

现代生物技术前沿

许国旺 等著

代谢组学

——方法与应用

 科学出版社
www.sciencep.com

现代生物技术前沿

许国旺 等 著

代谢组学

——方法与应用

科学出版社

北京

内 容 简 介

本书是国内第一部集基本理论和实际应用于一体的、极有价值的关于代谢组学的专著,对学科的发展现状、面临问题、应用前景、未来趋势和学科本身的价值都做了客观、科学的描述。除简要回顾代谢组学的发展历史、特点外,重点介绍了代谢组学技术平台及其在健康疾病、药物毒性、植物、微生物、营养科学和环境科学研究中的应用,使读者能在短时间内对最新的技术和国内外进展有一全面了解。为适应不同层次人员对代谢组学知识的需求,本书在全面阐述色谱、质谱、核磁共振谱和多变量数据分析方法在代谢组学中的应用的的同时,图文并茂地剖析了代谢组学在不同领域内的应用,使读者能很容易地应用本书解决相关领域中的问题。

本书既可作为从事代谢组学研究的专业人员的参考书,也可作为相关领域研究生的教材。对代谢组学感兴趣或具备一定的生物化学和分析化学背景的读者也可根据自己所从事的专业有选择地阅读部分章节。

图书在版编目(CIP)数据

代谢组学:方法与应用/许国旺等著. —北京:科学出版社,2008
(现代生物技术前沿)

ISBN 978-7-03-021528-4

I. 代… II. 许… III. 代谢-研究 IV. Q493.1

中国版本图书馆 CIP 数据核字(2008)第 043299 号

责任编辑:夏 梁 席 慧/责任校对:刘小梅

责任印制:钱玉芬/封面设计:陈 敬

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新蕾印刷厂印刷

科学出版社发行 各地新华书店经销

*

2008 年 8 月第 一 版 开本:787×1092 1/16
2008 年 8 月第一次印刷 印张:27 1/2 插页:2
印数:1—2 000 字数:624 000

定价:88.00 元

(如有印装质量问题,我社负责调换<新蕾>)

序

随着“基因组学”的提出，目前冠以“组学”的名字已有 200 多种。代谢组学是研究细胞和生物体的所有代谢中间体和终产物（即代谢组）的一门新兴科学。相对于 DNA 或蛋白质等生物高分子而言，代谢组学的研究对象一般为分子质量在 1000Da 以下的小分子。不同于基因和蛋白质具有相对严格的种属和细胞特异性，同一代谢物在任何其存在的物种中都具有相同的理化性质。即便如此，代谢物的功能却并不限于是代谢途径中某酶的底物或产物，它们具有结构单元、能量的载体和储存体、信号分子、神经递质、转录和翻译的调控因子、蛋白质功能的调控因子、辅酶、分子伴侣、肠道因子和诱变剂等诸多功效，在生命活动中以代谢网络的形式相互作用，参与生命活动的各个过程。由于代谢网络处于基因调控网络、信号转导网络和蛋白质互作网络的下游，因此代谢组学研究能反映基因组、转录组和蛋白质组受内外环境影响后相互协调作用的最终结果，更接近于反映细胞或生物的代表型。

代谢物的众多功能和代谢组学的重要性及其测试技术的适用性使代谢组学成为系统生物学中继基因组学、转录组学、蛋白质组学后的一个重要的组学平台，被广泛地应用于医学、药学、动植物学、微生物学、环境科学和食品科学等生命科学的各个研究领域。国际上代谢组学的研究萌芽于 20 世纪 80 年代，在来自不同领域的科学家们的共同努力下，于 90 年代末期得到迅猛发展，形成以核磁共振（NMR）和色谱-质谱为核心的两大技术平台，并逐渐得到广泛应用。英国帝国理工大学和辉瑞（Pfizer）等六大制药公司在 COMET 计划中率先采用代谢组学方法来评价药物的毒性，并取得了极大的成功；美国 FDA 已尝试将代谢组学技术作为药物安全性评价的一种方法。以中国科学院大连化学物理研究所许国旺研究员为首的研究小组是最早系统从事代谢组学研究的科研团体之一，在国际上也有较高的学术地位。该研究组以色谱质谱联用技术为主导，从技术平台的建立到代谢组学在疾病、药物、营养、植物和微生物研究中的应用，形成了一套较为完整的代谢组学研究体系。

《代谢组学——方法与应用》是国内第一部代谢组学基本理论和实践应用的专著。该书的出版将会极大地推进代谢组学技术在中国的普及和应用，同时也将为相关部门合理规划代谢组学的发展提供必要参考。

