



· 导读版 ·

The Handbook of Metabonomics and Metabolomics 代谢组学手册

John C. Lindon, Jeremy K. Nicholson, Elaine Holmes



科学出版社
www.sciencep.com

The Handbook of Metabonomics and Metabolomics

代谢组学手册

Edited by

John C. Lindon

Jeremy K. Nicholson

Elaine Holmes

Imperial College London

科学出版社

北 京

图字:01-2008-0364 号

This is an annotated version of

The Handbook of Metabonomics and Metabolomics

John C. Lindon, Jeremy K. Nicholson and Elaine Holmes

Copyright © 2007, Elsevier BV.

ISBN 13:978-0-44-452841-4

ISBN 10:0-44-452841-5

All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

AUTHORIZED EDITION FOR SALE IN P. R. CHINA ONLY

本版本只限于在中华人民共和国境内销售

图书在版编目(CIP)数据

代谢组学手册=The Handbook of Metabonomics and Metabolomics:英文/(英)林顿(Lindon,J.)编著. —北京:科学出版社,2008

ISBN 978-7-03-020777-7

I. 代… II. 林… III. 代谢-研究-英文 IV. Q493.1

中国版本图书馆 CIP 数据核字(2007)第 202464 号

责任编辑:田慎鹏 贾明月/责任印制:钱玉芬/封面设计:耕者设计工作室

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>

北京佳信达艺术印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2008年1月第 一 版 开本:787×1092 1/16

2008年1月第一次印刷 印张:36 1/4 插页:6

印数:1—1 500 字数:859 000

定价:128.00 元

如有印装质量问题,我社负责调换

导 读

自 20 世纪末基因组学概念提出以后,各种冠以“组学(-omics)”的概念如雨后春笋般地涌现出二百多种。同时也催生了关注生物分子之间途径和网络的动态变化和关系的“系统生物学(systems biology)”,及整合各种生物信息、从全局角度反映生物个体的遗传背景和所处环境交互作用于相应个体的结果的“全局系统生物学(global systems biology)”。这是生命科学向纵深发展的必然趋势,同时也是理解生命过程的必要途径。生物界的“中心法则”指出,基因转录出信使 RNA 再翻译成蛋白质(虽然有时可能仅是一条未经修饰过的多肽链),后者发挥各种生物学功能,其中绝大多数是酶的形式,催化底物生成相应代谢产物。基因组学的研究对象是一个生物个体内的全部基因;相应地,蛋白质组学是针对所有蛋白质;很自然,对所有代谢物(包括代谢中间体和终产物)的研究就是代谢组学的范畴。对于代谢组学概念,目前国际上存在两种观点,一种以英国帝国理工学院的 Jeremy K. Nicholson 等为代表,他们在 20 世纪 90 年代将“生命体系对病理生理刺激或遗传改造所产生的动态、多指标代谢响应的定量测定”定义为 metabonomics。另一种是 21 世纪初由原在德国、现在美国加州大学戴维斯分校的 Oliver Fiehn 提出的“全面、定量分析生物体系中所有代谢物”的 metabolomics。虽然字面上有一字母之差,但是两者的研究对象都是代谢物(metabolites),因此我们说本质上 metabonomics 和 metabolomics 没有区别。也有人试图从前者主要将核磁共振技术应用于健康疾病领域,多以来自动物的样品为研究对象;后者主要应用质谱联用技术,在植物、微生物领域进行研究,而将两者进行区分,其实也没有什么学术意义。一种学科可以应用不同的技术;同样,同一技术也可以服务于不同学科。因此国内读者大可不必为 metabolomics 和 metabonomics 的解释和区分而大伤脑筋,完全可以都称之为“代谢组学”。

《代谢组学手册》是英国帝国理工学院的 John C. Lindon、Jeremy K. Nicholson 和 Elaine Holmes 等人联合国际上代谢组学领域内的知名研究者共同撰写的一部代谢组学专著。以 Nicholson 为首的研究小组是国际上以核磁共振技术为主导进行代谢组学研究的前驱。他们利用该技术在疾病鉴别诊断、药物毒性研究、肠道菌群与宿主之间的关系、饮食环境等因素对机体机能状态影响等方面进行了广泛的探索,已在 *Nature* 等数十种杂志上发表相关论文 200 多篇。由于该研究组的卓越成就,美国辉瑞等六家国际知名制药公司与该组联合进行了一项 COMET 计划,旨在利用代谢组学技术对体液的检测快速筛查药物毒性,这标志着代谢组学开始被工业界所接受和认可。目前代谢组学的应用已经涉及到生命科学研究的各个领域。

