

## 红外光谱法结合系统聚类和 SIMCA 模式识别法快速鉴别羌活种子

沈 亮<sup>1,2</sup>, 蒋舜媛<sup>3\*</sup>, 黄荣韶<sup>2</sup>, 周 毅<sup>3</sup>, 李良波<sup>2</sup>, 孙 辉<sup>4</sup>, 甘凤琼<sup>2</sup>

1. 中国医学科学院 北京协和医学院药用植物研究所, 北京 100193
2. 广西大学农学院, 广西 南宁 530005
3. 四川省中医药科学院, 四川 成都 610041
4. 四川大学 环境科学与工程系, 四川 成都 610065

**摘要:** 目的 采用傅里叶变换红外光谱法鉴别羌活种子。方法 在 4 000~400 cm<sup>-1</sup>内测定羌活种子光谱吸收峰, 应用系统聚类和SIMCA模式识别法鉴别羌活和宽叶羌活的种子。结果 运用系统聚类法, 当聚类距离为 15 时, 两种羌活可明显分为 2 类; 运用SIMCA模式识别法, 所建模型对羌活和宽叶羌活的种子识别率分别达 96%和 97%, 拒绝率均为 100%, 两者盲样检测准确率也在 90%以上, 所建模型可用于样品检测。结论 应用红外光谱法结合系统聚类和SIMCA模式识别法, 可以简单、快速、准确地鉴别羌活种子, 该方法为鉴别羌活种子的来源及真伪提供了一种新的技术手段。

**关键词:** 羌活; 种子鉴定; 红外光谱; 系统聚类; SIMCA

中图分类号: R282.5 文献标志码: A 文章编号: 0253 - 2670(2011)10 - 2114 - 05

## Rapid seed identification of *Notopterygii Rhizoma et Radix* by SIMCA based on FTIR

SHEN Liang<sup>1,2</sup>, JIANG Shun-yuan<sup>3</sup>, HUANG Rong-shao<sup>2</sup>, ZHOU Yi<sup>3</sup>, LI Liang-bo<sup>2</sup>, SUN Hui<sup>4</sup>, GAN Feng-qiong<sup>2</sup>

1. Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China
2. College of Agriculture, Guangxi University, Nanning 530005, China
3. Sichuan Academy of Traditional Chinese Medicine Sciences, Chengdu 610041, China
4. Department of Environmental Science and Engineering, Sichuan University, Chengdu 610065, China

**Abstract: Objective** To identify the seeds of *Notopterygii Rhizoma et Radix* by Fourier transform infrared spectrometry (FTIR). **Methods** Based on the fingerprint infrared spectrum from 4 000 to 400 cm<sup>-1</sup>, hierarchical cluster and Soft Independent Modeling of Class Analogy (SIMCA) analyses were developed to identify the seeds of *Notopterygium incisum* and *N. franchetii*. **Results** The seeds could be divided into two categories with 15 Euclidean distance by using hierarchical cluster analysis. Using SIMCA analysis, the recognition rate of *N. incisum* and *N. franchetii* reached to 96% and 97%, respectively, and the rejection rate reached to 100%. When testing with the blind sample which the authors picked out from the chosen samples, the accuracy rate reached over 90%. The established model could be used for sample testing. **Conclusion** Both clustering analysis and FTIR combined with SIMCA provide an effective way to evaluate the seeds of *Notopterygii Rhizoma et Radix* rapidly and accurately. Therefore, this method provides a new technique for evaluating the source and authenticity of the seeds of *Notopterygii Rhizoma et Radix*.

**Key words:** *Notopterygii Rhizoma et Radix*; seed identification; FTIR; hierarchical cluster; Soft Independent Modeling of Class Analogy (SIMCA)

羌活*Notopterygii Rhizoma et Radix*为传统常用大宗中药材, 具有散寒、祛风、除湿、止痛的功效, 主治风寒感冒、骨节酸痛、痈疽疮毒等症<sup>[1]</sup>, 入方成药达 280 余种, 全国关联的制药企业超过 700 家。《中国药典》2010 年版规定其药材为伞形科羌活属

植物羌活*Notopterygium incisum* Ting ex H. T. Chang 和宽叶羌活*N. franchetii* H. de Boisseu的干燥根茎及根, 为生长于青藏高原东南部高海拔山地的中国特有属植物<sup>[2]</sup>。市售的羌活药材在商品规格和利用价值上历来都是羌活优于宽叶羌活, 近十余年高强