中国工程院院士 杨胜利

2008 年 1 月

前 言

1995年，本人获马普(Max-Planck-Institute)研究基金资助，前往德国图宾根大学医学院(University of Tuebingen)作高级访问学者，开始涉足生命科学的研究。作为年轻的科学工作者，本人深知分析化学在生命科学中的重要性。健康与疾病是生命科学研究中永恒的主题，也是国家和社会发展中需要长期关注和投入的领域，为此，我选择了尿液中代谢组分和恶性肿瘤的关系作为切入点，采用亲和色谱的方法提取尿液中的修饰核苷，并首次建立了分析尿液中核苷的毛细管电泳方法，使检测分辨率提高的同时，也提高了分析通量。通过对罹患十几种不同癌症的患者尿液进行检测发现，癌症患者的尿液中修饰核苷总量明显高于正常人。但是始终没有发现某种特定核苷与癌症的关联。事实上，在不同的癌症患者个体中，并非所有修饰核苷的排放量均产生有规律的增高现象，而是出现了有的显著增高、有的略有增高、有的甚至反而降低的复杂情况。通过对大量数据的研究表明，癌症患者尿液中没有一种修饰核苷的排放量总是高于正常人，以其中某种修饰核苷的排放量作为依据时，癌症的检出率仅为40%左右。由此判定，采用单一代谢物指标进行癌症诊断的思路是值得商榷的。

人体是个复杂的系统，存在于这一系统中的物质之间必然存在着联系，那么修饰核苷之间也可能存在着某种联系。由此，我们在1997年开始将多变量的模式识别方法用于癌症诊断，对近千个癌症患者尿样中核苷的浓度数据进行处理，原先无序可循的数据则奇迹般地变得有规律起来。最终我们证明，尿液中修饰核苷的代谢模式变化可用于癌症的早期发现和术后随访，并具有广谱性，尤其适用于手术效果评价及预后评估。

在此基础上我们意识到，以小分子的产生作为表象的代谢活动应能够更准确地反映生物系统的状态，易于与现实世界建立正确的联系。对代谢物必须进行综合分析，也即只有对整个代谢指纹按“组学”的方式进行研究才有意义，疾病早期发现和诊断也应从利用单一标志物向组合标志物转变。恩师卢佩章院士早年提出的希望我们能进行“建立人体体液化学指纹数据库”的构想就是要用分析化学的方法，把一个人从婴幼儿到成年的化学指纹变化记录下来，从中发现个体疾病易感性，实际上就是人群的“代谢组学”研究。

2001年，在中国科学院大连化学物理研究所骨干人员参加的科研规划研讨会上，本人受当时的所长包信和研究员的邀请做了专题报告，第一次在所里正式介绍“代谢组学”，并被列为中国科学院知识创新工程(二期)重点支持方向。同年，杨胜利院士出任我所生物技术研究部主任，将“代谢组学”研究作为部里的重要发展方向，并将我们的研究平台介绍给相关单位。本课题组也在此时拥有了新名字——高分辨分离分析及代谢组学，并成为后来成立的大连化学物理研究所“代谢组学研究中心”的核心。这是学科凝炼的结果，也忠实记录了我们进行代谢组学研究的发展轨迹。

与此同时，科技部做出了及时的支持，2003年设立了国内第一个针对代谢组学的国家高技术研究发展计划(“863”计划)项目——代谢组学技术平台的研究，由本人担

任项目负责人。2003年底,中国科学院在上海生命科学院召开了关于“植物、微生物代谢组学和代谢工程”学术研讨会。在会上,本人受邀做了题为“植物、微生物代谢组学技术平台的研究”的报告,会场座无虚席,与会科研人员和青年学生对代谢组学所表现出的热情与2002年的情景形成鲜明对比。那年,同样是我受吴家睿副院长邀请在上海生命科学院做类似报告,吴院长为避免“冷场”,不得不亲自向逐个部门打电话动员听众。短短一年半时间,代谢组学在国内的“知名度”有了如此飞跃使本人在感动于国内科研人员巨大热情和科研敏感性的同时,更加坚定了进行代谢组学研究、推广代谢组学科学的信心。

近年来代谢组学蓬勃的发展势头与多学科领域科学家以及政府部门的极大重视密不可分。特别是在“十一五”期间,科技部先后在国家重大研究计划“蛋白质科学”、国家重点基础研究发展规划(“973”)、“863”计划和“十一五”科技支撑项目、食品安全等方面支持了代谢组学的研究。代谢组学和系统生物学被科技部和欧盟第7框架计划列为促进中医药现代化的主要途径之一。国家自然科学基金委员会化学科学部连续多年将“组学分析中的新原理、新方法和新技术”列为重点项目支持方向。同时很多科学家也都急于进入这个新的研究领域。为了适应我国代谢组学高速发展的需要,满足广大科学工作者的要求,我们在前期工作的基础上,应科学出版社邀请,撰写了本书。

