



# 基于 LC-MS 分析的中药体内成分辨识技术

闫广利, 韩莹, 王喜军\*

(黑龙江中医药大学, 黑龙江 哈尔滨 150040)

**[摘要]** 中药血清药物化学是阐明中药药效物质基础直接而有效的方法,然而中药化学成分的复杂性和体内过程中成分间的相互作用决定了中药体内成分分析的艰巨性。液相色谱-质谱联用技术为中药体内成分的全息表征提供了有效的分析手段,然而由于生物样品中内源性物质的严重干扰难以在采集的色谱图中直观地辨识出中药体内成分,迫切需要引入专属的成分辨识技术,以避免对中药体内成分的人为漏检。该文结合黄连在大鼠血中的移行成分分析,介绍用于质谱数据处理的质量短缺过滤技术、Metabolynx 软件以及主成分分析、偏最小二乘判别分析、正交偏最小二乘判别分析等模式识别方法在中药体内成分辨识中的应用。

**[关键词]** 中药血清药物化学; 体内成分分析; 质量短缺过滤; Metabolynx; 模式识别

中药是在传统中医理论指导下应用的天然药物,包括中药材、中药材饮片和中成药。中药多来源于植物,其效应相关的物质基础多为植物的次生代谢产物,其数量往往达百种以上,含量也横跨很宽的浓度范围。基于中药体内直接作用物质的形成过程及中药有效成分体内过程研究而提出的中药血清药物化学认为,中药所含成分虽然复杂,但只有被吸收入血的成分才能发挥作用,这些成分包括中药所含成分的原型、代谢产物及生理活性物质;通过分析中药口服给药后血液中的成分,确定中药的体内直接作用物质,是快速、准确地研究确定中药药效物质基础的有效途径<sup>[1]</sup>。

中药体内成分研究最初是采用传统提取分离方法从生物样品中获得单体化合物,进而进行结构鉴定,或应用高效液相色谱-紫外光谱的指纹图谱检测方法,结果只能确定主要的或部分体内直接作用物质<sup>[2-4]</sup>。随着分析技术的发展,特别是超高效液相色谱串联质谱技术的应用,已经可以实现生物体内微量中药成分的在线分离和鉴定<sup>[5-6]</sup>,对于全息表征中药体内成分提供了有效的分析手段。然而,随之也带来了新的问题,即血液、尿液等生物样品所含有的

内源性物质同时也获得了良好的检测,而且其质谱响应值通常高于中药体内成分,从而在采集的色谱图中直观上难以发现中药体内成分,结果会人为造成一些成分的漏检。因此,在直观比较分析的基础上,需要引入一些有效的数据处理技术,从而比较完整地表征分析中药的体内成分。

本文以黄连含药血清 UPLC-Q-TOF-MS 图谱的处理为例,介绍用于质谱数据处理的质量短缺过滤(mass defect filter, MDF)技术、Metabolynx 软件以及主成分分析(principal component analysis, PCA)、偏最小二乘判别分析(partial least squared discriminant analysis, PLS-DA)、正交偏最小二乘判别分析(orthogonal partial least squared discriminant analysis, OPLS-DA)等模式识别方法在中药体内成分辨识中的应用。

## 1 质量短缺过滤技术

质量短缺(mass defect)是指一个化合物或离子的精确质量与整数质量之间的差值,例如葡萄糖的精确相对分子质量为 180.063 4,那么它的质量短缺值为 0.063 4。通过设定较窄的质量短缺范围,可以将背景化合物信号过滤掉,从而便于目标化合物的识别,这被称为质量短缺过滤(mass defect filter, MDF)技术<sup>[7]</sup>。该技术是随着高分辨质谱的发展而被提出的,最初被应用于单体化合物的代谢产物检测。通常的 I 相和 II 相代谢产物的质量短缺值都能落在相对于母化合物质量短缺值的  $\pm 50$  mDa,如羟基化产物质量短缺值变化 -5 mDa、去氢化产物为 -16 mDa、去甲基化产物为 -23 mDa、葡萄糖醛酸

