

分类号_____

密级_____

U D C_____

编号_____

厦 门 大 学

博 士 后 研 究 工 作 报 告

基于超高效液相色谱-质谱的疾病代谢组学研究

张 洁

工作完成日期 2009. 04. 1

报告提交日期 2009. 04. 1

厦门大学

2009 年 4 月

基于超高效液相色谱-质谱的疾病代谢组学研究

Disease-related Metabonomics Research using Ultra-performance Liquid
Chromatography – Mass Spectrometry

博 士 后 姓 名 张洁

流动站（一级学科）名称 化学

专 业（二级学科）名称 分析化学

研究工作起始时间 2007. 04

研究工作期满时间 2009. 04

厦 门 大 学

2009 年 04 月

厦门大学博士后研究报告著作权使用声明

本人完全了解厦门大学有关保留、使用博士后研究报告的规定。厦门大学有权保留并向国家主管部门或其指定机构送交该报告的纸质版和电子版，有权将该报告用于非赢利目的的少量复制并允许该报告进入学校图书馆被查阅，有权将该报告的内容编入有关数据库进行检索，有权将博士后研究报告的标题和摘要汇编出版。保密的博士后研究报告在解密后适用本规定。

本研究报告属于： 1、保密（ ）， 2、不保密（ ）

纸本在 年解密后适用本授权书；

电子版在 年解密后适用本授权书。

（请在以上相应括号内打“√”）

作者签名： 日期： 年 月 日

导师签名： 日期： 年 月 日

内 容 摘 要

代谢组学是功能基因组学和系统生物学研究的重要组成部分。作为一个发展中的代谢组学分析工具,超高效液相色谱正受到越来越多的重视。本文以UPLC/MS的代谢组学研究为中心,开展了以下几方面的工作:

我们针对8例糖尿病肾病患者、33例2型糖尿病患者和25例正常人发展了基于UPLC/MS的代谢组学研究方法。患者和正常人之间可以得到明显的区分,而且两种疾病患者都可以投影到不同的区间得到较好的区分。通过精确分子量、同位素分布、MS/MS信息和标准品对照,我们初步鉴定了亮氨酸、4-羟双氢鞘氨醇和二氢鞘氨醇等标志物。我们发现,在患者的血清中这三种标志物的水平发生了较明显的变化,这也说明代谢性疾病可能引起氨基酸代谢和磷脂代谢的异常,为临床诊断和治疗提供了一条新的思路。

我们还将UPLC/MS系统应用于前列腺增生和前列腺癌的血清的代谢组学研究。结果表明,在PLS-DA投影图中,两种疾病患者和正常人分布在三个独立的区域内,可以得到良好的区分。此外,我们还初步鉴定出PE(22:6)、LysoPC(18:0)、LysoPC(22:6)、4-羟双氢鞘氨醇、二氢神经酰胺和神经酰胺等可能的标志物。

在基于UPLC的代谢组学研究平台中,由于色谱分离的引入使得整个分析时间增加,导致了较低的分析通量。我们以糖尿病患者和正常人的血清和尿样为研究模型,应用UPLC/MS系统,发展并详细比较了快速梯度洗脱和慢速梯度洗脱模式对聚类效果和标志物筛选的影响。在两种模式中,血清的PLS-DA分析结果均优于尿样,而且正离子扫描能够提供更好的聚类结果。另外,在聚类分析方面,快速梯度洗脱模式可提供与慢速梯度洗脱模式相当的效果,而且前者具有较高的分析通量。在标志物筛选方面,虽然在高丰度标志物筛选方面两种模式具有相当的效果,但是考虑到快速梯度洗脱模式的不充分分离可能导致的离子抑制效应对于低丰度标志物的影响,慢速梯度洗脱模式是标志物筛选研究中更好的选择。

关键词: 超高效液相色谱, 质谱, 代谢组学, 疾病

Abstract

Metabonomics is gaining increasing interest in many fields recently, which is an important component of the function genomics and systems biology. As a developing and valuable tool for the metabolomics analysis of complex samples, ultra-high performance liquid chromatography (UPLC) is gaining more attention. Focusing on UPLC/MS, the following work in this thesis has been carried out.