本书第1、10章由许国旺撰写;第2章由路鑫撰写;第3章由杨军和王媛撰写;第4章由赵欣捷撰写;第5章由石先哲撰写;第6章由田晶撰写;第7章由孔宏伟撰写;第8章由王畅撰写;第9章由杨军撰写;第11章由袁凯龙、王畅、杨军撰写;第12章由尹沛源撰写;第13章由汪江山撰写;第14章由马晨菲撰写;第15章由高鹏撰写;第16章由赵春霞撰写。全书由许国旺统一汇总、定稿。这些作者都是奋战在代谢组学科研一线的青年,他们具有丰富的代谢组学研究经验,同时一直密切关注国际上代谢组学研究的发展,因此,此书中的许多内容反映了当前代谢组学的最新进展,也包含了本研究组从1996年以来在相关领域取得的研究成果。本书除阐述代谢组学的发展历史和特点外,重点介绍了代谢组学技术平台及其在疾病、药物、微生物、营养和环境研究中的应用,使读者对代谢组学研究的最新技术和国内外进展有一较为系统的了解。除全面地讲解了关于色谱(GC、LC)、质谱(MS)、核磁共振(NMR)和多变量统计方法在代谢组学中应用的基础知识外,本书力求图文并茂地剖析大量研究实例,展示代谢组学在不同领域内发挥的作用,使读者能够更好地应用本书解决生命科学面临的实际问题,以适应不同层次科研人员对代谢组学知识的需求。

本书在撰写过程中特别是在相关项目的进行过程中,得到了卢佩章院士、杨胜利院士、顾健人院士、张玉奎院士、J. van der Greef教授、J. Nicholson教授、刘昌孝院士、陈竺院士、王红阳院士、欧阳平凯院士、黎介寿院士、李兰娟院士、包信和研究员、张涛研究员、吴家睿研究员、马延和研究员、Mei Wang博士和其他同仁的大力支持和帮助,也得到了国家杰出青年基金“代谢组学技术平台”(No. 20425516)、国家重点基础研究发展规划(“973”)项目“生物系统催化的理论和方法:生物催化和生物转化中关键问题的基础研究”(2003CB716003)、“生物炼制细胞工厂的科学基础:细胞代谢网络的结构与动态特征分析”(2007CB707802)、“2型糖尿病发生发展机制的研究:2型糖尿病发病机制和营养干预的代谢组学研究”(2006CB503902)、国家“863”目标

导向类课题“妇科常见恶性肿瘤的生物标志物谱发现及临床应用”(2006AA02Z342)、“发酵法生产富马酸的技术研究”(2006AA02Z240)、“十一五”国家科技支撑计划“残留标示物高通量表征关键技术研究”(2006BAK02A12)和“中医药诊疗技术及特色产品应用研究”(2006BAI11B07)、科技部国际科技合作项目“代谢组学用于中药质量控制的方法学研究”(2007DFA31060)、中德中心“中德复杂样品的分离分析”联合研究小组(GZ364, NSFC 和 DFG)、国家自然科学基金“类脂分析的方法学研究”(20675082)、中国科学院重点方向性项目“植物、微生物的代谢组学”(KSCX2-SW-329)、“基于现代理论和技术的复方中药系统研究”(KGCX2-SW-213)、中国科学院重大方向性项目“抗2型糖尿病候选新药和干预新方法的系统性研究:糖尿病发生发展的营养干预和机理研究”(KSCX1-YW-02)和中国科学院创新基金“代谢组学在肝脏疾病研究中的应用”(K2006A12)、“代谢组学指纹图谱平台”(K2006A13)、“维医、维药的系统生物学研究”(K2006A14)的资助。本书的内容也是这些基金和项目支持下研究成果的集中体现。对此,我们谨表示最衷心的感谢。