**[稿件编号]** 20111116005

**[基金项目]** 国家自然科学基金项目(81173500);国家“重大新药创制”科技重大专项(2009ZX09502-005)

**[通信作者]** \*王喜军,博士,教授,主要从事中药血清药物化学及方剂药物代谢组学研究, Tel/Fax: (0451) 82110818, E-mail: xijun@sina.com

**[作者简介]** 闫广利,博士,副研究员

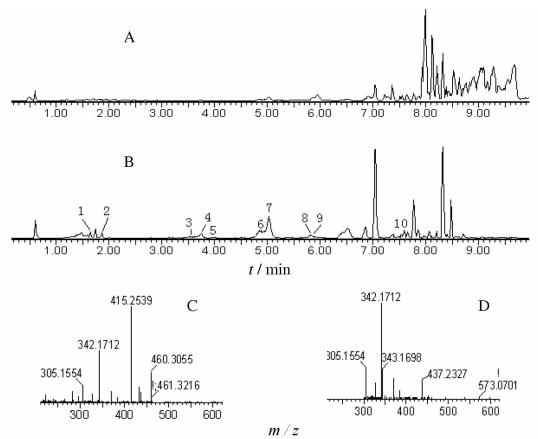


昔化产物为 + 32 mDa、硫酸酯化产物为 - 43 mDa。因此,通过设置较窄的质量短缺过滤范围(如  $\pm 50$  mDa),可以从含药生物样品的全扫描高分辨 LC-MS 图中过滤掉质量短缺值超出设定范围的离子,而保留质量短缺值在设定范围内的离子,从而简化 LC-MS 图和相应的质谱图,极易辨识代谢产物的色谱峰和相应质谱图中的分子离子峰<sup>[8]</sup>。目前,该技术已作为一个参数设置被整合到多个 LC-MS 数据处理软件中,例如下面将要介绍的 Metabolynx 软件。

中药含有的化学成分通常是其来源植物的次生代谢产物,它们在生物合成上可来源于几个途径,如莽草酸途径、乙酸-丙二酸途径、甲戊二羟酸途径等,因此在结构上通常具有相似的结构母核。通过对中药所含的同类化学成分的质量短缺值进行总结,并根据可能的体内代谢途径,可设置适当的质量短缺过滤范围,对含药生物样品的全扫描 LC-MS 图进行过滤,从而使该类化合物及其代谢产物易于辨识。例如,黄连所含的主要生物碱类成分为小檗碱、药根碱、巴马汀、小檗红碱、四氢巴马汀、非洲防己碱、黄连碱、木兰碱,总结它们的质量短缺值范围为 0.09 ~ 0.19 Da,进而根据其可能的代谢途径如羟基化、甲基化、硫酸化、葡萄糖醛酸化等,设定质量过滤范围为 300 ~ 600 Da,质量短缺过滤范围为 0.04 ~ 0.23 Da,过滤后的色谱图中黄连生物碱类成分及其代谢物峰清晰可见,相应的质谱图中其分子离子峰也明显暴露出来,见图 1,初步鉴定结果见表 1。

## 2 Metabolynx 软件

Metabolynx 是由 Waters 公司开发的高分辨率 LC-MS 数据处理软件,它通过对给药样品与空白样品的全扫描 LC-MS 数据文档的自动比较,根据设置的化合物代谢途径和适当的过滤参数,来提取化合物。这个软件是为了检测和鉴定单体化合物的代谢产物而设计的,需要预先设定母化合物,并通过对母化合物代谢产物的预测来设定代谢途径,即设定的相对于母化合物的精确质量偏差(根据代谢所引起的母化合物分子式变化进行计算所得),从而提取目标化合物(代谢产物)信息;同时 Metabolynx 能根据给药样品和空白样品 LC-MS 数据文档的比较,提供给药样品中存在而空白样品中不存在的非目标化合物供筛选,有助于实现相关代谢产物的全息检测<sup>[9]</sup>。



A. 过滤前的 UPLC-Q-TOF-MS 图; B. 过滤后的 UPLC-Q-TOF-MS 图, 峰号与表 1 中一致; C. 过滤前峰 1 ( $m/z$  342) 的质谱图; D. 过滤后峰 1 ( $m/z$  342) 的质谱图。