We developed metabonomics method to distinguish the global serum profiles of 8 diabetic nephropathy (DN) patients, 33 type 2 diabetes mellitus (T2DM) patients and 25 healthy volunteers using UPLC/MS system. The distinct clustering between the patients and controls was observed, and DN and T2DM patients were also separated into two individual groups. Several compounds were tentatively identified based on accurate mass, isotopic pattern and MS/MS information. In addition, significant changes in the serum level of leucine, dihydrosphingosine and phytosphingosine were noted, indicating the perturbations of amino acid metabolism and phospholipid metabolism in diabetic diseases, which having implications in clinical diagnosis and treatment.

We applied UPLC/MS in the metabonomics research of prostate diseases including prostatic cancer (PC) and prostatic hyperplasia (PH). The distinct clustering of PC patients, PH patients and the controls was observed. PE(22:6), LysoPC(18:0), LysoPC(22:6), phytosphingosine, dihydroceramide and ceramide were tentatively identified as potential markers based on accurate mass and isotopic pattern.

In UPLC-based metabonomics research, the introduction of separation step prolongs the total analysis time, consequently resulting in relatively lower throughput. We analyzed the serum and urine of the patients with type 2 diabetes mellitus and the controls with UPLC/TOF-MS, comparing fast gradient with slow gradient, using both positive and negative ionization modes. In both fast and slow gradients, the PLS-DA models generated with serum samples were more robust than those with urine samples, and positive ionization mode demonstrated better differentiation and higher classification rate than negative ionization mode. In addition, the fast gradient was found to have comparable differentiation capability to slow gradient. Considering the

sample throughput obtained by fast gradient mode was improved by 5 times, we recommend it to be employed for the differentiation analysis of clinical samples. The biomarker discovery capabilities of both gradients were compared using positive ionization mode. Some biomarkers obtained by slow gradient were observed in the loading plot of fast gradient, and the relative intensities of the corresponding biomarkers showed similar trends. However, the matrix effect induced by insufficient separation using fast gradient have deleterious effects on the detection of some important metabolites with low abundance. Therefore, slow gradient was the better choices than fast gradient for the biomarker discovery research.

Keywords: Ultra-high performance liquid chromatography; mass spectrometry; metabonomics, disease

目 录

第一章 文献综述.....	1
1.1 代谢组学的定义.....	1
1.2 代谢组学研究的层次.....	1
1.3 代谢组学的技术平台.....	2
1.3.1 样品制备.....	3
1.3.2 分析技术.....	3
1.3.2.1 核磁共振技术.....	4
1.3.2.2 色谱及其联用技术.....	6
1.3.2.2.1 液相色谱-质谱联用技术.....	6
1.3.2.2.2 毛细管液相色谱-质谱联用技术.....	7
1.3.2.2.3 超高效液相色谱-质谱联用技术.....	8
1.3.2.2.4 多维液相色谱.....	8
1.3.2.3 傅里叶变换离子回旋共振技术.....	9
1.3.2.4 直接输注大气压电离化质谱技术.....	9
1.3.2.5 基于多孔硅表面的解吸离子化技术.....	9
1.3.3 数据分析.....	9
1.3.3.1 数据预处理.....	9
1.3.3.2 模式识别.....	10
1.3.3.2.1 非监督方法.....	10
1.3.3.2.2 有监督方法.....	10
1.3.3.3 代谢物数据库.....	10
1.3.3.4 途径数据库.....	10
1.4 代谢组学的研究进展.....	11
1.4.1 临床代谢组学.....	11
1.4.2 药学代谢组学.....	11
1.4.3 植物代谢组学.....	12
1.4.4 微生物代谢组学.....	12
1.5 展望.....	12

参考文献.....	13
第二章 糖尿病和糖尿病肾病的代谢组学研究.....	25
2.1 前言.....	25
2.2 实验部分.....	26
2.2.1 试剂.....	26
2.2.2 血清来源及处理.....	26
2.2.3 超高效液相色谱.....	26
2.2.4 质谱.....	26
2.2.5 数据分析.....	27
2.3 结果与讨论.....	27
参考文献.....	35
第三章 前列腺疾病的代谢组学研究.....	37
3.1 前言.....	37
3.2 实验部分.....	38
3.2.1 试剂.....	38
3.2.2 血清来源及处理.....	38
3.2.3 超高效液相色谱.....	38
3.2.4 质谱.....	38
3.2.5 数据分析.....	39
3.3 结果与讨论.....	39
参考文献.....	49
第四章 基于超高效液相色谱的高通量筛选方法研究.....	52
4.1 前言.....	52
4.2 实验部分.....	53
4.2.1 试剂.....	53
4.2.2 血清和尿样来源及处理.....	53
4.2.3 超高效液相色谱.....	53
4.2.3.1 快速洗脱梯度.....	53