《代谢组学手册》一书共分 19 章,图文并茂、内容丰富、资料详实,大部分描述都围绕着国际上代谢组学方面具有代表性的研究而展开和论证,各章之间既遵循同一主线、交相呼应,又独立成文、自成体系。全书文献引用全面、广泛且重点突出、代表性强。为了便于读者有选择性地阅读,本书可概括为如下三部分:

第一部分：代谢组学分析技术及原理。这部分内容主要涉及什么是代谢组学、为什么需要代谢组学、代谢组学与其他组学的联系、应用领域和学科展望等（第1、19章）；相应地，本部分内容还就该领域内常用的核磁共振和质谱联用技术的原理和优势及其应用于代谢组学时需要注意的一些问题作了初步比较（第3~5章）。该部分还简要涉及代谢物、代谢途径、代谢通量到代谢网络之间的关联等内容（第2章）。

虽然现有的分析技术都可依研究对象和目的不同而应用到代谢组学中，但目前的代谢组学研究中应用最多的还是核磁共振和（气相、液相）色谱-质谱联用这两大主流技术。核磁共振技术的优势在于能够对样品实现无创性、无偏向的检测，具有良好的客观性和重现性，样品不需要繁琐预处理，具有较高的通量和较低的单位样品检测成本。此外， $^1\text{H-NMR}$ 对含氢化合物均有响应，能完成样品中大多数化合物的检测，满足代谢组学中对尽可能多的化合物进行检测的目标。它的缺点是检测灵敏度相对较低（采用现有成熟的超低温探头技术，其检测灵敏度在纳克级水平）、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。质谱检测具有较高的灵敏度和专属性，可以实现对多个化合物的同时快速分析与鉴定。气质联用方法的主要优点包括较高的分辨率和检测灵敏度，并且有可供参考的标准谱图库用于代谢产物定性。但是气相色谱不能直接得到体系中难挥发的大多数代谢组分的信息，对于挥发性较低的代谢产物需要衍生化处理，预处理过程繁琐。液相色谱质谱联用避免了繁杂的样品前处理，由于其较高的灵敏度和较宽的动态范围，已被越来越多地用于代谢组学研究。该书通过多种实例阐述了这些技术各自的优缺点，其中还涉及了高分辨魔角旋转核磁共振技术和超高效液相色谱等新近出现的分离分析技术的原理和应用。这一部分对于初次接触代谢组学并且缺乏相应分析化学和仪器分析技术基础的读者来说是个很好的入门材料。其中对包括代谢组学在内的各种组学技术与系统生物学、全局系统生物学之间的关系和作用等的阐述，可以帮助读者深入浅出地了解生命科学领域内的研究为什么和如何采取系统的策略。

第二部分：代谢组学中的数据和处理。这部分内容主要讨论代谢组学中所用到的化学计量学方法（第6章）、非线性数据处理技术（第7章）等多变量统计分析方法和模式识别技术，以及代谢组学分析中所需注意的实验和报告的标准化和相关数据库等问题（第8章）。具有生物统计学和生物信息学背景的读者，很容易对相关章节的内容做到触类旁通；初次接触该领域的读者如能仔细研读，也将对本书后续内容的理解更加得心应手。

代谢组学是一种高通量分析技术，一次分析会产生大量数据，由目前分析仪器上导出的原始数据大多不能直接用于多变量分析（主要是模式识别），因而需对数据进行必要处理，将其转变为适合于多变量分析的数据形式。主要的数据预处理包括滤噪、重叠峰解析、峰对齐、峰匹配、标准化和归一化等。在实际操作中，这些步骤并不都同时需要，而是根据实际情况，只做几种预处理。最后用于模式识别的数据多为二维矩阵形式。从这些庞杂的数据中提取有价值的信息并给予科学合理的解释使之用于指导所研究的问题，这个过程比实际样品分析的本身更具挑战性，对这一问题的认识应该从代谢组学的实验设计阶段就加以考虑。目前的数据分析过程中应用的主要模式识别技术，包括非监督学习方法（如PCA）和有监督学习方法（如PLS-DA），许多生产厂商在仪器开发的同时也推出配套的代谢组学分析软件以方便用户使用。另外，为了便于交流和协