图 1 大鼠口服给予黄连提取物后血清 UPLC-Q-TOF-MS 色谱质谱图质量短缺过滤前后的比较

Fig. 1 Comparison of MDF-processed and unprocessed UPLC-Q-TOF-MS chromatogram and the corresponding MS spectrum of the rat serum after oral administration of Coptidis Rhizoma extract

表 1 口服给予黄连提取物后大鼠血中移行成分的 UPLC-Q-TOF-MS 初步分析

Table 1 The primary result of analysis of the constituents absorbed into the blood of Coptidis Rhizoma extract administrated orally to rat by UPLC-Q-TOF-MS

No.	$t_R$ /min	测定 质量	计算 质量	偏差 / $\times 10^6$	分子式	化合物
1	1.62	342.170 0	342.170 5	-1.5	$C_{20}H_{24}NO_4$	-
2	1.85	342.167 5	342.170 5	-8.8	$C_{20}H_{24}NO_4$	木兰碱
3	3.71	336.122 5	336.123 6	-4.2	$C_{20}H_{18}NO_4$	表小檗碱
4	3.72	320.093 5	320.092 3	3.7	$C_{19}H_{14}NO_4$	黄连碱
5	3.93	338.138 9	338.139 2	-0.9	$C_{20}H_{20}NO_4$	药根碱
6	4.95	352.155 1	352.154 9	0.6	$C_{21}H_{22}NO_4$	巴马汀
7	4.99	336.121 3	336.123 6	-6.8	$C_{20}H_{18}NO_4$	小檗碱
8	5.82	382.092 3	382.092 7	-1.0	$C_{20}H_{16}NO_7$	-
9	5.83	364.081 7	364.082 1	-1.1	$C_{20}H_{14}NO_6$	-
10	7.68	352.117 9	352.118 5	-1.7	$C_{20}H_{18}NO_5$	-

Metabolynx 通过 2 种方式提取离子,一种是靶向性的,要求设置特定的代谢途径,根据设定的代谢途径提取目标化合物(expected metabolites);另一种是非靶向性的,提取在给药样品中存在而空白样品中不存在的非目标化合物(unexpected metabolites),同时通过设置一些过滤参



