



Precision
Medicine

精准医学出版工程

精准医学基础系列

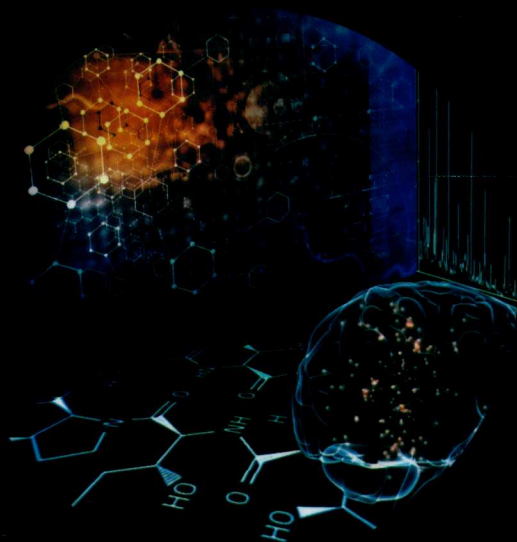
总主编 詹启敏

“十三五”国家重点图书出版规划项目

代谢组学与精准医学

Metabolomics and Precision Medicine

贾伟等
著



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS



Precision
Medicine

精准医学出版工程

精准医学基础系列

总主编 詹启敏

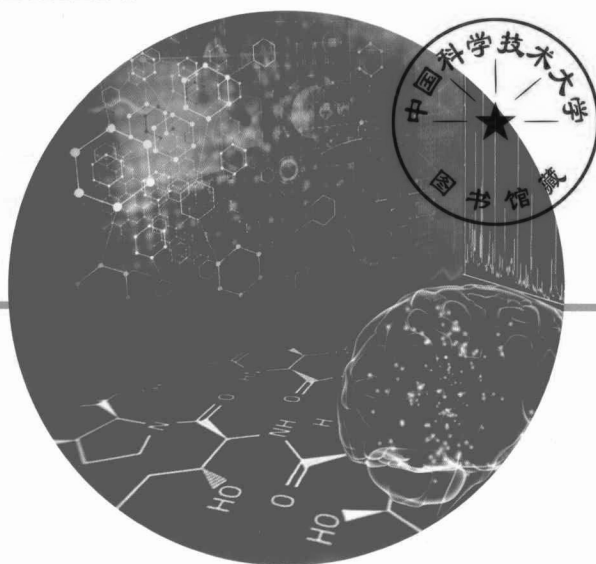
“十三五”国家重点图书出版规划项目

代谢组学与精准医学

Metabolomics and Precision Medicine

贾伟 等

著



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

内容提要

本书系统介绍了代谢组学在最近5年内的新进展,汇聚了国际代谢组学最新的临床研究成果,面向未来,引领精准医学研究的发展方向。本书重点聚焦在各类疾病的临床代谢组学研究,关注恶性肿瘤、代谢性疾病、药物不良反应、肝病、心脑血管疾病等传统方向的同时,还增加了生殖、儿童及老年性疾病以及脑疾病等新兴和热点领域的代谢组学应用和发展。同时,本书也针对本领域近年来出现的问题,详细介绍了所面临的挑战:现有方法学问题,数据标准化、行业标准化问题,各实验室和研究中数据的可重复性问题,以及如何整合等。

本书可以作为具有一定专业背景、从事代谢组学研究的临床医生和科研人员的参考书,也可作为相关领域研究生和高年级本科生的教材。具备一定临床医学、基础医学、生物化学和分析化学背景的读者可根据自己的兴趣选择性阅读部分章节。

图书在版编目(CIP)数据

代谢组学与精准医学/贾伟等著. —上海:上海交通大学出版社,2017

精准医学出版工程

ISBN 978-7-313-18400-9

I. ①代… II. ①贾… III. ①代谢病—数据处理 IV. ①R589-39

中国版本图书馆CIP数据核字(2017)第278594号

代谢组学与精准医学

著 者:贾伟等

出版发行:上海交通大学出版社

邮政编码:200030

出版人:谈毅

印 制:苏州市越洋印刷有限公司

开 本:787mm×1092mm 1/16

字 数:412千字

版 次:2017年12月第1版

书 号:ISBN 978-7-313-18400-9/R

定 价:248.00元

地 址:上海市番禺路951号

电 话:021-64071208

经 销:全国新华书店

印 张:24.5

印 次:2017年12月第1次印刷

版权所有 侵权必究

告读者:如发现本书有印装质量问题请与印刷厂质量科联系

联系电话:0512-68180638

编委会

总主编

詹启敏(北京大学副校长、医学部主任,中国工程院院士)

