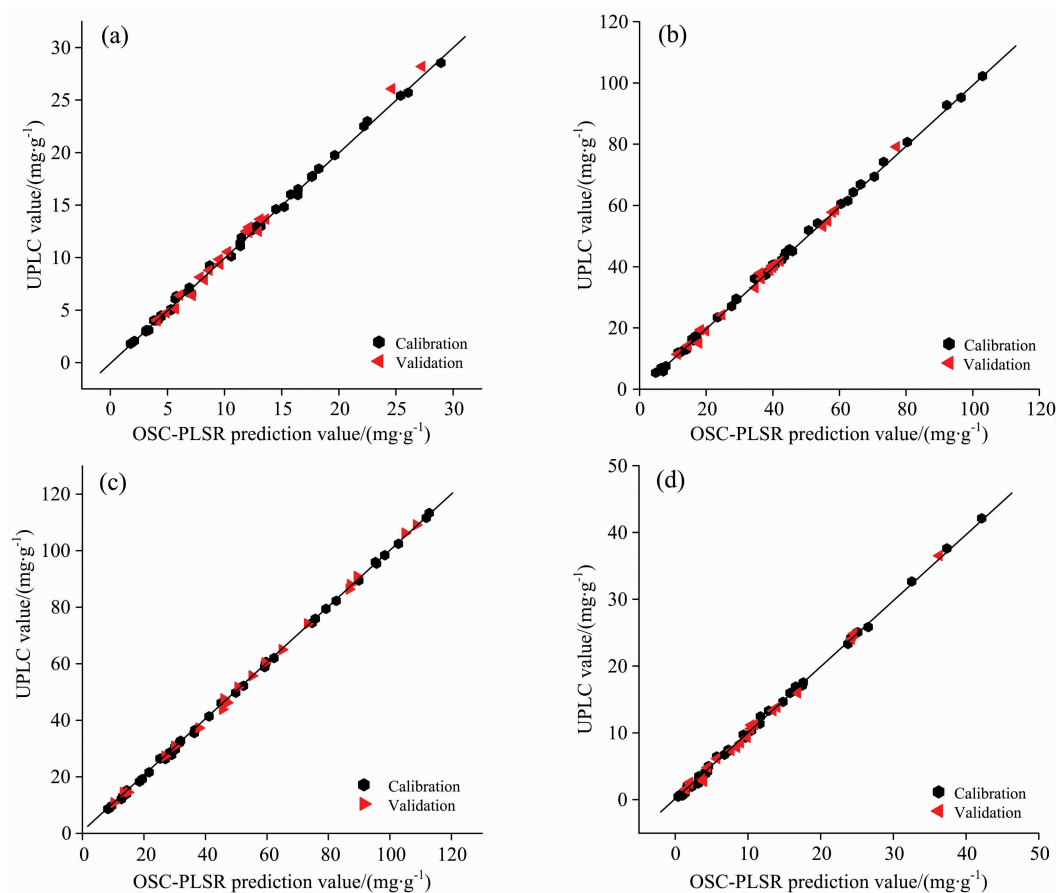


图 3 样品皂苷含量 UPLC 检测值与 OSC-PLSR 模型预测值

(a): 三七皂苷 R_1 ; (b): 人参皂苷 R_{g_1} ; (c): 人参皂苷 R_{b_1} ; (d): 人参皂苷 R_d

Fig. 3 The UPLC value and OSC-PLSR prediction value of samples

(a): Notoginsenoside R_1 ; (b): Ginsenoside R_{g_1} ; (c): Ginsenoside R_{b_1} ; (d): Ginsenoside R_d

3 结 论

利用 UPLC 对三七样品中三七皂苷 R_1 、人参皂苷 R_{g_1} 、人参皂苷 R_{b_1} 和人参皂苷 R_d 含量进行检测, 剪口样品中四种皂苷以及总皂苷含量均最高, 总皂苷含量为主根的 1.8 倍, 须根的 4.6 倍。二阶导数红外光谱结合 SVM 建立判别模型对三七主根、剪口和须根粉末进行鉴别, 预测集数据的判别正确率为 100%, 20 个预测样本均被正确鉴别。红外光