数,其中包括上述介绍的质量短缺过滤,可以缩小提取的非目标化合物的数量,从而方便进一步

的手动筛选<sup>[10]</sup>。这 2 种提取结果在 Metabolynx 浏览器中被给出,见图 2。

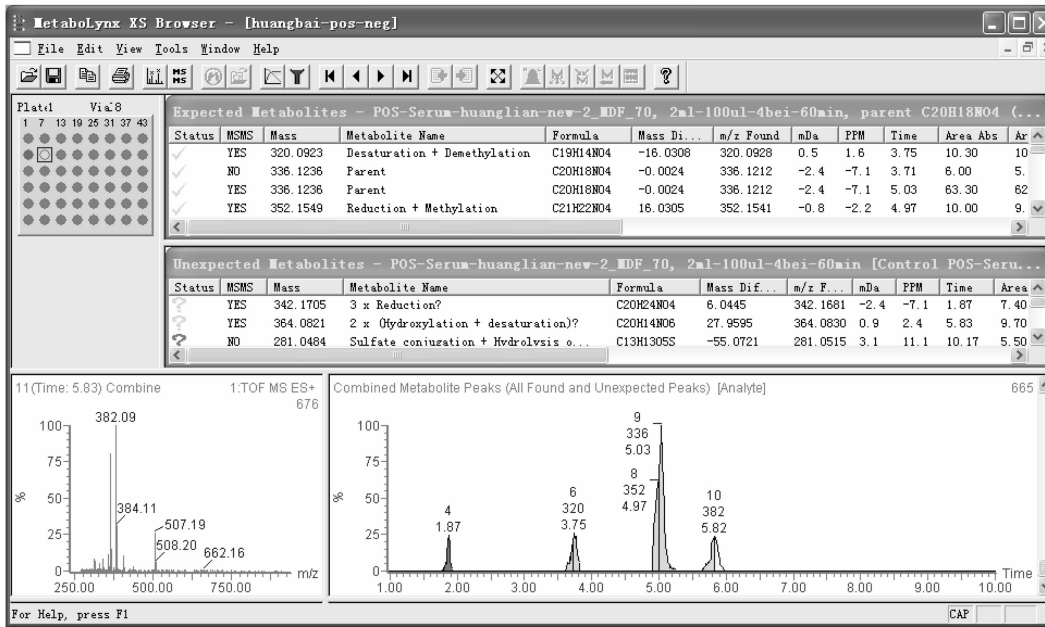


图 2 展示大鼠口服给予黄连提取物后血清 UPLC-Q-TOF-MS 图处理结果的 Metabolynx 浏览器

Fig. 2 The Metabolynx browser showing the processing result of UPLC-Q-TOF-MS chromatogram of the rat serum after oral administration of Coptidis Rhizoma extract

Metabolynx 同样可用于中药多组分代谢产物的辨识。如前所述,虽然中药化学成分复杂,但是根据其结构母核可以将其分成几个化合物类型。对于每个结构类型,可选择一个或几个母化合物,设置适当的代谢途径和过滤参数,可从中药或其含药生物样品的全扫描质谱图中提取特定类型的中药成分<sup>[11-12]</sup>。例如,利用 Metabolynx 软件对大鼠口服给予黄连后血清的 UPLC-Q-TOF-MS 图进行处理,设置小檗碱(分子式  $C_{20}H_{18}NO_4$ )为母化合物(Mass A),设定小檗碱类成分的已知代谢途径和  $\pm 50$  mDa 的质量短缺过滤范围,所得结果在 Metabolynx 浏览器中列出,见图 2。图中 Expected Metabolites 栏中为根据设定的代谢途径提取的目标离子,Unexpected Metabolites 栏为意外检测到的离子,它们仅在含药血清样品中存在而在空白样品中不存在,通过设定质量短缺过滤范围等过滤参数可缩小此检测范围,进一步通过 MS/MS 分析进行确定。

当然,除了本文介绍的 Metabolynx 软件外,还有其他多种高分辨 LC-MS 数据处理软件均有代谢物识别功能,如 MetaSite, Metabolite Predict, MetWorks