编委

(按姓氏拼音排序)

陈超(西北大学副校长、国家微检测系统工程技术研究中心主任,教授)

方向东(中国科学院基因组科学与信息重点实验室副主任、中国科学院北京基因组研究所“百人计划”研究员,中国科学院大学教授)

郜恒骏(生物芯片上海国家工程研究中心主任,同济大学医学院教授、消化疾病研究所所长)

贾伟(美国夏威夷大学癌症研究中心副主任,教授)

钱小红(军事科学院军事医学研究院生命组学研究所研究员)

石乐明(复旦大学生命科学学院、复旦大学附属肿瘤医院教授)

王晓民(首都医科大学副校长,北京脑重大疾病研究院院长,教授)

于军(中国科学院基因组科学与信息重点实验室、中国科学院北京基因组研究所研究员,中国科学院大学教授)

赵立平(上海交大学生命科学技术学院特聘教授,美国罗格斯大学环境与生物科学学院冠名讲席教授)

朱景德(安徽省肿瘤医院肿瘤表观遗传学实验室教授)

学术秘书

张华(中国医学科学院、北京协和医学院科技管理处副处长)

《代谢组学与精准医学》

编委会

主 编

贾 伟(美国夏威夷大学癌症研究中心副主任,教授)

副主编

张永煜(上海中医药大学交叉科学研究院教授)

王晓艳(上海交通大学系统生物医学研究院副研究员)

编委会

(按姓氏拼音排序)

陈天璐(上海交通大学附属第六人民医院转化医学中心副研究员)

傅君芬(浙江大学医学院附属儿童医院副院长,教授)

黄荷凤(国际和平妇幼保健院院长,中国科学院院士)

李后开(上海中医药大学交叉科学研究院教授)

林景超[麦特绘谱生物科技(上海)有限公司研究员]

倪 艳(美国夏威夷大学癌症研究中心助理教授)

苏明明[麦特绘谱生物科技(上海)有限公司研究员]

王晓柠(上海中医药大学交叉科学研究院研究员)

吴 蔚(浙江大学医学院附属儿童医院主任医师)

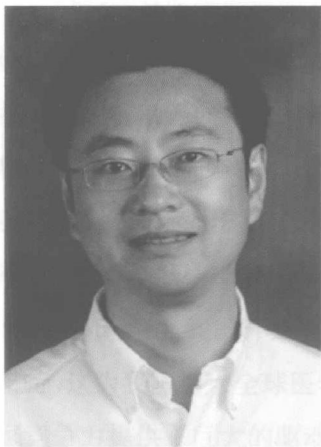
谢国祥(美国夏威夷大学癌症研究中心助理教授)

徐晨明(国际和平妇幼保健院生殖遗传科主任,副研究员)

于浩泳(上海交通大学附属第六人民医院副主任医师)

郑晓皎(上海交通大学附属第六人民医院转化医学中心副研究员)

主编简介



贾伟，美国密苏里大学放射药理学博士，现任美国夏威夷大学终身教授、夏威夷大学癌症研究中心副主任，同时兼任上海交通大学附属第六人民医院教授、转化医学中心主任。主要研究方向为肠道菌群和肝脏的代谢(相互)作用及其对肥胖、糖尿病与消化道肿瘤发病机制的影响。自2004年建立了基于质谱的医学代谢组学技术平台以来，其研究一直处于国际领先地位。现主持美国国立卫生研究院等机构资助的多项基金研究项目，曾担任国家科技部973计划项目“代谢性疾病的蛋白质功能和代谢组学研究”的首席科学家。目前担任美国国立卫生研究院基金项目等的评委，中国香港科技署基金、中国国家自然科学基金评委。同时担任9家国际学术期刊编委。迄今已在国际学术期刊发表科技论文200余篇，出版专著5部。