最后,我还要感谢我的同事、学生及其相关家属的支持和理解,大家在极度繁忙的情况下牺牲很多休息时间完成了有关章节的写作。

由于时间仓促、水平有限,书中存在不妥之处在所难免,祈望广大读者不吝指正。

许国旺

2008年1月于大连

目 录

序

前言

第 1 章 绪论	1
1.1 代谢组学简介	1
1.1.1 代谢组学发展的时代背景	1
1.1.2 代谢组学研究现状	4
1.1.3 代谢组学与系统生物学	8
1.2 代谢组学的研究方法	8
1.2.1 样品采集与制备	10
1.2.2 代谢组数据的采集	10
1.2.3 数据分析平台	12
1.2.4 代谢组学数据库	13
1.3 代谢组学的应用	14
1.3.1 药物研发	14
1.3.2 疾病研究	15
1.3.3 植物代谢组学	16
1.3.4 微生物代谢组学	17
1.3.5 代谢组学与中医药现代化	17
1.4 代谢组学发展展望	18
参考文献	22
第 2 章 气相色谱-质谱技术在代谢组学中的应用	29
2.1 气相色谱-质谱联用技术	29
2.1.1 GC-MS 的工作原理	30
2.1.2 全二维气相色谱-飞行时间质谱联用技术 (GC×GC-TOFMS)	31
2.1.3 质谱谱图库	34
2.2 基于 GC-MS、GC×GC-TOFMS 的代谢组学技术平台	34
2.2.1 样品的采集与制备	34
2.2.2 数据采集	36
2.2.3 基于 GC-MS、GC×GC-TOFMS 技术的代谢组学数据处理	38
2.2.4 代谢物的结构鉴定	41
2.3 代谢组学应用	42
2.3.1 GC-MS 在代谢组学研究中的应用	42
2.3.2 全二维气相色谱在代谢组学研究中的应用	42
2.4 发展与展望	45

参考文献	46
第 3 章 液相色谱及液相色谱-质谱联用技术在代谢组学研究中的应用	49
3.1 液相色谱及液相色谱-质谱联用技术	49
3.1.1 液相色谱	49
3.1.2 液相色谱-质谱联用	50
3.1.3 基于液相色谱及液相色谱-质谱联用技术的代谢组学方法的优势和发展趋势	52
3.2 用于极性小分子分析的二维液相色谱分离系统	55
3.3 液相色谱及液相色谱-质谱联用技术在代谢组学研究中的应用实例	58
3.3.1 液相色谱代谢轮廓分析方法用于肝脏疾病严重程度的诊断	59
3.3.2 基于液相色谱-质谱联用技术的代谢组学方法用于慢性乙型肝炎的急性发作疾病研究	63
3.3.3 全二维亲水作用色谱质谱联用技术用于极性复杂样品中结构类似组分的分离和鉴定	74
参考文献	84
第 4 章 超高效液相色谱-质谱及在代谢组学中的应用	87
4.1 超高效液相色谱的理论基础	87
4.1.1 van Deemter 方程	87
4.1.2 超高效液相色谱仪器	88
4.2 超高效液相色谱与 HPLC 分析性能的比较	88
4.2.1 超高效液相色谱与 HPLC 色谱分离性能的理论比较	88
4.2.2 超高效液相色谱、HPLC 与质谱联用性能的比较	91
4.3 超高效液相色谱在代谢组学中的应用	95
4.3.1 超高效液相色谱-质谱用于代谢组学研究的优势	95
4.3.2 超高效液相色谱与质谱联用于代谢组学研究实例	97
参考文献	103
第 5 章 毛细管电泳-质谱联用技术在代谢组学研究中的应用	105
5.1 毛细管电泳-质谱联用技术简介	105
5.1.1 同轴液体鞘流	106
5.1.2 无鞘接口	106
5.1.3 液体连接	106
5.2 基于毛细管电泳-质谱联用技术的代谢组学平台	106
5.2.1 阳离子代谢物 CE-MS 分析方法	107
5.2.2 阴离子代谢物 CE-MS 分析方法	107
5.2.3 多价阴离子代谢物 CE-MS 分析方法	109
5.3 毛细管电泳在代谢组学中的应用研究	112
5.3.1 微生物菌株细胞提取物的代谢分析	112
5.3.2 植物细胞提取物的代谢分析	115
5.3.3 疾病诊断和生物标志物发现	117
5.3.4 食品安全	118