4.2.3.2 慢速洗脱梯度	54
4.2.4 质谱	54
4.2.5 数据分析	54
4.3 结果与讨论	54
参考文献	66
致 谢	67
博士生期间发表的学术论文、专著	68
博士后期间发表的学术论文、专著	71
作者简介	72
联系地址	73

厦门大学博硕士学位论文摘要库

第一章 文献综述

1.1 代谢组学的定义

生命的代谢活动是个永不停止的过程,代谢活动是生命活动的本质特征和物质基础。代谢组学是指生物体内源性代谢物质的动态整体[1-3]。传统的代谢概念既包括生物合成也包括生物分解,因此理论上代谢物包括核酸、蛋白质、脂类以及其他小分子代谢物质。为了有别于基因组、转录组和蛋白质组,代谢组目前只涉及相对分子质量约小于1000的小分子代谢物质。

目前代谢组学的定义普遍接受的有两种:① Metabolomics[4] (Steven Oliver):通过考察生物体系受刺激或扰动后(如将某个特定的基因变异或环境变化后)其代谢产物的变化或其随时间的变化,来研究生物体系的代谢途径的一种技术;② Metabonomics[1] (Jeremy K. Nicholson):生物体对病理生理刺激或基因修饰产生的代谢物质的质和量的动态变化的研究。一般认为,前者一般以细胞为研究对象,广泛地应用于植物和微生物研究领域;后者则更注重动物的体液和组织,广泛地应用于药物开发和疾病诊断研究领域。

1.2 代谢组学研究的层次

根据研究对象和目的的不同, Oliver Fiehn 将代谢组学分为 4 个层次[5]:

- 1) 代谢物靶标分析(metabolite target analysis):对某个或某几个特定组分的分析。在这个层次中,需要采取一定的预处理技术,除掉干扰物,以提高检测的灵敏度。
- 2) 代谢轮廓分析(metabolic profiling analysis):对少数所预设的一些代谢产物的定量分析。如某一类结构、性质相关的化合物、某一代谢途径的所有中间产物或多条代谢途径的标志性组分。
- 3) 代谢组学(metabonomics):对限定条件下的特定生物样品中所有代谢组分的定性和定量。进行代谢组学研究时,样品的预处理和检测技术必须满足对所有的代谢组分具有高灵敏度、高选择性、高通量的要求,而且基体干扰要小。代谢组学涉及的数据量非常大,因此需要有能对其数据进行解析的化学计量学技术。
- 4) 代谢指纹分析(metabolic fingerprinting analysis):不分离鉴定具体单一组分,而是对样品进行快速分类。

从严格意义而言,只有第 3 层次才是真正的代谢组学研究。由于目前还没有发展出一种真正的代谢组学技术可以涵盖所有的代谢物而不管分子大小和性质,

所以代谢组学的最终目标还是不可完成的任务。但是，在具体的实验中，代谢组学研究必须在目前的技术条件下设法解析所有的可见峰（即代谢物）。

1.3 代谢组学的技术平台

代谢组学研究的对象可以是细胞、组织或者生物机体整体。代谢组学要求能对限定条件下特定样本中的所有代谢组分进行定性和定量分析。完整的代谢组学分析的流程包括样品的采集和制备、样品的分析及数据分析。代谢组学力求分析生物体系中的所有代谢产物，整个分析过程应尽可能地保留和反映总的代谢产物的信息。代谢组学方法得到的数据是海量的数据。对代谢组学产生的海量数据必须使用各种化学计量学软件和生物信息学方法对其进行数据挖掘，只有这样，我们才能从数据中得到所需的知识。