作，制定相应的标准用于规范数据的获取、分析和比较也是十分必要的，否则任何数据库的存在价值都会大打折扣。第一、二两部分是构成和理解代谢组学的核心。

第三部分：代谢组学的应用实例。本书用近 50% 的篇幅来描述代谢组学在生物医药-临床前药物发现和开发、临床药物研发（第 9、10 章）、疾病健康-群体研究、心血管病、癌症、先天性代谢疾病、糖尿病和肥胖、关节炎、神经性疾病、肝病、器官移植、传染性疾病（第 11~15、17 章）、植物科学（第 16 章）和环境毒理（第 18 章）等方面的应用和取得的成就。

目前，系统生物学的数据绝大多数来自于基因组学、转录组学、蛋白质组学和代谢组学。基因组学研究反映了什么是可以发生的，转录组学则反映的是将要发生的，蛋白质组学指出的是赖以发生的，而只有代谢组学才真正反映出业以发生的。作为表型区分、差异表述和标志物发现的有力工具，与转录组学和蛋白质组学等其他组学相比，代谢组学具有以下优点：①基因和蛋白表达的微小变化会在代谢物水平得到放大；②代谢组学的研究不需进行全基因组测序或建立大量表达序列标签的数据库；③代谢物的种类远少于基因和蛋白质的数目；④生物体液的代谢物分析可反映机体系统的生理和病理状态。代谢组学的另一优势还在于实验成本相对低廉。因为基因表达、蛋白质翻译和代谢物变化等分子事件的发生不是遵循一成不变的线性关系，还要受到诸多外界条件的影响，所以目前的生命科学研究与全局系统生物学的要求还有相当距离。

《代谢组学手册》是一本集基本理论和实际应用于一体的极有价值的代谢组学学科专著，书中对学科的发展现状、面临问题、应用前景、未来趋势和学科本身的价值都作了客观、科学的描述。本书既可作为从事代谢组学研究的专业人员的工具书，也可作为相关领域研究生的教材使用。对代谢组学感兴趣或具备一定的生物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性地阅读部分章节。目前国际上关于代谢组学方面的专著已经开始不断推出，在平台技术和应用领域方面各有侧重。比如 *Metabolomics: The Frontier of Systems Biology* (M. Tomita, T. Nishioka 著, 2005) 在公共数据库和虚拟细胞系统的描述方面好于本书；*Metabolomics: Methods and Protocols* (Wolfram Weckwerth 著, 2006) 增加了毛细管电泳与质谱联用和代谢通量分析的内容；*Concepts in Plant Metabolomics* (Basil J. Nikolau, Eve Syrkin Wurtele 著, 2007) 侧重于植物学方面的应用。国内目前尚无类似出版物出现（我们自己正在写的《代谢组学：原理和应用》的专著，预计春节前可交给科学出版社在 2008 年出版）。

中国科学院大连化物所、天津药物研究院、军事医学科学院、上海交大等是国内最早从事代谢组学研究的机构。代谢组学作为系统生物学的分支学科目前还没有写入教科书并走入学生的课堂，但是代谢组学的潜在价值已经引起国内相关科研机构和政府部门的关注。科技部和国家自然科学基金委在“十一五”的多个项目中设立了代谢组学研究专题。2006 年 9 月 13~14 日，第 284 次香山科学会议在中国科学院大连化物所召开，其主题就是“医学代谢组学”。多数学者认为，代谢组学与中医药现代化目标的结合将是我国民族医药发展历程中的战略性突破，不仅如此，生命科学领域内的诸多学科都可或多或少得益于代谢组学。