谱结合 OSC-PLSR 建立三七中四种皂苷含量预测模型, 考察 1D, 2D 以及 SG5, SG7, SG9, SG11 预处理对模型拟合能力及预测能力的影响, 并根据预测能力筛选最优模型。结果显示, 最优模型拟合能力及预测能力均较好, 预测值与 UPLC 检测值接近。红外光谱结合 SVM 判别能成功的对三七主根、剪口和须根粉末进行鉴别, 结合 OSC-PLSR 能够成功地对三七主根、剪口和须根中三七皂苷 R_1 、人参皂苷 R_{g_1} 、人参皂苷 R_{b_1} 和人参皂苷 R_d 含量进行预测, 可作为三七快速质量控制的方法, 为三七产业健康规范发展提供了基础。

References

- [1] Xiao C, Wu M, Chen Y, et al. *Phytochemical Analysis*, 2015, 26(1): 86.
- [2] Le T H V, Lee G J, Vu H K L, et al. *Chemical and Pharmaceutical Bulletin*, 2015, 63(11): 950.
- [3] Pan Y, Zhang J, Shen T, et al. *Biomedical Chromatography*, 2016, 30(2): 232.
- [4] Uzayisenga R, Ayeka P A, Wang Y. *Phytotherapy Research*, 2014, 28(4): 510.
- [5] Wang Z, Chen Y, Pan H, et al. *Chinese Herbal Medicines*, 2015, 7(2): 179.
- [6] Yang C, Wang J, Zhao Y, et al. *Journal of Ethnopharmacology*, 2010, 130(2): 231.
- [7] Uzayisenga R, Ayeka P A, Wang Y. *Phytotherapy Research*, 2014, 28(4): 510.
- [8] Sun S, Wang C Z, Tong R, et al. *Food Chemistry*, 2010, 118(2): 307.
- [9] Sun S, Qi L W, Du G J, et al. *Food Chemistry*, 2011, 125(4): 1299.
- [10] Xu L, Liu J, Liu N, et al. *Journal of Ethnopharmacology*, 2011, 137(1): 226.
- [11] Fan J S, Liu D N, Huang G, et al. *Journal of Ethnopharmacology*, 2012, 142(3): 732.
- [12] Zhou N, Tang Y, Keep R F, et al. *Phytomedicine*, 2014, 21(10): 1189.
- [13] Ng T B. *Journal of Pharmacy and Pharmacology*, 2006, 58(8): 1007.
- [14] Chinese Pharmacopoeia Commission(中华人民共和国药典委员会). *Pharmacopoeia of the People's Republic of China, Part One(中华人民共和国药典一部)*. Beijing: China Medical Science Press(北京: 中国医药科技出版社), 2015.
- [15] Wan J B, Li P, Yang R L, et al. *Journal of Liquid Chromatography & Related Technologies*, 2013, 36(3): 406.
- [16] Wang J R, Yau L F, Gao W N, et al. *Journal of Agricultural and Food Chemistry*, 2014, 62(36): 9024.
- [17] Jia X H, Wang C Q, Liu J H, et al. *Journal of Natural Medicines*, 2013, 67(2): 339.
- [18] Cheng C, Liu J, Zhang C, et al. *Applied Spectroscopy Reviews*, 2010, 45(2): 148.
- [19] Luo C, He T, Chun Z. *Journal of Molecular Structure*, 2013, 1037: 40.
- [20] Li B, Wei Y, Duan H, et al. *Vibrational Spectroscopy*, 2012, 62: 17.
- [21] Kyomugasho C, Christiaens S, Shpigelman A, et al. *Food Chemistry*, 2015, 176: 82.
- [22] Luo Q, Yun Y, Fan W, et al. *RSC Advances*, 2015, 5(7): 5046.
- [23] Li D, Jin Z, Zhou Q, et al. *Journal of Molecular Structure*, 2010, 974(1): 68.
- [24] Shen F, Yang D, Ying Y, et al. *Food and Bioprocess Technology*, 2012, 5(2): 786.
- [25] Chang C C, Lin C J. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2011, 2(3): 27.