收稿日期: 2011-01-15

基金项目: 国家科技部“重大新药创制”科技重大专项(民口)课题; “中药材种子种苗和种植(养殖)标准平台”课题(2009ZX09308-002); 广西科学研究与技术开发计划项目(桂科攻 0992003A-20)资助

作者简介: 沈 亮(1986—), 男, 山东枣庄人, 在读博士, 主要从事药用植物栽培育种及保护研究。

Tel: 13667882455 E-mail: shenliang08@126.com

\*通讯作者 蒋舜媛 Tel: 18982072450 E-mail: jsy007@vip.sina.com

度的采挖已致其资源严重濒危。羌活种子发芽育苗技术的突破为实现羌活的人工栽培和可持续利用提供了技术支撑,但两个种的种子在形态上难以分辨,种子前处理及育苗时间长达 9 个月以上<sup>[3]</sup>,使形态法、幼苗法等常用的种子鉴定方法难以实施,采用电泳法等其他生物技术手段也因实验材料缺乏稳定性和准确性,羌活种子真伪的快速准确鉴别成为困扰羌活种子质量控制和人工种植的一个关键问题。

红外光谱结合系统聚类分析及SIMCA模式识别法在药材鉴定方面已有相关报道<sup>[4]</sup>,该方法具有快速、准确、经济等特点。本实验尝试运用该项技术对羌活种子进行鉴别研究,为快速、准确鉴定不同来源的羌活属植物种子探索一种新的分析方法。

## 1 材料与方法

### 1.1 材料

羌活和宽叶羌活种子共 17 份,由四川省中医药科学院蒋舜媛博士提供并鉴定(表 1)。其中 1~6 号为 6 份不同来源的宽叶羌活种子(定为 A 种子),9~16 号为 8 份不同来源的羌活种子(定为 B 种子),7、8 和 17 号为待测盲样。将各样品清洗后,放入烘箱,40 °C 低温烘干至恒质量,去种皮,粉碎,过 200 目筛,将样品包装好后放于恒温烘箱或干燥器内保存,待测定。

Nicolet 5700傅氏转换红外线光谱分析仪,检测器为美国尼高力公司生产;FW—4A型压片机(天津光学仪器厂);分析天平XSI05DU(梅特勒-托利多仪器公司);溴化钾(天津光复精细化工研究所,色谱纯)。

### 1.2 方法

精确称取 1.5 mg样品,将其与完全干燥的 200.0 mg溴化钾混合研磨均匀,压制成厚度约 1 mm的透明锭片。扫描范围 4 000~400 cm<sup>-1</sup>,光谱分辨率 4 cm<sup>-1</sup>,累积次数 32 次。每个产地的羌活种子各制备 20 个溴化钾压片,每个待测样品随机扫描 3 次,收集所得光谱图。

原始光谱数据经过 Omnic 8.0 软件进行基线校正,13 点平滑,平均图谱后,将光谱数据导入 Unscrambler 10.0 软件中,对谱图进行标准归一化处理,以消除样本称量不均带来的光谱峰吸光度的差异,保证谱图规格的统一性和可比性。系统聚类分析使用 SPSS 11.5 软件,聚类采用内平方距离法

表 1 两种羌活种子样品的产地及采收时间

Table 1 Habitat and collected time of *N.franchetii* and *N. incisum*

序号	来源	采收时间	拉丁名
1	四川小金县家种	2010-09	<i>N. franchetii</i>
2	四川壤塘县家种	2010-09	<i>N. franchetii</i>
3	四川小金县两河乡野生移栽	2010-09	<i>N. franchetii</i>
4	四川甘孜县野生移栽	2010-09	<i>N. franchetii</i>
5	四川壤塘县阿斯玛乡野生	2010-09	<i>N. franchetii</i>
6	甘肃渭源县家种	2010-08	<i>N. franchetii</i>
7	待测盲样 1	2010-09	
8	待测盲样 2	2010-09	
9	四川小金县家种	2010-09	<i>N. incisum</i>
10	四川壤塘县家种	2010-09	<i>N. incisum</i>
11	四川甘孜县泥柯乡野生移栽	2010-09	<i>N. incisum</i>
12	四川小金县梦笔山野生	2010-09	<i>N. incisum</i>
13	四川壤塘县阿斯玛乡野生	2010-09	<i>N. incisum</i>
14	四川壤塘县野生	2010-09	<i>N. incisum</i>
15	四川壤塘县家种 1	2010-09	<i>N. incisum</i>
16	四川壤塘县家种 2	2010-09	<i>N. incisum</i>
17	待测盲样 3	2010-09	