代谢组学是进行生物表型研究的主要手段之一，是系统生物学研究不可或缺的部分，在疾病分型、药物毒性评价、植物基因功能和表型的研究等方面取得了极大的成

功。但从总体来看，它仍然处于发展阶段，在方法学和应用两方面均面临着极大的挑战，需要其他学科的配合和交叉。

在平台技术和方法学研究方面，生物样本的复杂性使得代谢组学研究对分析技术的灵敏度、分辨率、动态范围和通量提出了更高的要求。代谢组学研究的深入得益于分析技术的不断发展，如高分辨质谱、超高效液相色谱/质谱、毛细管液相色谱/质谱、多维色谱质谱联用技术和多维核磁共振技术等的使用。生物标记物的结构鉴定也是目前代谢组学研究的重点和难点问题之一，由于缺乏标准的可通用的质谱数据库，一定程度上制约了基于 LC-MS 的技术在代谢组学研究中的应用。理论上讲，LC-MS-NMR 可提供较好的关于组分结构的信息，但仪器复杂，操作繁琐，灵敏度和通量急需改进和提高。

另外由于代谢物仅占细胞的很少一部分（肠杆菌科细菌中，代谢物仅占菌体干重的 3%~5%），有些代谢物极不稳定（如 ATP 的“半寿期”仅为 0.1 秒左右），因此取样技术也需要不断探索和改进。再者代谢物数据库的建立也明显落后于其他相关组学，这不仅与学科发展历程有关，更受化合物结构解析技术的局限。功能完善的代谢产物数据库的构建及代谢组学研究的标准化等问题越来越受到关注。

体液代谢组学研究与细胞生物学和动物模型数据及知识的整合、代谢组学数据与其他组学数据的整合、代谢组学与计算生物学的整合以及构建代谢网络和代谢流动态变化的数学模型等，在代谢组学研究领域内有着广泛的发展前景，也是下一步研究的重点。

在应用方面，代谢组学要生存和发展，必然要有特色，从表型着手回答其他组学不能回答的生物问题。与其他组学一样，如何克服瓶颈，从大量的代谢产物中找出特异性的生物标记物（特别是低丰度的标记物），是决定此技术能否在药物和临床领域广泛应用的一个重要因素。代谢组学在个性化药物治疗和中医药的研究上也将有灿烂的前景。

本书无疑是帮助你克服代谢组学研究瓶颈和面对上述挑战的最好参考书。

许国旺

2007 年 12 月 8 日于大连

前 言

我们解析生理过程和分析其相应组分能力提高的同时，生命赖以存在的生物系统的奥秘也在不断展现。近来生命科学的进展主要集中于生物大分子，即 DNA、RNA 和蛋白质。许多人曾乐观地认为，通过了解染色体的组成和基因的表达差异，将会使我们对疾病有更加深入的认识，因而很快就能实现预测某个体在其一生中患某疾病的可能性，从而更有效地调整防治措施。

我相信这种乐观是有其坚实基础的，但问题是对大分子行为的研究只能提供观测复杂有机体（如人类）对环境变化应答所需的部分信息。如若力求完整，则要能考察生物学状态的动态标记物——反映有机体整体功能的实时信号，以资诊断与预测，这方面正是现今代谢组学的强大用武之地。

“代谢组学”（metabonomics）一词是以 Jeremy Nicholson 为首的帝国理工学院的一批科学家首先提出来的。虽然利用核磁共振技术对体液的相关研究可追溯到 20 世纪 80 年代中期，但直至今日，该技术平台才被认为是一种重要的研究工具。代谢组学基于如下实证：将代谢物代谢池的波动与整体代谢途径功能的变化相关联，可以预测诸如药物应答强度、潜在毒性甚至病程进展等结果。

生物制药工业及相关政策机构所表现出的极大兴趣，是发展中的代谢组学概念及相关数学算法重要性的极好例证。鉴定有效的具有生理和临床意义的标志物，以之作为廉价、快捷的筛选对特定人群有效或有毒药物的方法，已成为大家的共识。事实上，由于候选分子开发成本及人们开发能力的提高，新药发现-开发环节中的消耗变成制药工业所面临的巨大挑战之一。