如图1-1所示，代谢组研究一般可以包括以下流程：

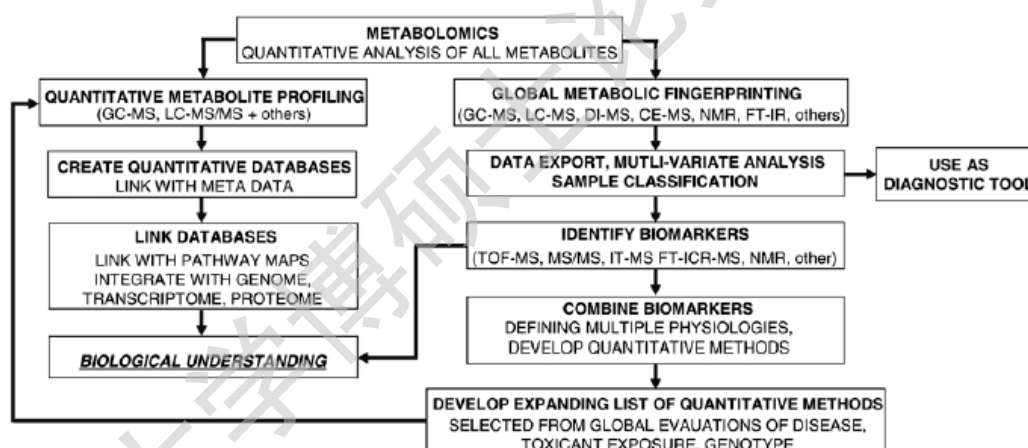


Figure 1-1 Flow chart of a typical metabolomics platform.

- 给研究对象引入一定的外源性刺激，该刺激既可以是基因的改变(剔除或导入)、转录水平的更改、蛋白质水平的变化，也可以是并不会导致基因和转录水平发生变化的某环境因素。
- 采集相关的如尿液、血液、组织、细胞和培养液，甚至整个生物体的生物样品，以反映时空信息。实验设计中样品收集的时间、部位、种类、样本群体等应给予充分考虑。
- 用磁共振、质谱、色谱等分析手段检测其中代谢物的种类、含量、状态及其变化，建立代谢组数据。
- 使用多变量数据分析方法，表征代谢组特征的动态模型，确定和研究相关代谢物变化涉及的代谢途径，进而联系该变化规律，从不同层次和水平上阐述

生物体对相应刺激响应。

1.3.1 样品制备

针对不同的研究对象和分析技术，样品的提取和预处理方法也就多种多样。对于基于 NMR 的技术平台，对样品进行很少的预处理就可达到要求。对于基于色谱的技术平台，一般都需要进行较为复杂的处理才能进行分析。比如通常用水或有机溶剂分别提取，获得水提取物和有机溶剂提取物，从而把非极性相和极性相分开，以便进行分析。分析之前，常常还用固相微萃取、固相萃取和亲和色谱等预处理。在做气相色谱或气相色谱质谱联用时常常需要进行衍生化，使样品适合于气相分析的需要。

由于某种提取条件往往只适合某些化合物，但对于另外一些化合物的稳定性却不利。截至目前，没有一种能够适合所有代谢产物的提取方法，许多重要的化合物都被忽视了。实际的做法是，应该根据不同的化合物选择不同的提取方法，并对提取条件进行优化。

1.3.2 分析技术

代谢物整体水平的检测分析必须依赖分析化学中的各种谱学技术，包括磁共振波谱、质谱、色谱、红外和拉曼光谱、紫外-可见光谱等及其偶合联用方法获取代谢组数据(图 1-2)。其中，色谱以其高分离度、高通量，质谱以其普适性、高灵敏度和特异性，核磁共振技术是因其对含氢代谢产物的普适性而成为最主要的分析工具；由于液质联用和气质联用能分析范围很广的代谢组分，因此也成为代谢组学研究分析中最重要的工具。