(Word法),采用欧氏距离法<sup>[5]</sup>计算距离,SIMCA建模预测分析在Unscrambler 10.0 软件中进行。

## 2 结果与分析

### 2.1 羌活与宽叶羌活种子的红外光谱分析比较

红外光谱的特点是能够解析出样品化学成分的官能团,确定其基本骨架,从而判断出样品所含化学成分<sup>[6]</sup>。图1为羌活和宽叶羌活种子的原始红外光谱图,参考相关文献<sup>[7-9]</sup>数据对其光谱峰进行指认和归属可知,3 305 cm<sup>-1</sup>波数附近是多糖、苷类化合物中的羟基O-H伸缩振动峰;2 925、2 854 cm<sup>-1</sup>附近为蛋白质和油脂类甲基、亚甲基C-H伸缩振动峰;1 741 cm<sup>-1</sup>附近是羧酸类或酯类物质中C=O的伸缩振动峰;这部分主要体现了羌活种子中蛋白质、挥发油、香豆素等脂类物质的光谱表征,而从 2 852、1 741 cm<sup>-1</sup>的吸收峰可以推断其种子中含有大量的油脂类物质。1 660 cm<sup>-1</sup>附近为C=O或C=C双键的伸缩振动区,可能为苯环及双键的骨架振动,此外还包括由蛋白质C=O伸缩振动产生的酰胺I带;而在 1 616 cm<sup>-1</sup>波数附近出现了氨基酸-NH<sub>2</sub>的剪式振动酰胺II带特征吸收峰;1 518、1 456 cm<sup>-1</sup>附近出现了芳环的骨架振动吸收峰;这部分主要体现了羌活种子含有氨基酸、香豆素类物质的光谱

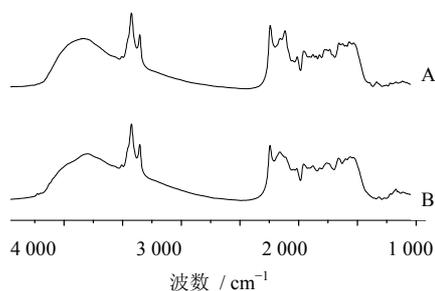


图 1 宽叶羌活 (A) 和羌活 (B) 种子的红外原始光谱图  
Fig. 1 FTIR spectra of *N. franchetii* (A) and *N. incisum* (B) seeds

表征。1 157、1 070  $\text{cm}^{-1}$  重叠而成的宽峰，归属为多糖、苷类等碳水化合物的C-O弯曲振动，而 857  $\text{cm}^{-1}$  是C-H的弯曲振动峰，这部分主要体现羌活种子含有多糖类物质的特征吸收峰。虽然两种羌活种子的红外光谱图较为相似，但是某些吸收峰的峰位、峰形及峰强的差异还是为鉴别两种羌活种子提供了理论依据。如羌活种子在 1 741、1 616  $\text{cm}^{-1}$  处的吸收峰峰强明显弱于宽叶羌活种子；在 1 660  $\text{cm}^{-1}$  处羌活种子的酰胺I带吸收峰又明显强于宽叶羌活种子；此外二者在 1 518、1 456  $\text{cm}^{-1}$  附近的芳环振动吸收也存在明显的差异，在 1 279、1 261、857、673  $\text{cm}^{-1}$  附近也有不同程度的吸收峰，通过原始光谱可以有效地对两种羌活种子进行鉴别及化学成分分析。

### 2.2 羌活和宽叶羌活种子的系统聚类分析

对羌活样品进行归一化处理，将 1 800~400  $\text{cm}^{-1}$  波段的吸光度进行系统聚类分析，结果如图 2 所示。当聚类距离为 15 时，羌活样品明显的分为 2 类，而不出现交叉。其中 1~8 号聚为一类，9~17 号聚为另一类，它们分别代表宽叶羌活和羌活种子的光谱信息，系统聚类方法可以直观地反映出样品之间的亲缘关系和相似程度。待测盲样 7、8 号与宽叶羌活类群聚到一起，而待测盲样 17 与羌活类群聚到一起，这与本课题组采用其他常规方法及分子鉴定方法所得结果一致<sup>[10]</sup>。另外系统聚类法还可以推断出样品亲缘关系的远近，由图 2 可知两种羌活的差异较大，当聚类距离接近 25 时，两者才出现较为相似的光谱信息而聚为一类。