5.3.5 细胞基因和蛋白质功能研究	118
参考文献	119
第6章 核磁共振在代谢组学研究中的应用	121
6.1 核磁共振波谱分析原理	121
6.1.1 核磁共振概论	121
6.1.2 核磁共振波谱仪	123
6.1.3 核磁共振氢谱	123
6.1.4 核磁共振碳谱	125
6.1.5 ^{31}P 及 ^{15}N 核磁共振	126
6.2 代谢组学中的核磁共振波谱分析方法	127
6.2.1 样品的准备	127
6.2.2 实验数据的获得	128
6.2.3 数据分析	132
6.3 核磁共振在代谢组学研究中的应用	133
6.3.1 在药物毒理学评价研究中的应用	134
6.3.2 在人体代谢和动物代谢研究中的应用	135
6.3.3 在病理和生理研究中的应用	136
6.3.4 在环境监测方面的应用	136
6.3.5 在植物化学研究中的应用	137
6.3.6 在微生物系统研究中的应用	137
6.4 展望	137
参考文献	138
第7章 代谢组学研究中常用的化学计量学方法	142
7.1 数据预处理方法	143
7.1.1 原始数据矩阵的获得	143
7.1.2 自变量筛选	144
7.1.3 数据的标度化及滤波	144
7.2 常用的模式识别方法	146
7.2.1 无监督的模式识别方法	146
7.2.2 有监督的模式识别方法	152
7.3 数据库及专家系统	157
参考文献	159
第8章 类脂和类脂组学	161
8.1 类脂分子的结构和功能	162
8.1.1 非极性类脂	162
8.1.2 极性类脂	163
8.1.3 类脂代谢物	166
8.2 类脂的分子生物学 ^[19]	166
8.2.1 表面化学和界面催化	167

8.2.2	膜类脂分子是信号转导分子的前体	167
8.2.3	类脂-蛋白质相互作用	167
8.3	类脂研究中的分析方法	168
8.3.1	类脂的提取方法	168
8.3.2	类脂的分析方法	169
8.4	ESI-MS 在类脂分析中的定量方法	171
8.4.1	内标化合物的选择	171
8.4.2	MS 在定量分析类脂分子中存在的问题	172
8.4.3	ESI-MS 在类脂研究中的定量方法	172
8.5	类脂组学及其应用	178
8.5.1	类脂组学	178
8.5.2	类脂组学中类脂分子识别的策略	179
8.5.3	基于质谱技术的几种类脂组学方法	180
8.5.4	类脂组学在微生物研究中的应用	181
8.5.5	类脂组学在宿主-病原体相互作用研究中的应用	182
8.5.6	类脂组学在 $n-3$ 多不饱和脂肪酸影响细胞膜微区域研究中的应用	182
8.5.7	类脂组学在疾病和药物开发研究中的应用	183
8.5.8	类脂组学在功能研究中的应用	187
8.5.9	类脂组学在药物滥用研究中的应用	188
8.5.10	与类脂组学相关的数据库	189
	小结	190
	参考文献	190
第9章	代谢组学在疾病分型和标志物发现研究中的应用	195
9.1	引言	195
9.2	代谢组学用于心血管疾病严重程度的诊断	197
9.3	基于修饰核苷的代谢组学用于肝脏疾病的研究	201
9.3.1	代谢物的靶标分析	202
9.3.2	顺二醇类代谢物的轮廓分析	203
9.3.3	潜在标志物的寻找	205
9.4	代谢组学用于恶性肿瘤的研究	206
9.5	代谢组学在其他疾病分型和标志物发现研究中的应用	208
	参考文献	210
第10章	修饰核苷在癌症诊断和随访中的应用	212
10.1	概论	212
10.1.1	癌症诊断的方法概况	212
10.1.2	肿瘤标志物的概况	214
10.1.3	修饰核苷的来源及在癌症患者中修饰核苷增高的根源	218
10.1.4	修饰核苷作肿瘤标志物的可行性	221
10.2	体液中修饰核苷的分析和数据处理	222

10.2.1	体液中修饰核苷的分析	222
10.2.2	多变量可视化数据分类的方法	226
10.3	尿中核苷在恶性肿瘤诊断、治疗随访中的应用	239
10.3.1	尿中修饰核苷分布模式的建立	239
10.3.2	核苷作为肿瘤标志物在多种肿瘤中的应用研究	246
10.3.3	良性、恶性肿瘤尿中修饰核苷排放差异的研究	269
10.3.4	修饰核苷用于检测手术和治疗效果的评价	271
10.3.5	化疗对尿中核苷排放水平的影响	273
10.3.6	尿中核苷作肿瘤标志物实际应用范例	274
	参考文献	276
第 11 章	代谢组学在糖尿病研究中的应用	281
11.1	基于 GC 的有机酸轮廓分析对 2 型糖尿病研究	282
11.1.1	实验部分	282
11.1.2	2 型糖尿病的有机酸轮廓分析	283
小结	286
11.2	尿样“全”指纹分析及应用于科素亚对糖尿病的疗效评价	286
11.2.1	实验部分	286
11.2.2	科素亚对糖尿病的疗效	287
小结	292
11.3	2 型糖尿病中血浆磷脂代谢轮廓分析和标记物的识别	293
11.3.1	实验部分	293
11.3.2	2 型糖尿病中血浆磷脂代谢轮廓和标志物	294
小结	303
11.4	血清中脂肪酸的靶标分析方法用于 2 型糖尿病患者与正常人的区分	304
11.4.1	实验部分	304
11.4.2	结果与讨论	307
小结	311
	参考文献	311
第 12 章	代谢组学在营养学研究中的应用	313
12.1	简介	313
12.2	代谢组学全组分分析在营养学中的应用	314
12.2.1	代谢组学用于食物和营养对机体影响的评价	315
12.2.2	肠道菌群与宿主的相互作用	320
12.3	代谢轮廓分析及代谢物靶标分析在营养学中的应用	322
12.4	代谢组学在营养学研究中的应用展望	323
	参考文献	323
第 13 章	代谢组学与药物研究	326
13.1	药物开发现状	326
13.2	代谢组学在药物研究中的应用概况及特点	328