既要协助有价值的新药尽快到达急需的患者手中，又要最大程度地保障公众的安全，使得相应政策机构举步维艰。毫无疑问，任何能快速、经济、有效地预测药物对特定人群潜在毒性的工具，都会被仔细审视，看它们是否可作为药物审批流程标准中的一部分。

经济上的考虑也会提升代谢组学在临床上的潜在重要性，用只需几美元的无创检测试验替代通常较为昂贵的有创检测的可能性还是大大存在的。

代谢组学的价值不只限于人类卫生保健方面，事实上，相关技术可用于植物科学、动物健康和模式生物（酵母、细菌或体外细胞系）的开发。目前的代谢组学使得原本只作为一些（也许更多）可能性而存在的东西正在转变为现实。因而，出版一本全面的有关该学科的图书是很适时的。对于那些致力于该学科的人们，这只是一本参考书；对于那些首次涉足该领域的人们，这本书将打开一个蕴藏诸多应用可能的迷人新世界。

Richard B. Sykes FRS

于伦敦

（许国旺 译）

序

从 20 世纪 90 年代,尤其是人类基因组的测定开始,应用于分子生物学与生物化学的技术与方法发生了巨大变革。相应的,出现了一个颇具广泛性的观念转变,即遗传差异可以解释所有疾病过程,并由此产生新的诊断方法和更好的靶向治疗策略。自动微阵列方法对基因表达改变的检测的确带来了分子生物学的革命性变化,并催生了一个被称为“转录组学”的新概念。随后,主要以基于质谱方法分析鉴别某一系统中的蛋白质为代表的分析和鉴定能力的提高,导致了“蛋白质组学”的形成。然而,过去数年间,人们已经意识到了分子生物学的整体复杂性以及基因构成与环境因素之间的复杂交互作用,对这些相互作用的理解在蛋白质组学层次已是困难重重,更不可能从转录组学水平去揭示。

事实上,参与生物化学过程的小分子能提供大量有关被研究的生命系统功能和状态的信息,其中既涉及差异基因表达引起的,又包括人类、哺乳动物因生活方式、饮食不同所造成的差别。评价、监测这种变化的过程就定义为“metabonomics”。类似的策略在模式生物和植物系统中的应用被称为“metabolomics”。代谢组学始于 20 世纪 80 年代中期利用核磁共振技术对生物体液的检测研究,随后又与复杂数据的模式识别和多变量统计分析技术相结合。直到 1999 年 Jeremy Nicholson 和他的同事才把“生命体系对病理生理刺激或遗传修饰所产生的动态、多指标代谢响应的定量测定”正式定义为 metabonomics。2001 年,Oliver Fiehn 引入了定义稍有不同的 metabolomics 概念,即“全面、定量分析体系中的所有代谢物”的科学。尽管概念上有差异,但两者在基本理念与方法学上多有重叠,两个术语也被很多科学家和机构等同使用。本书中,我们尊重了编者们的使用偏好,对 metabolomics 和 metabonomics 不加区分。

通过同时监测多种分子的浓度(某些情况下包括分子动力学和区域化)变化,代谢分析为生物化学提供了一个数据密集型的研究手段。为有助于观察理解不同研究类别间的显著差异,多变量统计分析的深入使用十分必要,相应计算结果的解释将直接指向不同类别样品间的真实生化差异所在,继而鉴定出与研究对象相关的生物标志物。

这本书的主要目的是展示代谢组学的现状,对所用的分析与统计技术进行权威性诠释与指导,纵览和指明主要的应用领域。书中详细阐述了核磁共振与色谱质谱联用这两种主要的分析技术和广为采用的各种化学计量学与统计学方法。另外,也借此机会介绍了一些近年来在标准化设计与报告代谢实验方面的相关尝试。我们收录了大量生物学应用方面的文章,范围涵盖植物、模式生物研究,临床前与临床药物学研究以及人类疾病和流行病学研究。

我们非常感激所有作者付出的努力,感谢这些学识渊博并为本书提供最新最重要文献综述的人们。我们希望,同时也相信这种做法会从所探讨方法学的正反两方面为本书的主题提供一个平衡的观点。书中每章含有对应、全面的参考文献,以供进一步阅读。