### 2.3 羌活和宽叶羌活种子的 SIMCA 模式识别分析

为了更加直观地区分两种羌活种子并验证以上系统聚类分析结果，本研究选用SIMCA模式识别法<sup>[10]</sup>对羌活种子做进一步研究，以实现两种羌活

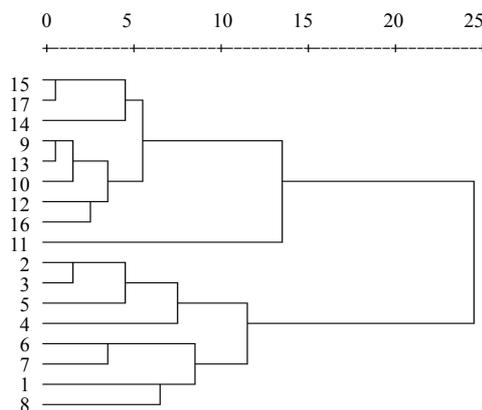


图 2 17 个不同产地的羌活种子红外光谱系统聚类图  
Fig. 2 Hierarchical cluster of FTIR spectra recorded of *N. franchetii* and *N. incisum* seeds from 17 habitats

种子的智能识别。

2.3.1 羌活和宽叶羌活种子模型的建立 本研究经过数据预处理后，选择羌活种子光谱信息丰富的 1 800~400  $\text{cm}^{-1}$  波段进行SIMCA模式识别分析，以宽叶羌活种子的样品 (1~6 号) 为纵坐标，羌活种子的样品 (9~15 号) 为横坐标，经过处理得到羌活和宽叶羌活种子的SIMCA聚类分析图。由图 3 可见，羌活和宽叶羌活的聚类结果比较理想，两者各自聚到一起，没有出现重叠。

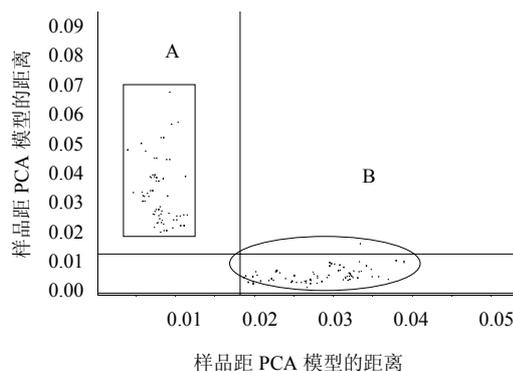


图 3 宽叶羌活 (A) 和羌活 (B) 的 SIMCA 聚类分析图  
Fig. 3 SIMCA classification of *N. franchetii* (A) and *N. incisum* (B) seeds

2.3.2 模型的诊断 模型建好后，需要对其进行进一步的模型诊断，即求出每种羌活种子的识别率和拒绝率。所谓识别率就是考察某类样品有多少落在该类模型的范围中，而拒绝率就是考察某类样品模型对于不属于该类的其他样品的拒绝程度，即其他样品是否都落在该类模型的范围外。当这两个值都

为 100%时,表明这两种样品之间差异很大,可以完全分开。经过多次重复试验及数据处理,所建模型对羌活和宽叶羌活种子的识别率分别达到了 96%和 97%,拒绝率均为 100%,说明所建模型较为成功,准确性较高。

**2.3.3 模型的检验** 为检验所建模型是否具有较全面的样品代表性,本研究采用测试集样本对模型进行了测试。测试集是由 40 个非建模集样品构成,其中宽叶羌活为 2 个产地(为待测种子 7、8 号)的 20 个种子样品,羌活为 2 个产地(16 号和待测种子 17 号)的 20 个种子样品。经过预测得知,宽叶羌活种子判别准确率为 90%,而羌活种子的判别准确率为 95%,测试正确率较高,说明建模样品数量合理,所建模型具有一定的代表性,可以用于两种羌活及未知羌活种子的鉴别研究。

### 3 讨论

通过对原始光谱图分析发现,羌活和宽叶羌活种子的化学成分存在一定的差异,但尚不能从原始图谱上进行直接区分。进一步使用系统聚类法进行分析,两种羌活种子样品各自聚到一起,而不出现交叉,当聚类距离接近 25 时,二者聚为一类,较好地实现了不同种羌活种子的区分。采用 SIMCA 模式识别法对羌活种子和宽叶羌活种子的识别率分别达到了 96%和 97%,拒绝率均为 100%。用待测盲样对所建模型检验得出,二者的聚类判别准确率分别为 95%和 90%,可见该方法能够用于对未知样品的检测。本研究表明,借助羌活种子的红外光谱信息,从整体上可以解析出其种子内含有的化学成分,在此基础上采用系统聚类和 SIMCA 模式识别法,可以实现对外观相似的羌活和宽叶羌活种子的快速鉴别。