“精准”是医学发展的客观追求和最终目标，也是公众对健康的必然需求。“精准医学”是生物技术、信息技术和多种前沿技术在医学临床实践的交汇融合应用，是医学科技发展的前沿方向，实施精准医学已经成为推动全民健康的国家发展战略。因此，发展精准医学，系统加强精准医学研究布局，对于我国重大疾病防控和促进全民健康，对于我国占据未来医学制高点及相关产业发展主导权，对于推动我国生命健康产业发展具有重要意义。

2015年初，我国开始制定“精准医学”发展战略规划，并安排中央财政经费给予专项支持，这为我国加入全球医学发展浪潮、增强我国在医学前沿领域的研究实力、提升国家竞争力提供了巨大的驱动力。国家科技部在国家“十三五”规划期间启动了“精准医学研究”重点研发专项，以我国常见高发、危害重大的疾病及若干流行率相对较高的罕见病为切入点，将建立多层次精准医学知识库体系和生物医学大数据共享平台，形成重大疾病的风险评估、预测预警、早期筛查、分型分类、个体化治疗、疗效和安全性预测及监控等精准预防诊治方案和临床决策系统，建设中国人群典型疾病精准医学临床方案的示范、应用和推广体系等。目前，精准医学已呈现快速和健康发展态势，极大地推动了我国卫生健康事业的发展。

精准医学几乎覆盖了所有医学门类，是一个复杂和综合的科技创新系统。为了迎接新形势下医学理论、技术和临床等方面的需求和挑战，迫切需要及时总结精准医学前沿研究成果，编著一套以“精准医学”为主题的丛书，从而助力我国精准医学的进程，带动医学科学整体发展，并能加快相关学科紧缺人才的培养和健康大产业的发展。

2015年6月，上海交通大学出版社以此为契机，启动了“精准医学出版工程”系列图

书项目。这套丛书紧扣国家健康事业发展战略,配合精准医学快速发展的态势,拟出版一系列精准医学前沿领域的学术专著,这是一项非常适合国家精准医学发展时宜的事业。我本人作为精准医学国家规划制定的参与者,见证了我国精准医学的规划和发展,欣然接受上海交通大学出版社的邀请担任该丛书的总主编,希望为我国的精准医学发展及医学发展出一份力。出版社同时也邀请了刘彤华院士、贺福初院士、刘昌孝院士、周宏灏院士、赵国屏院士、王红阳院士、曹雪涛院士、陈志南院士、陈润生院士、陈香美院士、金力院士、周琪院士、徐国良院士、董家鸿院士、卞修武院士、陆林院士、乔杰院士、黄荷凤院士等医学领域专家撰写专著、承担审校等工作,邀请的编委和撰写专家均为活跃在精准医学研究最前沿的、在各自领域有突出贡献的科学家、临床专家、生物信息学家,以确保这套“精准医学出版工程”丛书具有高品质和重大的社会价值,为我国的精准医学发展提供参考和智力支持。

编著这套丛书,一是总结整理国内外精准医学的重要成果及宝贵经验;二是更新医学知识体系,为精准医学科研与临床人员培养提供一套系统、全面的参考书,满足人才培养对教材的迫切需求;三是为精准医学实施提供有力的理论和技术支撑;四是将许多专家、教授、学者广博的学识见解和丰富的实践经验总结传承下来,旨在从系统性、完整性和实用性角度出发,把丰富的实践经验和实验室研究进一步理论化、科学化,形成具有我国特色的精准医学理论与实践相结合的知识体系。

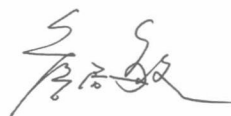
“精准医学出版工程”丛书是国内外第一套系统总结精准医学前沿性研究成果的系列专著,内容包括“精准医学基础”“精准预防”“精准诊断”“精准治疗”“精准医学药物研发”以及“精准医学的疾病诊疗共识、标准与指南”等多个系列,旨在服务于全生命周期、全人群、健康全过程的国家大健康战略。