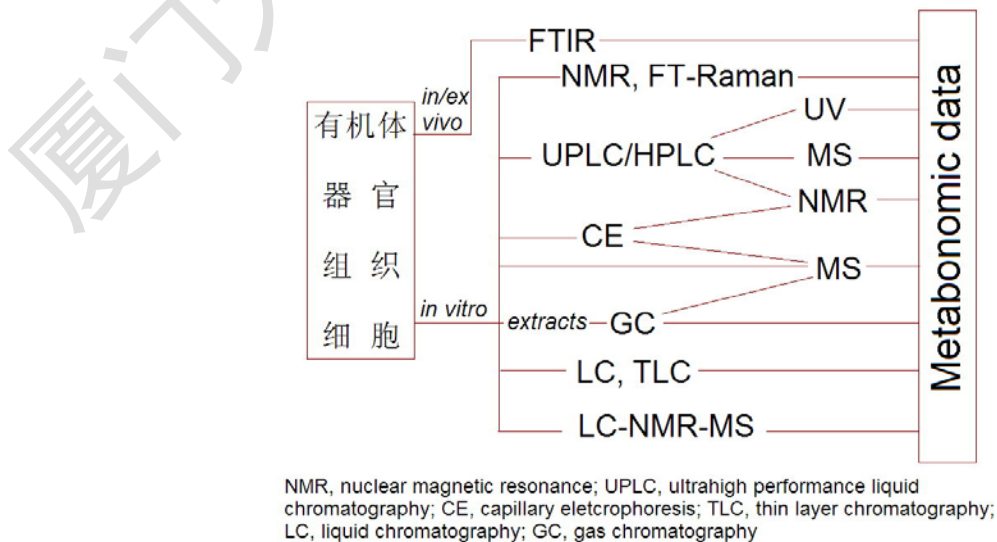


Figure 1-2 Data-acquisition techniques in metabonomics research

在代谢组学的分析方法的选择上，需要在分析速度、无偏性和灵敏度上综合考虑[6-9]。如图 1-3 所示，核磁共振具备快速和无偏向性的特点，但与其它技术相比灵敏度则太低；其它方法（如 CE/LIF）有着很高的检测灵敏度，但为选择性检测。而色谱质谱联用能够很好的兼顾灵敏度和无偏性，但需要较长的分析时间，在分析速度即通量上不具备优势。

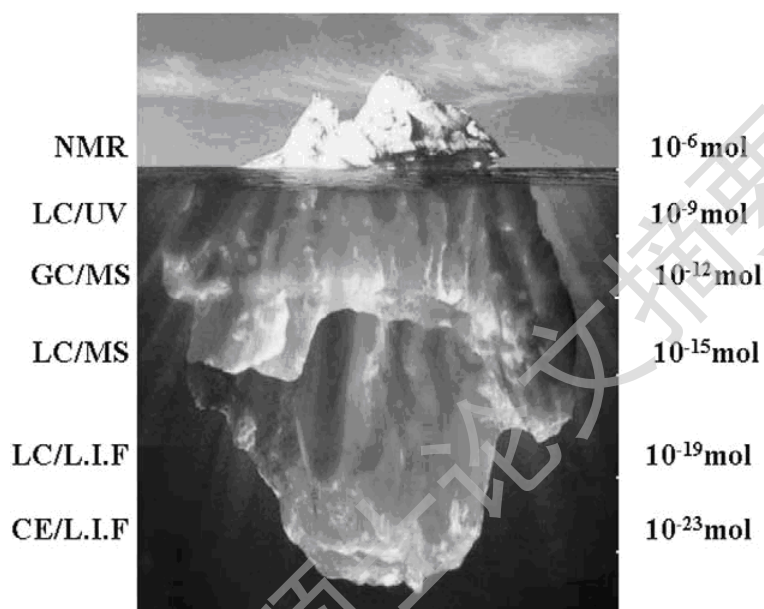


Figure 1-3 A comparison of the relative sensitivities of various metabolomic tools. NMR has rapid analysis times but suffers from lower sensitivity thus allowing visualization only of the more concentrated metabolites (i.e. the tip of the iceberg). GC/MS and HPLC/MS provide good selectivity and sensitivity. CE/LIF provides very high sensitivity but lower selectivity.

1.3.2.1 核磁共振技术

基于核磁共振技术的代谢组学研究，是近几年发展起来的一种新的组学技术。它主要是利用生物体液的核磁共振谱图所提供的生物体内全部小分子代谢物的丰富信息，通过对这些信息的多元统计分析和模式识别处理，了解相关生物体在功能基因组学、病理生理学、药理毒理学等方面的状况及动态变化，以及它们所揭示的生物学意义，并从分子水平来认识生命运动的规律。

核磁共振波谱技术具有以下优势[10-16]:

- ① 良好的客观性和重现性，因而便于不同实验室之间数据的交换和比较。
- ② 样品不需要繁琐处理，可在接近生理条件下进行实验。
- ③ 具有无创性，不破坏样品的结构和性质，因而便于活体、原位的动态检测。