代谢组学研究领域正在迅速扩大，但是我们相信现在是对代谢组学领域的成就进行回顾和科学界为未来的研究制定标准的好时机。我们的目的是让这本书成为那些正涉足该领域和已经开始进行代谢研究的科学家们的权威性参考著作。我们相信本书还有其重要的教育用途并可以为研究生所使用。这本书应会引起分析科学家们和工作在植物科学、环境科学和人类临床医学等诸多应用领域的科学家的兴趣。总之，我们相信这是一本教育性强、信息丰富、意义重大并可以给人以启迪的专著。

John C. Lindon

Jeremy K. Nicholson

Elaine Holmes

英国伦敦

2006 年 5 月

(许国旺 译)

有奖征集反馈意见



尊敬的读者：

科学出版社科爱森蓝文化传播有限公司（简称“科爱传播”）立足国际合作，致力于为科技专业人士提供优质的信息服务。我们很想通过自己的努力最大限度地满足您的需求，您的哪怕是一点点的建议和意见，都将成为我们改进工作的重要依据。

我们将在每年的6月份、12月份各一次从半年的参与者中抽取幸运者10名，幸运者可以从“科爱传播”的出版物中任选价值1000元的图书（10册以内）作为奖品（全部出版物信息可在我们的网站上查到）。

1. 您所购买的图书书名：《_____》

您于_____年__月__日在（通过）_____购买到此书。

你认为本书的定价：☐偏高 ☐合适 ☐偏低

你认为本书的内容有约____%对您有用。

2. 你认为我们出版物的质量：

内容质量（学术水平、写作水平）：☐很好 ☐较好 ☐一般 ☐较差

译介质量（翻译水平、文字水平）：☐很好 ☐较好 ☐一般 ☐较差

印制质量（印制、包装）：☐很好 ☐较好 ☐一般 ☐较差

3. 您所在的专业领域：_____

4. 在你获取专业知识和专业信息的主要渠道中，排在前三位的是：

1. _____ 2. _____ 3. _____

A.网络 B.期刊 C.图书 D.报纸 E.电视 F.会议 G.内部交流 H.其他：_____

5. 你还需要哪些类型的图书？

☐专著 ☐教材 ☐实验手册 ☐辞典工具书 ☐文集 ☐其他：_____

☐书摘（原版书摘编） ☐刊摘（国外学术期刊重要文献摘编）

6. 您还希望我们从国外引进哪些专业方向的图书（或期刊）？

7. 您建议采用何种引进形式？

- ☐ 翻译 ☐ 影印 ☐ 摘编（只从原书中选部分内容引进）
☐ 导读（原文影印加少量中文介绍） ☐ 注解（原文影印加大量中文介绍）

8. 您是否愿意与我们合作，参与编写、编译、翻译图书或其他科技信息？

9. 请列举您近两年看过的，您认为最有参考价值、对您帮助最大的 1~2 本书：

书名	著作者	出版社	出版日期	定价

10. 请列出 1~2 本（如果有的话）您近两年购买的由科学出版社出版的图书：

书名	出版日期	著作者	定价	购买地

11. 您还有什么别的意见、建议？（可另附纸）

● 请告诉我们您准确的地址和联系办法：

姓名：_____ 性别：_____ 生日：_____年__月__日

单位：_____ 职务/职称：_____

地址：_____

E-mail：_____ 电话：_____

传真：_____ 手机：_____

回寄地址（也可以通过 E-mail 反馈）：

100717 北京东黄城根北街 16 号 科学出版社 科爱传播中心 杨 琴（收）

联系电话：010-64006871；传真：010-64034056

E-mail: yangq@kbooks.cn, keai@mail.sciencep.com

（注：本反馈单复印有效，也可以在线下载：<http://www.kbooks.cn>）

Foreword

Biological systems which underpin life itself are continuously yielding their secrets as our ability to isolate processes and analyse their component parts improves. Recent advances have tended to focus on macromolecules – DNA, RNA and proteins – and there is considerable optimism that understanding the make up of our chromosomes and the differential expression of genes will lead to much better understanding of disease. It is also believed that we will soon be able to predict a person's likelihood of developing a disease during their lifetime and tailor curative or preventative treatment more effectively.