预计这套丛书的总规模会达到 60 种以上。随着学科的发展,数量还会有所增加。这套丛书首先包括“精准医学基础系列”的 11 种图书,其中 1 种为总论。从精准医学覆盖的医学全过程链条考虑,这套丛书还将包括和预防医学、临床诊断(如分子诊断、分子影像、分子病理等)及治疗相关(如细胞治疗、生物治疗、靶向治疗、机器人、手术导航、内镜等)的内容,以及一些通过精准医学现代手段对传统治疗优化后的精准治疗。此外,这套丛书还包括药物研发,临床诊疗路径、标准、规范、指南等内容。“精准医学出版工程”将紧密结合国家“十三五”重大战略规划,聚焦“精准医学”目标,贯穿“十三五”始终,力求打造一个总体量超过 60 本的学术著作群,从而形成一个医学学术出版的高峰。

本套丛书得到国家出版基金资助,并入选了“十三五”国家重点图书出版规划项目,体现了国家对“精准医学”项目以及“精准医学出版工程”这套丛书的高度重视。这套丛书承担着记载与弘扬科技成就、积累和传播科技知识的使命,凝结了国内外精准医学领域专业人士的智慧和成果,具有较强的系统性、完整性、实用性和前瞻性,既可作为实际工作的指导用书,也可作为相关专业人员的学习参考用书。期望这套丛书能够有益于精准医学领域人才的培养,有益于精准医学的发展,有益于医学的发展。

此次集束出版的“精准医学基础系列”系统总结了我国精准医学基础研究各领域取得的前沿成果和突破,内容涵盖精准医学总论、生物样本库、基因组学、转录组学、蛋白质组学、表观遗传学、微生物组学、代谢组学、生物大数据、新技术等新兴领域和新兴学科,旨在为我国精准医学的发展和实施提供理论和科学依据,为培养和建设我国高水平的具有精准医学专业知识和先进理念的基础和临床人才队伍提供理论支撑。

希望这套丛书能在国家医学发展史上留下浓重的一笔!



北京大学副校长
北京大学医学部主任
中国工程院院士
2017年11月16日

序

精准医学(precision medicine)是一种预防和治疗疾病的新方法,它考虑个体基因差异、环境因素及个人生活方式差异等影响,目前主要在某些癌症筛查方面取得明显进展,但并未应用到其他大多数疾病的诊治中。

2003年,“人类基因组计划”宣布完成。之后,相继应运而生的“组学”如基因组学、转录组学、蛋白质组学和代谢组学等,为从基因组生物学进入精准医学时代奠定了基础,“个性化医疗”的医学实践对分子水平信息的使用将使医学更精准成为现实。

2015年1月,美国总统奥巴马提出的精准医学计划,将彻底改变医学现状并加快精准医学理念融入日常临床实践的步伐。在生物技术和信息技术与医学临床实践的交汇融合应用中,我国自2015年以来将精准医学作为医学科技发展的前沿方向,目标明确地强调:系统加强精准医学研究布局,对于加快重大疾病防控技术突破,占据未来医学及相关产业发展主导权,打造我国生命健康产业发展的新驱动力至关重要。在计划实施上,我认为其发展的本质是通过“组学”技术和医学前沿技术,对于大样本人群和特定疾病类型进行生物标志物的分析与鉴定、验证与应用,从而精确寻找到疾病的原因和治疗的靶点,并对一种疾病的不同状态和过程进行精确的亚分类,最终实现对疾病和特定患者进行个性化精准治疗的目的,提高疾病诊治与预防的效益。精准医学是因人因病而异的、更加精确的个体化医疗,其进步之处是将人们对疾病机制的认识与生物大数据和信息科学相结合,精确进行疾病分类及诊断(精准诊断),精准应用现有药物和创新药物(精准药物),为患者提供更具针对性和有效性的预防和治疗措施,最终让精准医学(精准医疗)更好地为患者服务。