Study of the Underground Parts Identification and Saponins Content Prediction of *Panax Notoginseng* Based on FTIR Combined with Chemometrics

LI Yun^{1,2,3}, ZHANG Ji^{1,2}, JIN Hang^{1,2}, WANG Yuan-zhong^{1,2*}, ZHANG Jin-yu^{1,2,3*}

1. Institute of Medicinal Plants, Yunnan Academy of Agricultural Sciences, Kunming 650200, China

2. Yunnan Technical Center for Quality of Chinese Materia Medica, Kunming 650200, China

3. College of Traditional Chinese Medicine, Yunnan University of Traditional Chinese Medicine, Kunming 650500, China

Abstract Phenomenon of adulterated traditional Chinese medicine (TCM) are still common in TCM market today. Unscrupulous traders used fibrous root powder pretending to be main root and rhizome powder of *Panax notoginseng*, and such behavior has serious influence on the quality and efficacy of *Panax notoginseng*. In this study, we have established a rapid method to discriminate the main root, rhizome and fibrous root powder and detect saponins content of *Panax notoginseng* in order to provide some research bases for rapid quality assessment of *Panax notoginseng*. A total of 60 Fourier transform infrared (FTIR) spectra of the main root, rhizome and fibrous root powder of *Panax notoginseng* were collected, and ultra-high performance liquid chromatography (UPLC) was used for measuring the content of notoginsenoside R₁, ginsenoside Rg₁, ginsenoside Rb₁ and ginsenoside Rd of samples. The origin data of identify model were processed by ordinate normalization and second derivative, and 2/3 of the 60 individuals were selected to form the calibration set by using Kennard-stone algorithm as well as the other 1/3 were used as validation set. Calibration set data were used to establish the discriminant model of support vector machine (SVM) and the cross-validation was used for screening optimal parameters *c* and *g*, and validation set data were used to verify the results of SVM discriminant model for external validation. The origin data used to predict saponins content were calculated by first (1D) and second derivative (2D), Savitsky-Golay smoothing with five, seven, nine, and eleven points. 2/3 of the 60 individuals were selected to form the calibration set and the rest were used as validation set. The orthogonal signal correction-partial least squares regression (OSC-PLSR) model was established by calibration set and the validation set was utilized to verify the results of the model for external validation. Results showed that, (1) with second derivative processing, the overlapped peak of FTIR spectra were efficiently separated and the resolution of the spectra has been improved. (2) The optimal parameters *c* and *g* of support vector machine calculated by cross-validation were 2.828 43 and $4.882\ 81 \times 10^{-4}$ respectively and the optimal accuracy rate of calibration set was 100%. (3) The parameter of support vector machine model was set as the optimal parameter and the accuracy rate of validation set was 100%, and all samples in validation set have been identified correctly. (4) The prediction content of greatest model of notoginsenoside R₁, ginsenoside Rg₁, ginsenoside Rb₁ and ginsenoside Rd was close to the content measured by UPLC. The result indicated that, FTIR combined with support vector machine could effectively identify the main root, rhizome and fibrous root powder of *Panax notoginseng*. OSC-PLSR could accurately predict the content of four saponins of *Panax notoginseng*. In summary, the FTIR spectroscopy could provide a rapid and effective method for the quality control of *Panax notoginseng*.

Keywords Fourier transform infrared spectroscopy; *Panax notoginseng*; Main root; Rhizome; Fibrous root; Powder identification; Saponins content prediction

(Received Apr. 3, 2016; accepted Oct. 19, 2016)

* Corresponding authors