目前有关种子鉴定的方法较多,对于区别性特征很显著的种子,可采用传统的外观形态鉴定法,外观相似度较高的种子则需要依赖长期积累的鉴别经验为基础,且难保证高的准确性<sup>[11]</sup>。幼苗观察法可确保鉴定结果准确,对于发芽周期短的植物种子适用<sup>[11]</sup>,但对发芽周期较长的种子,则需花费大量的人力、物力、财力及时间,尤其对从种子前处理到发芽出苗,并长到形态可鉴别的幼苗所需时间长达 9 个月以上的羌活种子来说,幼苗法可作为复核的技术手段,但对种子质量控制和人工种植生产则缺乏实用价值。近年在种子鉴定方面发展起来的电泳法、ISSR 标记、rDNA 标记等方法,对某些类群

的种子能实现快速鉴定,准确性也较高<sup>[12-14]</sup>,但 these 方法一方面对仪器及人员的技能要求较高,花费较高<sup>[11]</sup>,另一方面对实验材料的要求对某些种类来说难以满足。

红外光谱鉴定法只需少量粉碎样品,使用光谱仪收集谱图进行数据处理即可得出鉴定结果,具有方便、快速、准确、灵敏、经济等特点。其不足之处在于大型红外光谱仪价格较高,同时由于实验用材较少,检测灵敏性较高,取样必须有较好的代表性。目前已有运用红外光谱法对砂仁种子及紫苏子的真伪进行快速鉴别研究的相关报道<sup>[15-16]</sup>,而红外光谱法结合化学计量学在种子鉴定领域更具优势,如系统聚类法、SIMCA 模式识别法具有直观、准确、可验证性等优点,可以快速客观的区分形态和解剖手段较难鉴别的不同来源及真伪的种子<sup>[17-19]</sup>。可见,在种子鉴定领域里,红外光谱鉴定法与其他鉴定方法相比,具有操作简单、鉴定快速、准确,花费较少等优点而具有较大的应用价值。随着科学技术的发展,红外光谱鉴定法可以在种子鉴定领域发挥更大的作用。

### 参考文献

- [1] 中国药典 [S]. 一部. 2010.
- [2] 王幼平, 溥发鼎, 王萍莉, 等. 中国特有属—羌活属的系统分类研究 [J]. 云南植物研究, 1996, 18(4): 424-430.
- [3] 史 静, 蒋舜媛, 马小军, 等. 羌活种子发芽及实生苗生长发育的研究 [J]. 中国中药杂志 2007, 32(18): 1841-1844.
- [4] 徐永群, 黄 昊, 周 群, 等. 红外指纹图谱和聚类分析法在赤芍产地分类鉴别中的应用 [J]. 分析化学, 2003, 31(1): 5-9.
- [5] 孔德鑫, 黄庶识, 黄荣韶, 等. 基于双指标分析法和聚类分析法的鸡骨草红外指纹图谱比较研究 [J]. 光谱学与光谱分析, 2010, 30(1): 45-49.
- [6] 翁诗甫. 傅里叶变换红外光谱分析 [M]. 北京: 化学工业出版社, 2010.
- [7] 陈 君, 孙素琴, 徐 荣, 等. 应用红外光谱法鉴别肉苁蓉及其混淆品草苁蓉和锁阳 [J]. 光谱学与光谱分析, 2009, 29(6): 1502-1507.
- [8] 孙素琴, 周 群, 秦 竹. 中药二维相关红外光谱鉴定图集 [M]. 北京: 化学工业出版社, 2003.
- [9] 黄冬兰, 陈小康, 徐永群, 等. 纹党参与白条党参红外光谱的 SIMCA 聚类鉴别方法研究 [J]. 分析测试学报, 2009, 28(12): 1440-1443.
- [10] Xu K J, Jiang S Y, Zhou Y, *et al.* Discrimination of the