Whilst I believe that much of this optimism is well founded, it is clear that study of the behaviour of macromolecules will only give part of the information we need to observe the response of a complex organism, such as man, to a changing environment. To give a more complete picture, we need to be able to observe dynamic markers of biological status – real-time signals which reflect the integrated function of the organism in ways which allow diagnosis and prediction. This is where metabonomics is now demonstrating enormous potential.

Metabonomics was pioneered by a group of scientists now based at Imperial and headed by Jeremy Nicholson. Whilst the technological platform is only now starting to be recognised as a tool of major importance, the foundation studies involving NMR spectroscopy of biological fluids dates back to the mid-1980s. Metabonomics is based on the demonstration that correlation of changes in metabolite pool patterns with changes in the function of integrated biological pathways can lead to ways of predicting outcomes such as the magnitude of response to a drug, potential toxicity or even the path of a disease process.

That the concepts and mathematical algorithms being developed are of major significance is well evidenced by the keen interest being shown by the biopharmaceutical industry and regulatory agencies. Both sectors share the goal of identifying surrogate markers of biological or clinical outcome which are well validated and lead to speedier, cheaper ways of identifying which drug candidates may be efficacious or toxic and in which populations. In fact, the management of attrition in the drug discovery–drug development pipeline is one of the most challenging for the industry as development costs and our ability to produce candidate molecules escalate.

The regulators are caught in a difficult position – needing to help important new medicines reach needy patients quickly and also protecting the public at large. Any

tools which can help to predict potential toxicity in particular populations, in speedy cost-effective ways, will undoubtedly be scrutinised closely to see whether they should become a standard part of the drug approval process.

Economic considerations will also raise the potential importance of metabonomic tests in the clinical setting. Huge opportunities exist wherever non-invasive tests costing just a few dollars can be shown to replace invasive tests which are nearly always more expensive.

It would be wrong to leave the impression that metabonomics is only of value in the field of human healthcare. In fact, the technologies are applicable to plant science, animal health and development of model organisms for research (yeast, bacteria or *in vitro* cell systems). The state of the art of metabonomics is now such that some of these possibilities – and more – are verging on being realities. It is thus timely for a comprehensive book on the subject to appear in print. For those closely engaged in the subject, the book will serve as a definitive reference; for those seeking to learn for the first time, the book will open a fascinating new world containing a plethora of possibilities for application.

Richard B. Sykes FRS
London

Preface

Since the 1990s and particularly since the determination of the human genome, there have been dramatic changes in the scientific techniques and approaches used in molecular biology and biochemistry. The accompanying change in mindset has appeared to lead to a pervasive attitude that genetic differences might be able to account for all disease processes, and that this would lead to new diagnostic approaches and thence to much better targeted therapies. The revolution in molecular biology really took off with the availability of automated micro-array methods for detecting changes in gene expression, leading to the new discipline of transcriptomics. The subsequent expansion in the ability to assay and then identify, using mostly mass spectrometry-based methods, the proteins in a system has led to the term “proteomics” being coined. However, during the past few years the full complexity of molecular biology has been realised and the complex interactions between genetic make-up and environmental factors have now been recognised. It is now accepted that understanding of these interactions is impossible at the transcriptomic level and difficult at the proteomic level.

The reality is that the small molecules involved in biochemical processes provide a great deal of information on the status and functioning of a living system under study both from effects caused by changes in gene expression, and also by differences in life style and diet in humans and other mammals. The process of monitoring and evaluating such changes is termed “metabonomics”. A parallel approach has also been under way, mostly in model organisms and in plant systems, and that has led to the term “metabolomics”. Metabonomics really grew out of work using NMR spectroscopy of biofluids going back to the mid-1980s, and which was subsequently combined with the use of pattern recognition and multivariate statistics investigation of the complex data sets. The term was not coined until much later and was formally defined in 1999 by Jeremy Nicholson and colleagues as “the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification”. A little later, in 2001, the term “metabolomics” was introduced by Oliver Fiehn and defined somewhat differently as “a comprehensive and quantitative analysis of all metabolites” in a system. Although there remain some differences in concept, there is now a great deal of overlap in the philosophies and methodologies, and the two terms are often used interchangeably by scientists and organisations. In this volume, we have allowed authors to use their term of preference.