④ 代谢组中代谢物质的响应系数相同，因此可以进行一次性同步、无偏向的检测而且具有良好的原位定量效果。

⑤ 检测具有优异的重现性，其信号携带着原子之间连接关系、动力学性质和相互作用等丰富的分子信息，便于确定未知代谢物质的结构和性质。

⑥ 可以对细胞和组织等进行原位无创的检测分析而不受样品具体形态的限制，而且具有较高的通量和较低的单位样品检测成本等优点。

当然，核磁共振技术在检测生物样品中也存在一些问题，比如：

① 灵敏度低：相对于其它光谱和色谱仪器，核磁共振有较低的灵敏度。对于 500 MHz(或以上)谱仪采用 ^1H NMR 的检测限理论上为 $10\ \mu\text{M}$ ，但实际上 $<100\ \mu\text{M}$ 的物质已不能很准确定量。

② 水峰过强：生物分子主要是以水作为介质而生存的，离了水，就几乎无活性可言或者说有活性也难以发挥。因此生物 NMR 所研究的样品都是溶在水中或重水中的。体液中的生物分子处于丰富的水环境中，一般其浓度为毫摩尔级，而水中质子浓度约为 $100\ \text{mol/L}$ ，是生物分子浓度的 10^5 倍。如果水信号的信噪比为 10^5 ，生物分子的信噪比就只有 1。这样就使得 ^1H NMR 谱中水峰过强。

③ 谱峰重叠：大量生物分子的共振处于一个很窄的化学位移范围中，使得生物分子的 NMR 信号不可避免地相互发生重叠。因此在多数情况下，必须用化学位移、偶合方式以及体液的种类等多方面的信息进行综合评价。

④ 谱线较宽：体液和组织中某些生物分子如蛋白质、糖等在 NMR 谱上为宽峰或宽带，而一些小分子在通常的溶液状态下也会与蛋白偶联，致使其谱峰在 NMR 谱图中“不可见”。这种现象在血浆和胆汁中更为明显。

NMR 是当前代谢组学研究中的主要技术，NMR 技术广泛地应用于药物毒性 [17-25]、基因功能 [26-27]、疾病的临床诊断 [28-37] 中。随着高场傅里叶变换核磁共振波谱仪灵敏度和分辨率的增加、化学位移的扩展、多脉冲实验技术的发展以及计算机技术的不断进步，NMR 技术在生物和医学领域的应用有了飞速的发展，本身所存在的问题目前已有解决的方法。对体液而言，在 400 MHz 上观察体液代谢物就会得到较好的信号，并且在 500 和 600 MHz 仪器上能获得更高的灵敏度，可观测到相对较低浓度的代谢物的信号，而 2D NMR 的使用还可提供更多信息。生物组织富含蛋白质、脂类等生物大分子，由于偶极相互作用和化学位移各向异性等因素，以及分子运动受限和磁化率的不均匀性导致谱线的增宽不能被有效平均掉，但可以利用固体核磁中的魔角旋转(Magic Angle Spinning, MAS)方

法可以克服这些谱线增宽因素[38-42]。让生物组织以魔角(54.7°)的方向以 $2\text{k}\sim 5\text{k}$ Hz 左右频率快速旋转, 就可以有效平均掉生物组织内各种相互作用和磁化率不均匀引起的谱线变宽, 提高分辨率。

NMR 虽然可对化学组成知之甚少的复杂样品如尿液、血液等进行非破坏性分析, 但 NMR 的灵敏度和动态范围有限, 很难同时测定生物体系中共存的浓度相差较大的代谢产物, 同时所需硬件的投资也较大。