如何精准认识“组学”并利用“组学”推进精准医疗的发展,是“组学”研究者的使命。

组学的特点就是数据量大,而研究其差异是难点,如能高效发现差异基因、差异蛋白、差异代谢物,就能对功能(原因、过程、结果)解释、新功能发现(治疗靶点)、新药物疗效(生物标志物)等提供可靠的指标。基因组学虽然在基因活性和疾病的相关性方面为人类提供了有力依据,但实际上大部分疾病并不是基因改变所致。并且,基因的表达方式错综复杂,同样一个基因在不同条件、不同时期可能会起完全不同的作用,基因组学对其后果是无法回答的,只有发现基因差异(突变)才能分析突变所致的可能性。蛋白质组学是通过寻找各种因素引起的蛋白质表达差异,解释细胞生理和病理机制,即主要通过比较分析不同状态下或近似物种间蛋白质的表达图谱,实现对体系内代谢调控的动态监测,从而揭示机体对内外环境变化产生反应的本质规律。代谢组学是继基因组学和蛋白质组学之后新近发展起来的一门学科,是系统生物学的重要组成部分。其后,代谢组学又迅速发展并渗透到多个领域,如疾病诊断、医药研制开发、营养食品科学、毒理学等与人类健康密切相关的领域。基因组学和蛋白质组学分别从基因和蛋白质层面探寻生命的奥秘,而细胞内许多变化是发生在代谢物层面的,如细胞信号释放、能量传递、细胞间通信等都是受代谢物调控的。代谢组学正是研究代谢组——在某一时刻细胞内所有代谢物的集合——的一门学科。基因与蛋白质的表达紧密相关,而代谢物则更多地反映了细胞所处的环境,这又与细胞的营养状态、药物和环境污染物的作用及其他外界因素的影响密切相关。因此,正如美国加州大学戴维斯分校的学者 Bill Lasley 所说:“基因组学和蛋白质组学告诉你什么可能会发生,而代谢组学则告诉你什么确实发生了。”

贾伟教授和他带领的团队坚持从事代谢组学研究 10 余年,从美国到天津,之后到上海,再回到美国,一直坚守这个阵地辛勤耕耘。即将出版的新作《代谢组学与精准医学》就是贾伟教授和他带领的团队辛勤耕耘的结晶。

本书系统介绍了代谢组学在最近 5 年内的新进展,汇聚了国际代谢组学最新的临床研究成果。新作的特点之一是,重点聚焦在各类疾病的临床代谢组学研究,既关注恶性肿瘤、代谢性疾病、药物不良反应、肝病、心脑血管疾病等传统方向,又增加了生殖、儿童及老年性疾病、脑疾病等新兴和热点领域的代谢组学研究成果。因此,它能面向未来,引领精准医学研究的发展方向。特点之二是,本书也针对本领域近年来出现的问题,详细介绍了代谢组学研究所面临的挑战:现有方法学的问题,数据标准化、行业标准化的问题,各实验室和研究中数据可重复性的问题,以及如何通过整合使其更具实用性

的问题。

基于本书的上述特点,我真诚地将本书介绍给读者。对于具备一定临床医学、基础医学、生物化学和分析化学背景的读者,如有选择地阅读部分章节,更能达到事半功倍的效果。当然,对于具有一定专业背景并从事代谢组学研究的临床医生和科研人员来说,本书是一本难觅的参考书。此外,本书也可作为相关领域研究生和高年级本科生的教材。读者读后一定会真正认识“基因组学和蛋白质组学告诉你什么可能会发生,而代谢组学则告诉你什么确实发生了。”的含义。

研究对象的代谢组学研究结果告诉研究者确实发生了什么变化,更能体现代谢组学在精准医学中的作用,这是基因组学和蛋白质组学所不能替代的。本人自2003年接触代谢组学以来,一直倡导中国学者开展相关研究,自己也在此领域进行过一点试验性研究。本人认为代谢组学还处于发展阶段,人们能够了解的代谢物的数量还很有限(不到10%)。也就是说,现实就是如此。相信随着研究方法的不断完善和优化,同行们会通过代谢组学研究使代谢组学成为人类更高效、精准诊断疾病和治疗疾病的有力工具。