13.3	代谢组学在动物模型评估方面的应用	330
13.4	代谢组学在药物安全性研究中的应用	333
13.4.1	代谢组学在药物毒性研究中的应用概况及其特点	333
13.4.2	COMET 计划介绍	336
13.5	代谢组学在中药研发中的应用	339
13.5.1	代谢组学在中药质量研究中的应用	340
13.5.2	代谢组学在中药药效研究中的应用	347
13.5.3	代谢组学在中药安全性研究中的应用	350
	参考文献	351
第 14 章	代谢组学在植物研究中的应用	354
14.1	功能基因组时代的植物代谢组学	354
14.1.1	植物功能基因组学	354
14.1.2	植物代谢组学	354
14.1.3	植物代谢工程简介	356
14.2	植物代谢组的分析	357
14.2.1	植物样品的保存和预处理方法	357
14.2.2	植物代谢组的分析检测	364
14.2.3	代谢组学研究的数据处理	365
14.3	植物代谢组学的应用	366
14.3.1	植物代谢指纹图谱及轮廓分析应用实例	366
14.3.2	植物代谢分析应用实例——多维色谱/联用技术	376
14.3.3	转基因植物及受环境影响植物代谢分析应用实例	379
14.4	植物代谢组学的未来——主要问题与发展前景	384
	参考文献	385
第 15 章	代谢组学在微生物学中的应用	387
15.1	概述	388
15.2	微生物代谢组学实验设计的基本原则	388
15.2.1	代谢组学的优势	388
15.2.2	实验设计原则	389
15.2.3	分析平台	393
15.3	代谢组学在工业微生物学中的应用	398
15.4	代谢组学在医学微生物学中的应用	402
15.5	代谢组学在环境微生物学中的应用	406
15.6	展望	407
	参考文献	408
第 16 章	代谢组学在环境科学中的应用	410
16.1	简介	410
16.2	代谢组学用于脊椎动物环境毒物暴露及其代谢的研究	411
16.2.1	环境中重金属污染对脊椎动物的毒性作用研究	411

16.2.2 环境中有机污染物对脊椎动物的毒性作用研究	414
16.3 代谢组学用于无脊椎动物环境毒物暴露及其代谢的研究	417
16.4 环境代谢组学的发展前景	420
参考文献	420

第1章 绪 论

随着人类基因组测序工作的完成, 基因功能的研究逐渐成为热点, 随之出现了一系列的“组学”研究, 包括研究转录过程的转录组学 (transcriptomics)、研究某个生物体系中所有蛋白质及其功能的蛋白质组学 (proteomics) 及研究代谢产物的变化及代谢途径的代谢组学 (metabolomics 或 metabonomics) (图 1-1)。

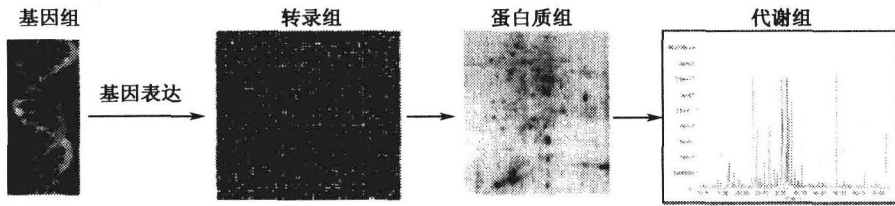


图 1-1 组学时代 4 种最重要的组学

对基因组 (genome)、转录组 (transcriptome)、蛋白质组 (proteome) 及代谢组 (metabolome) 的研究分别对应基因组学、转录组学、蛋白质组学及代谢组学