等<sup>[13]</sup>。本文仅结合作者的应用着重介绍了 Metabolynx 软件,其他相关软件可参考具体应用文献。

### 3 模式识别方法

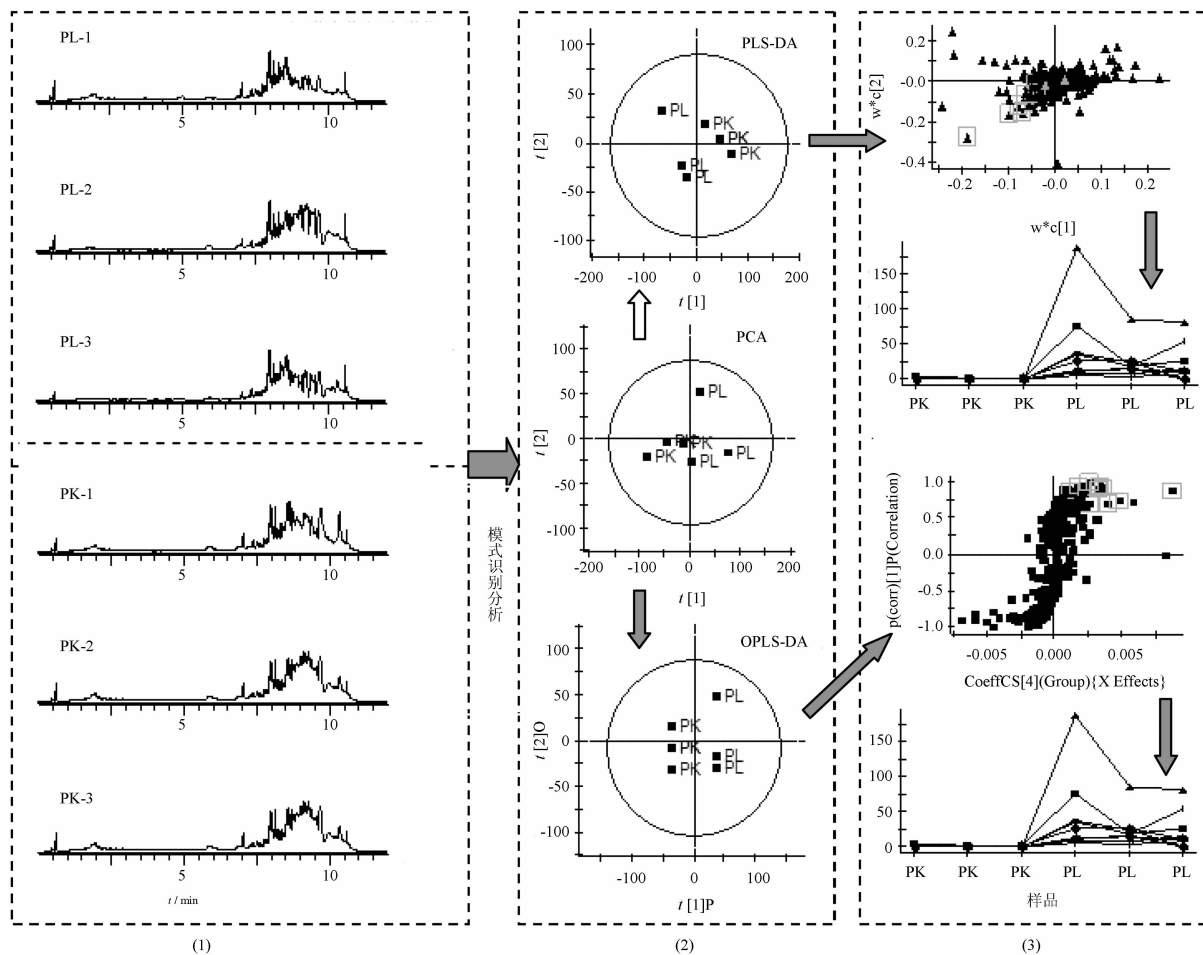
模式识别方法主要有“无监督”的主成分分析法(PCA)、“有监督”的偏最小二乘判别分析法(PLS-DA)、正交偏最小二乘判别分析法(OPLS-DA),目前在代谢组学研究中获得有效应用,用于生物体整体代谢组变化轨迹的描述和判别,进而挖掘使产生代谢组变化的潜在生物标志物<sup>[14-15]</sup>。

以往的直观分析方法是空白样品色谱图为对照,在含药样品色谱图中筛选相同保留时间处空白样品色谱图中没有的峰,即认为是中药体内成分峰<sup>[16-17]</sup>。基于此原则,通过采集空白样品和含药样品 2 组数据,利用模式识别方法对采集的 LC-MS 图谱进行处理,挖掘在含药样品组中存在而在空白样品组中为零的离子,即为潜在的中药体内成分,进而通过进一步的鉴定进行确认。因此,对中药体内成分分析时,需要采集空白生物样品(血清、胆汁、尿液等)和给药后的含药样品 2 组数据,每组数据包

括至少 3 个平行样品。首先对 2 组数据进行 PCA 分析,绘制出反映组间离散程度的得分图 (score plot),进而进行加入分组信息的 PLS-DA 分析,强化组间差异,尽可能的减少或消除组内差异,得 scores plot 和 loadings plot,在 loading plot 中观察在外围离散的离子的 trend plot,选择在空白样品中含量为零或接近于零(误差范围内)而在含药样品中较高的离子,即为潜在的中药体内成分。值得提及的是,由于分析对象本身就是不同的 2 组数据,因而进行有

监督的 OPLS-DA 分析,在所得的 S-Plot 中能更为清晰快速地筛选潜在的中药体内成分,选择的原则同 PLS-DA。最后,通过与中药体外样品色谱图的比较和 MS/MS 质谱数据分析,对挖掘的潜在中药体内成分进行确认。