- seeds of *Notopterygium incisum* and *Notopterygium franchetii* by validated HPLC-DAD-ESI-MS method and principal component analysis [J]. *J Pharm Biomed Anal*, 2011, 56: 1089-1093.
- [11] 白雁, 鲍红娟, 王东, 等. 红外光谱和聚类分析法在药用菊花产域分类鉴别中的应用 [J]. *中药材*, 2006, 29(7): 663-665.
- [12] 王芳. 种子纯度鉴定方法及其评述 [J]. *中国种业*, 2008, 10: 62-63.
- [13] 高燕会, 李慧慧, 朱玉球, 等. 基于 ISSR 的栝楼遗传多样性分析 [J]. *中草药*, 2011, 42(2): 363-366.
- [14] 蒲汉丽, 李乃坚, 黄爱兴, 等. RAPD 技术鉴定番茄一代杂种粤星种子纯度的研究 [J]. *种子*, 2000, 11(54): 14-15.
- [15] 郝剑瑾, 程舟, 梁洪卉, 等. 基于 rDNA ITS 序列探讨我国冬虫夏草的遗传分化及分布格局 [J]. *中草药*, 2009, 40(1): 112-116.
- [16] 程存归, 阮永明, 李冰岚. 傅里叶变换红外光谱法应用于中药砂仁真伪鉴别的研究 [J]. *光谱学与光谱分析*, 2004, 24(11): 1355-1358.
- [17] 程存归. FTIR 直接鉴定紫苏子及其伪品的研究 [J]. *光谱学与光谱分析*, 2003, 23(2): 282-284.
- [18] 姜科声, 孔黎春, 余鹏, 等. 不同花生品种的红外光谱分析 [J]. *光谱实验室*, 2008, 25(4): 550-553.
- [19] 周群, 孙素琴, 梁曦云. 枸杞产地的红外指纹图谱与聚类分析法研究 [J]. *光谱学与光谱分析*, 2003, 23(3): 509-511.

欢迎订阅

### Chinese Herbal Medicines (CHM, 中草药英文版)

我国第一份中药专业的英文期刊——*Chinese Herbal Medicines* (CHM, 中草药英文版) 经国家新闻出版总署批准, 已于 2009 年 10 月正式创刊, 国内统一连续出版号为: CN12—1410/R。

CHM 由天津药物研究院和中国医学科学院药用植物研究所主办, 天津中草药杂志社出版。中国工程院院士、中国医学科学院药用植物研究所名誉所长肖培根教授担任主编; 中国工程院院士、天津药物研究院刘昌孝研究员, 天津药物研究院院长汤立达研究员, 中国医学科学院药用植物研究所所长陈士林研究员共同担任副主编; 天津药物研究院医药信息中心主任、《中草药》杂志执行主编常青研究员担任编辑部主任。

**办刊宗旨** 以高起点、国际化为特点, 继承和发扬祖国医药学遗产, 报道和反映中草药研究最新进展, 宣扬我国中草药的传统特色, 加强与世界各国在传统药物研究的经验交流, 在中医和西医、传统与现代、东方与西方之间架起一座理解和沟通的桥梁, 促进中药现代化、国际化。

**主要栏目** 综述与述评、论著、快报、简报、文摘、信息和国际动态、人物介绍、来信、书评等栏目。

**读者对象** 国内外从事中医药研究、管理、监督、检验和临床的专业技术人员。

CHM 邀请相关领域的院士和国内外知名专家加盟, 组建一支国际化、高水平、精干的编委会队伍 (第一届编辑委员会由 49 位专家组成, 其中院士 10 名, 国际编委 19 名)。吸引国内外高质量的稿件, 提高期刊的学术质量; 坚持按照国际标准编排, 加强刊物规范化和标准化, 充分利用计算机、网络技术和英语, 加强与国际知名科技期刊的交流合作; 充分发挥中医药特色, 争取在较短时间内进入国际最著名的检索系统——美国科学引文索引 (SCI), 把 CHM 办成国际知名期刊之一。

欢迎广大作者踊跃投稿! 欢迎广大读者积极订阅! 自办发行, 直接与编辑部订阅!

本刊已正式开通网上在线投稿系统。欢迎投稿、欢迎订阅! 网址: [www.tiprpress.com](http://www.tiprpress.com)

### Chinese Herbal Medicines (CHM, 中草药英文版) 编辑部

天津编辑部

地址: 天津市南开区鞍山西道 308 号

邮编: 300193

E-mail: [chm@tiprpress.com](mailto:chm@tiprpress.com)

Tel: (022)27474913

Fax: (022)23006821

开户银行: 兴业银行天津南开支行

北京编辑部

地址: 北京市海淀区马连洼北路 151 号

邮编: 100193

E-mail: [bjchm@tiprpress.com](mailto:bjchm@tiprpress.com)

Tel: (010)57833035

Fax: (010)57833035

账号: 44114010010081504

户名: 天津中草药杂志社