代谢组学是众多组学中的一种, 是随着生命科学的发展而发展起来的。与其他组学不同, 代谢组学是通过考察生物体系 (细胞、组织或生物体) 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, 来研究生物体系的一门科学^[1]。所谓代谢组 (metabolome) 是基因组的下游产物也是最终产物, 是一些参与生物体新陈代谢、维持生物体正常功能和生长发育的小分子化合物的集合, 主要是相对分子质量小于 1000 的内源性小分子。代谢组中代谢物的数量因生物物种不同而差异较大, 据估计, 植物王国中代谢物的数量在 200 000 种以上, 单个植物的代谢物数量在 5000~25 000, 甚至简单的拟南芥 (*Arabidopsis thaliana*) 也产生约 5000 种代谢产物, 远远多于微生物中的代谢产物 (约 1500 种) 和动物中的代谢产物 (约 2500 种)^[2]。实际上, 在人体和动物中, 由于还有共存的微生物代谢、食物及其代谢物本身的再降解, 到目前为止, 还不能估计出到底有多少种代谢产物, 浓度分布范围有 7~9 个数量级。因此对代谢组学的研究, 无论从分析平台、数据处理及其生物解释等方面均面临诸多挑战。本章对代谢组学发展的历史、国内外现状、研究方法、典型应用领域及研究热点等给予了介绍。

1.1 代谢组学简介

1.1.1 代谢组学发展的时代背景

生命科学是研究生命现象、生命活动的本质、特征和发生、发展规律, 以及各种生物之间和生物与环境之间相互关系的科学。自从 1953 年 Watson 和 Crick 建立了 DNA

双螺旋结构模型后，生命科学研究的面貌便焕然一新。在此基础上发展的分子生物学使得生命的基本问题，如遗传、发育、疾病和进化等，都能从分子机制上得到诠释。生物学研究进入了对生命现象进行定量描述的阶段。分子生物学的飞速发展极大地推动了人们从分子组成水平对生物系统进行深入的了解。基因组计划向人们展示了包括大肠杆菌、酵母、线虫、果蝇、小鼠等模式生物以及人类的所有遗传信息的组成，生命的奥秘就存在于这些序列中。技术上的突破使得基因组数据的获得已经不再是生命科学的难点。人类基因组计划的基本完成标志着后基因组时代的到来，在这一时期，基因组功能分析成为生命科学的主要任务，核心思想是以整体和联系的观点来看待生物体内的物质群，研究遗传信息如何由基因经转录向功能蛋白质传递，基因功能如何由其表达产物蛋白质以及代谢产物来体现。继基因组 (genome) 后、转录组 (transcriptome)、蛋白质组 (proteome) 等相继出现，并相应形成“omics”学说，如转录组学 (transcriptomics)、蛋白质组学 (proteomics) 等。但是基因与功能的关系是非常复杂的，还不能用转录组、蛋白质组来表达生物体的全部功能。生物体内存在着十分完备和精细的调控系统以及复杂的新陈代谢网络，它们共同承担着生命活动所需的物质与能量的产生与调节。在这一复杂体系中，既有直接参与物质与能量代谢的糖类、脂肪及其中间代谢物，也有对新陈代谢起重要调节作用的物质。这些物质在体内形成相互关联的代谢网络，基因突变、饮食、环境因素等都会引起这一网络中某个或某些代谢途径的变化，这类物质的变化可以反映机体的状态。起调节作用的代谢物，从生理功能上来说包括神经递质、激素和细胞信号转导分子等，从化学组成上来说包括多肽、氨基酸及其衍生物、胺类物质、脂类物质和金属离子等，这些调节物质绝大部分都是小分子物质，在植物与微生物中还存在着大量的次生代谢产物。这些分子广泛分布于体内，对多种生理活动都具有普遍和多样的调节作用，仅微量存在就能够发挥很强的生物效应。不同活性的分子或协同、或拮抗、或修饰而相互影响，在生物学效应以及信号转导和基因表达调控上形成复杂的网络，承担着维持机体稳态的重要使命，是神经内分泌和免疫网络调节的物质基础和自稳态调节的最重要成分。转录组、蛋白质组的研究很难涵盖这些非常活跃而且非常重要的生命活性物质，然而对这类物质的生理和病理生理学意义如果不能充分认识，就不可能真正阐明生命功能活动的本质。传统研究方法是以生理学和药理学实验方法为主，缺乏高通量的研究技术，难以建立生物小分子物质复杂体系的研究模式。在这种情况下，代谢组 (metabolome) 和代谢组学 (metabolomics 或 metabonomics) 应运而生了，并成为系统生物学的一个重要突破口^[3]，代谢处于生命活动调控的末端，因此代谢组学比基因组学、蛋白质组学更接近表型。