应用模式识别方法,作者快速实现了大鼠口服给予黄连后血中移行成分的全息检测和辨识,见表 1,与前 2 种辨识技术处理所得结果基本一致,分析流程见图 3。



1. 给药组和空白组的 LC-MS 图;2. PCA, PLS-DA, OPLS-DA 识别方法的得分图;3. PLS-DA, OPLS-DA 分析的载荷图和所选变量对应的趋势图; PK. 正离子模式下检测的空白组; PL. 正离子模式下检测的黄连给药组。

图 3 模式识别方法应用于 LC-MS 图中中药体内成分辨识的分析流程

Fig. 3 The analysis flowchart of pattern recognition methods used for discrimination of the constituents of traditional Chinese medicine in vivo in the LC-MS chromatograms

#### 4 综合概论

质量短缺过滤技术通过设定适当的质量短缺范围而对全扫描高分辨 LC-MS 图谱进行处理,过滤掉

大量内源性干扰物质峰,从而使获得的图谱中代谢物峰能够明显地暴漏出来而易于识别。Metabolynx 软件一方面通过代谢途径的设定能够靶向提取预定



代谢物峰,另一方面通过比较含药样品和空白样品图谱而提取只有在含药样品中存在的潜在代谢物峰供筛选,能有效地避免漏检。比较而言,前者是根据质量短缺值的变化范围进行成分的提取,而后者是根据设定的代谢途径进行靶向地提取,同时辅以质量短缺过滤等多种参数的设置以提供一些潜在的成分供筛选。然而,由于这 2 种技术均是针对单体化合物的代谢物进行辨识,因此,对于含有复杂化学成分的中药进行体内成分的 LC-MS 分析和辨识时,就需要对每个组分或结构类型进行单独提取,操作上就略显繁琐;并且需要对提取的成分进行质量短缺值的总结以及代谢途径的预测,对未知代谢物及其来源的辨识就略显不足。模式识别方法是一种化学计量学方法,该方法(尤其是 OPLS-DA)能将在给药组中存在而空白组中不存在的离子最大程度显露出来,因此,该方法不需要对中药的化学成分及其代谢途径进行总结和预测,仅通过对 2 组样品多个数据文档的多变量统计分析,挖掘造成 2 组数据分组的成分,并选择只有在给药组中出现的离子作为潜在的中药体内成分,是更为快速、简便、全息的中药体内成分辨识技术。总而言之,在进行中药体内成分分析时,可根据特定的分析目的,选择最适合的辨识技术,对于有特定结构类型的成分,可选择前 2 个,而对于成分复杂的中药要进行体内全成分分析,则选择模式识别方法能获得更完善的结果。