Metabolic analyses provide a data-dense approach to biochemistry by monitoring simultaneously the changes in concentrations (and in some cases, molecular dynamics and compartmentation) of a wide range of molecules. To aid the observation and understanding of those differences that are really significant between classes, it has been necessary to use multivariate statistics extensively and interpretation of such calculations leads directly to the real biochemical differences between sample classes and hence identification of biomarkers of the process under study.

The main aim of this book is to provide a state-of-the-art picture of where metabonomics and metabolomics stand today, to give authoritative education and guidance on the analytical and statistical techniques used, and to identify and review the main current areas of application. The main analytical techniques of NMR spectroscopy and chromatography linked to mass spectrometry are explained and reviewed in detail, as are the various chemometrics and statistics approaches that are widely used. In addition we have taken the opportunity to provide information on the recent attempts at setting standards in designing and reporting metabolic experimental studies. We have invited articles to cover a wide range of biological applications. These range from plants and model organisms through to pre-clinical and clinical pharmaceutical studies and human disease and epidemiological investigations.

We greatly appreciate the efforts of all of the authors of the chapters who have consistently provided excellent articles. These have a high educational content as well as providing reviews of very current and cutting-edge literature. We hope and believe that this provides a balanced view of the subject with both the advantages and the shortcomings of the various methodologies being explored. Each chapter also includes a relevant and comprehensive set of references for further reading.

The field of endeavour is expanding rapidly, but we believe that now is a good time to review the achievements in the area of metabonomics and metabolomics and for the scientific community to set standards for the future. Our intention is that this book will serve as the authoritative reference work for scientists entering the field and for those already conducting metabolic studies. We believe that it also has a substantial educational role and will be of use to postgraduate students. It should be of interest to analytical scientists and to those working in application areas as diverse as plant science, environmental science and human clinical medicine. In summary we believe that this volume should prove educational, informative, critical and thought-provoking.

John C. Lindon
Jeremy K. Nicholson
Elaine Holmes
London, UK
May 2006

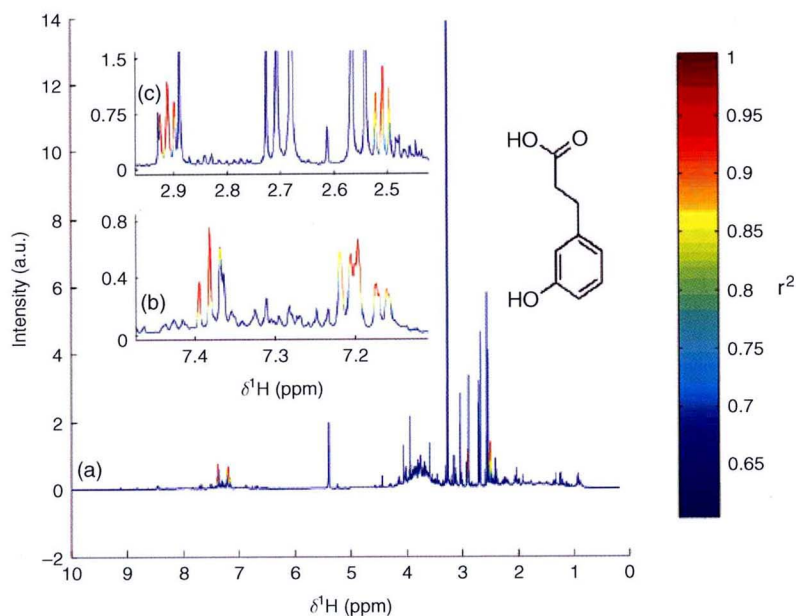


Figure 1.5. One-dimensional STOCSY analysis to identify peaks correlated to that at the chemical shift, δ 2.51. The degree of correlation across the spectrum has been colour-coded and projected on the spectrum. (a) Full spectrum; (b) partial spectrum between δ 7.1 and 7.5; (c) partial spectrum between δ 2.4 and 3.0. The STOCSY procedure enabled the assignment of this metabolite as 3-hydroxyphenylpropionic acid. Reproduced with permission from Cloarec *et al.* [63].