从广义的代谢组学的意义上来说，代谢组学的历史是相当长的，很早以前人们就已经对生物样品中的某些靶标化合物进行分析以了解生命机体的状态。目前代谢组学所采用的一些技术平台，如 NMR 和色谱技术以及质谱技术也有比较长的应用历史。严格意义上的代谢组学 (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) 从提出到现在只有短短数年的时间。现在一般认为代谢组学源于代谢轮廓 (metabolic profiling) 分析，在代谢轮廓分析中体现了代谢组学的“尽可能多地分析生物样本中的代谢产物”这一理念的萌芽。在这里，我们对从代谢轮廓分析发展到代谢组学这一过程^[4] (图 1-2) 做一简单的介绍。

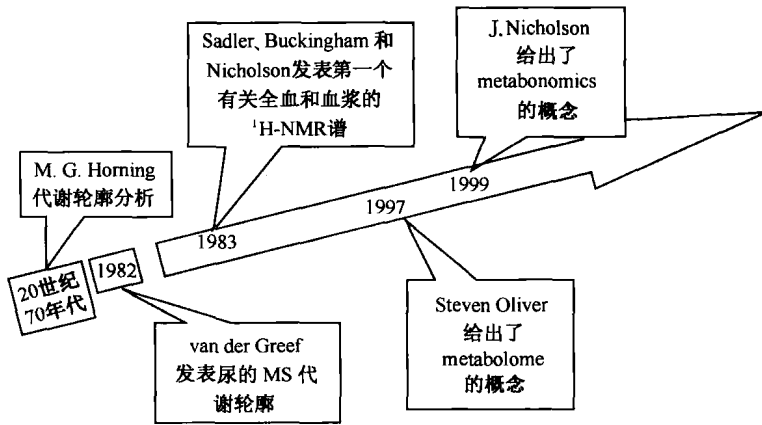


图 1-2 代谢组学的发展历史

早在 20 世纪 70 年代初，Baylor 医学院就发表了有关代谢轮廓分析方面的论文，在他们的工作中采用了 GC-MS 的方法对多种类固醇、有机酸以及尿中药物的代谢物进行了分析，并将这种多组分分析的方法称为代谢轮廓分析，开创了对复杂样品进行代谢轮廓分析的先河。此后代谢轮廓分析广泛应用于血、尿等生物样本中代谢物的定性与定量分析，以对疾病进行筛选和诊断。在临床上使用 GC-MS 的方法来诊断疾病的方法一直沿用到今天。紧接着，人们把重点主要放在分析的自动化上，并将 GC 的方法用于其他类型化合物的分析。进入 20 世纪 80 年代，人们开始使用高效液相色谱和核磁共振的技术来进行代谢轮廓的分析，如 1982 年，荷兰应用科学研究所（TNO）的 van der Greef^[5]在国际上首先采用质谱对尿中代谢指纹进行研究。1983 年，Sadler、Buckingham 和 Nicholson 发表了第一个有关全血和血浆的 $^1\text{H-NMR}$ 谱^[6]。在 1986 年，色谱杂志 *Journal of Chromatography* 发表了一期有关代谢轮廓（metabolic profiling）分析的专辑。进入 90 年代，代谢轮廓分析技术一直平稳发展，每年都有 10~15 篇的论文发表，不过这一时期人们的目标更多地集中于某些特定的标靶化合物上。在 90 年代初，Sauter 等人用基于 GC-MS 代谢轮廓分析的方法研究了不同除草剂对大麦的影响，这种用代谢轮廓分析来研究各种因素对生物功能的影响的研究思路随即被人们认可。1997 年，Steven Oliver 研究小组提出了通过对代谢产物的数量和定性来评估酵母基因的遗传功能及其冗余度，并率先提出了代谢组的概念^[7]。1999 年，J. Nicholson 等提出 metabonomics 的概念^[8]，并在疾病诊断、药物筛选等方面做了大量卓有成效的工作^[1,9~11]。接着，德国的 Max-Planck-Institut 的科学家们开始了植物代谢组学的研究^[12]，使代谢组学得到了极大的充实。

代谢组学的特点为：

- (1) 关注内源化合物。
- (2) 对生物体系中的小分子化合物进行定性定量研究。
- (3) 上述化合物的上调和下调指示了与疾病、毒性、基因修饰或环境因子的影响。
- (4) 上述内源性化合物的知识可以被用于疾病诊断和药物筛选。

与转录组学和蛋白质组学比较，代谢组学有以下优点^[13]：

- (1) 基因和蛋白质表达的微小变化会在代谢物上得到放大，从而使检测更容易。