1.3.2.2 色谱及其联用技术

在过去的几十年里, 色谱技术因其卓越的分离性能、高灵敏度已被广泛用于复杂体系的分离分析。近年来, 越来越多的研究者将色谱及其联用技术用于代谢组学的研究。

1.3.2.2.1 液相色谱-质谱联用技术

气质联用(GC/MS)技术和液质联用(LC/MS)可以同时检测出数百种化合物, 包括糖类、有机酸、氨基酸、脂肪酸和大量植物的次级代谢产物[43]。GC/MS 技术需要先对样品进行衍生化预处理, 而且容易引起样品的变化。受此限制, GC/MS 技术不能分析热不稳定物质和一些大分子的代谢产物。而 LC/MS 技术不受此限制, 又经济实用, 适用于那些热不稳定, 不易挥发、不易衍生化和分子量较大的物质[44]。质谱多通道监测的功能和色谱卓越的分离能力使 LC/MS 技术对检测样品的浓度和纯度要求与 NMR 技术相比明显降低, 甚至对含量极低的物质也能通过优化质谱的扫描模式给出可视化响应。同时, LC/MS 技术又有较好的选择性和较高的灵敏度, 得到了越来越广泛的应用。

Plumb等[45]采用梯度HPLC/TOF-MS和MS/MS对大鼠和小鼠的不同种系的尿液进行检测, Symmetry C18 $3.5\mu\text{m}$ HPLC柱, 质量数范围 $50\sim 850$, 柱后分流 $120\mu\text{L}/\text{min}$ 进入质谱检测, 详细研究了黑鼠、白鼠和裸鼠种系和性别对尿样代谢轮廓的影响。

Idboorg-Bjorkman[46]用抗抑郁药西酞普兰处置大鼠, 诱导磷酸酶的生物标记物。用反相HPLC/MS正负离子电喷雾离子检测给药组和对照组大鼠的尿液。样品SPE萃取, 色谱分离用Xterra C18柱梯度洗脱, 用固相萃取(SPE)对样品预处理或液液萃取进行代谢指纹谱研究时, 必须仔细优化避免重要分析物的丢失。SPE预处理有利于样品的浓缩, 减少盐的干扰, 并为HPLC/MS的分析提供合适的

形式。

Lafaye等[47]采用HPLC/MS进行代谢组学研究，调查了重金属对大鼠的毒性。大鼠用氯化镉或硝酸双氧铀处置后检测尿液。反相梯度HPLC， XTerra C18 (5 μ m, 2.1 mm \times 15 cm)，柱温30 $^{\circ}$ C，电喷雾离子正负离子分析样品；扫描质量数范围100~1000。两种重金属毒物的不同剂量给药，其内源性代谢物也不同。他们采用的手工处理数据而非模式识别技术，发现了许多异常的内源性代谢物。近年来多数代谢组学研究使用高场NMR来完成；目前对HPLC/MS在代谢组学研究的有效性也进行了评价。

Lenz等[48]对同一样品采用HPLC/MS对比NMR检测的结果。WD大鼠模型药肾毒性环孢霉素A给药后，收集尿样，对小分子有机物用HPLC/MS和NMR研究。¹H NMR观测到内源性物质发生明显的变化。包括葡萄糖、醋酸盐、三甲胺、琥珀盐的增加和三甲胺氧化物的减少。相应的HPLC-MS研究尿样中犬尿烯酸、黄尿烯酸、柠檬酸和核黄素含量下降，同时有大量的未确定化合物量的减少。

Williams[49]采用HPLC/MS和¹H NMR研究雌雄Zucker肥胖鼠和正常WD大鼠的尿液进行代谢组学分析，得到了代谢指纹谱，用化学计量学对数据处理。研究表明，不同种系、性别和昼夜差异，其内源性物质均有明显差异；HPLC/MS和¹H NMR对内源性物质的检测能力各有优势，能够相互补充。

HPLC-MS技术应用于代谢组学研究的例子还有很多，Hollywood和Dettmer等人[50-52]做了很详细的综述，这里就不再赘述了。

1.3.2.2.2 毛细管液相色谱-质谱联用技术

目前，虽然使用传统的HPLC/MS能够检测Zucker大鼠的尿液中1500多个峰[49]，但是为了减少峰重叠和离子抑制效应，同时又希望得到更多的复杂代谢指纹谱，分析时间较长，分析通量较低。因此需要找到其它的分离技术来达到更高的通量以符合实际需求。毛细管HPLC/MS具有强大的分离能力，能够有效减少离子源的离子抑制效应，而且分析速度较快，能够提高一定的分析通量。Tolstikov等[53]采用90 cm毛细管液相分离能够检测出几百个峰。另外，Wilson等[54]将毛细管HPLC/MS和传统HPLC/MS对比，表明前者能达到更好的分离和更好的聚类结果。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库