#### [参考文献]

- [1] 王喜军. 中药血清药物化学的研究动态及发展趋势[J]. 中国中药杂志, 2006, 31(10): 789.
- [2] 王喜军, 孙文军, 张宁, 等. 六味地黄丸血中移行成分的分离及结构鉴定[J]. 中国天然药物, 2007, 5(4): 277.
- [3] 姚新生. 中药天然药物活性成分的研究方法[J]. 药学服务与研究, 2003, 3(4): 205.
- [4] Sun H, Lv H T, Zhang Y M, et al. HPLC method for preliminary analysis of constituents in rat blood after oral administration of the extract of *Acanthopanax senticosus* [J]. J Sep Sci, 2007, 30: 3120.
- [5] Wang X J, Sun W J, Sun H, et al. Analysis of the constituents in the rat plasma after oral administration of Yin Chen Hao Tang by UPLC/Q-TOF-MS/MS [J]. J Pharm Biomed Anal, 2008, 46: 477.
- [6] Wang X J, Sun H, Zhang A H, et al. Ultra-performance liquid chromatography coupled to mass spectrometry as a sensitive and powerful technology for metabolomic studies [J]. J Sep Sci, 2011, 34(24): 3451.
- [7] Zhang H Y, Zhang D L, Ray K, et al. Mass defect filter technique and its applications to drug metabolite identification by high-resolution mass spectrometry [J]. J Mass Spectrom, 2009, 44: 999.
- [8] Zhu M S, Ma L, Zhang D L, et al. Detection and characterization of metabolites in biological matrices using mass defect filtering of liquid chromatography/high resolution mass spectrometry data [J]. Drug Metab Dispos, 2006, 34(10): 1722.
- [9] Mortishire-Smith R J, O'Connor D, Castro-Perez J M, et al. Accelerated throughput metabolic route screening in early drug discovery using high-resolution liquid chromatography/quadrupole time-of-flight mass spectrometry and automated data analysis [J]. Rapid Commun Mass Spectrom, 2005, 19(18): 2659.
- [10] Peters R J, van Engelen M C, Touber M E, et al. Searching for in silico predicted metabolites and designer modifications of (cortico) steroids in urine by high-resolution liquid chromatography/time-of-flight mass spectrometry [J]. Rapid Commun Mass Spectrom, 2009, 23(15): 2329.
- [11] Yan G L, Sun H, Sun W J, et al. Rapid and global detection and characterization of aconitum alkaloids in Yin Chen Si Ni Tang, a traditional Chinese medical formula, by ultra performance liquid chromatography-high resolution mass spectrometry and automated data analysis [J]. J Pharm Biomed Anal, 2010, 53: 421.
- [12] Guo J, Shang E X, Duan J A, et al. Fast and automated characterization of major constituents in rat biofluid after oral administration of *Abelmoschus manihot* extract using ultra-performance liquid chromatography/quadrupole time-of-flight mass spectrometry and MetaboLynx [J]. Rapid Commun Mass Spectrom, 2010, 24(4): 443.
- [13] Prasad B, Singh S. Identification of rat urinary metabolites of rifabutin using LC-MS<sup>n</sup> and LC-HR-MS [J]. Eur J Pharm Sci, 2010, 41(1): 173.
- [14] Wang P, Sun H, Lv H T, et al. Thyroxine and reserpine-induced changes in metabolic profiles of rat urine and the therapeutic effect of Liu Wei Di Huang Wan detected by UPLC-HDMS [J]. J Pharm Biomed Anal, 2010, 53(3): 631.
- [15] Yang B, Zhang A H, Sun H, et al. Metabolomic study of insomnia and intervention effects of Suanzaoren decoction using ultra-performance liquid-chromatography / electrospray-ionization synapt high-definition mass spectrometry [J]. J Pharm Biomed Anal, 2012, 58: 113.
- [16] Schneider R P, Zhang H, Mu L, et al. Utility of multivariate analysis in support of *in vitro* metabolite identification studies: retrospective analysis using the antidepressant drug nefazodone [J]. Xenobiotica, 2010, 40: 262.
- [17] Sun J, Von Tungeln L S, Hines W, et al. Identification of metabolite profiles of the catechol-O-methyl transferase inhibitor tolcapone in rat urine using LC/MS-based metabolomics analysis [J]. J Chromatogr B Analyt Technol Biomed Life Sci, 2009, 877: 2557.



## Identification technique for *in vivo* ingredients of traditional Chinese medicines based on LC-MS analysis

YAN Guangli, HANG Ying, WANG Xijun\*

(Heilongjiang University of Chinese Medicine, Harbin 150040, China)

[**Abstract**] Serum pharmacochimistry of traditional Chinese medicine (TCM) is a direct and effective method for determining efficacious substance foundation of TCM. However, the complexity of chemical constituents of TCM and the interaction among ingredients in the *in vitro* process make the analysis on *in vitro* ingredients of TCM arduous. Liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) has become the cornerstone in detection and characterization of *in vivo* ingredients of TCM because of its sensitivity and ability to analyze complex mixtures. However, due to significant interference from endogenous species, detection and identification of the constituents of TCM in the biological matrices are often difficult. There is a crying need for introducing specialized ingredient identification techniques to avoid artificial omission of *in vivo* ingredients of TCM. On the basis of the analysis on transitional ingredients in rat blood, this essay introduces the application of such pattern recognition methods as mass defect filter, Metabolynx software and principal component analysis, partial least squared discriminant analysis and orthogonal partial least squared discriminant analysis in identifying *in vivo* ingredients of TCM.

[**Key words**] serum pharmacochimistry of traditional Chinese medicine; analysis of the constituents *in vivo*; mass defect filter; metabolynx; pattern recognition

doi:10.4268/cjmm20121216

[责任编辑 马超一]