回顾代谢组学的发展,欣喜于《代谢组学与精准医学》的出版,特以继往开来的心境写成此序,以资恭贺。

刘昌孝

天津药物研究院新药评价研究中心

中国工程院院士

2017年11月于天津

前言

最近,代谢组学的创始人、英国帝国理工大学的 Jeremy Nicholson 教授一直在宣传“表型组学”,即 phenomics。表型组学的概念目前还不是那么清晰,可以笼统地理解为研究某一生物或细胞除了基因组以外的所有组学的集合,而其中最核心的部分,就是代谢组!2016年8月复旦大学的唐惠儒教授告诉我,复旦的金力教授正在牵头开展基于表型组的大型队列研究。在队列研究的范畴内,他们把表型组定义为个体从胚胎发育到出生、成长、衰老及死亡过程中的形态特征、功能行为、分子组成规律,分成三个层面——生物特征、物理特征、化学特征进行系统的测量。

我个人很推崇这个表型组研究的策略,因为基因组学无法独力承担精准医学的任务。如果说疾病是遗传因素和包括生活方式在内的所有环境因素共同作用的一种结果的话,那么表型组反映的信息更接近疾病本身。

细胞内的生命活动是由众多基因、蛋白质以及小分子代谢物共同承担的,而上游的核酸、蛋白质等大分子的功能性变化最终会体现于代谢层面,如神经递质的变化、激素调控、受体作用效应、细胞信号释放、能量传递和细胞间通信等,所以代谢组处于基因调控网络和蛋白质作用网络的下游,所提供的是生物学的终端信息。如同在长江的上游建大坝或让江水改道,这些项目对生态的影响会在下游的河道和地域体现出来一样,我们经常说:“基因组学和蛋白质组学告诉你可能发生什么,而代谢组学则告诉你已经发生了什么。”

我们称细胞内的代谢物特征性变化为代谢指纹(metabolic fingerprints),分泌到细胞外的代谢物为代谢足迹(metabolic footprints)。与基因组学、转录组学和蛋白质组学相比较,代谢组学还具有以下特点。首先,基因和蛋白质表达在功能水平上的微小变化

会在代谢物上得到放大,从而使检测更容易;其次,许多基因和蛋白质的非功能性变化不会在代谢物上反映出来,从而在上游信息向下游传递过程中达到“噪声过滤”的效果;再次,代谢物的种类数要远小于基因和蛋白质的数目,物质的分子结构也要简单得多,因而代谢组学研究所采用的代谢物信息库,远没有全基因组测序及大量表达序列标签的数据库那么复杂。另外,常见代谢物在各个生物体系[如植物(初级代谢)、微生物、动物]中都相似,所以代谢组学研究中采用的平台技术可以在不同的生物体系中得到应用。

唐代诗人孟浩然有两句诗:“人事有代谢,往来成古今。”从万物皆有兴衰代谢的角度来看,我们的生物世界其实是由代谢组组成的,是这些不同的代谢组让我们的生物界呈现出五彩缤纷、气象万千的表型。地球上的各种植物含有几十万种代谢物(25万~50万种植物化学分子),微生物界有几万种代谢物,而哺乳动物体内常见的代谢物(分子量小于1500)有5000~7000种。这三类代谢组相互渗透,植物和微生物的代谢物通过食物、营养补充、药物等形式进入我们人体的代谢网络,也使我们每一个人的代谢表型呈现各种不同的特征。

我曾经在以前的一篇博文中把人体的代谢网络比喻成我们所在城市的交通网络。从市中心(如上海市的人民广场)到城市外围的任意一点(如上海浦东国际机场)在理论上有无数条路径可走,但大家都知道最可行的路径也就是少数几条。而某一刻我们究竟通过哪条路径去机场,主要看这一时刻的交通工具、交通状况,以及我们的时间和资金情况。生命活动其实也是一样的。人类有3万多个基因,尽管功能基因所占比例不大,但它们在排列组合之下,就会出现无穷多种可能。奇妙的是,在指挥系统近乎无穷多种可能的指令下,仅仅产生出几万种蛋白质。而下游的代谢物和代谢途径更少,尤其是主要的代谢途径(交通主干线)更是屈指可数,可以在一张白纸上画清楚。这说明什么呢?这说明再复杂的生物系统在它的功能层面都有着简单的、共性的一面。有人对收集到的各种癌细胞进行检测,发现了大约500万种基因突变方式,但这些变化再复杂还是有章可循的。它们无非是要在功能层面(如代谢层面)实施调整和转换,达到一个或几个简单的目的——要么获取更多的能量,要么获得更多的物质,或者设法排除更多的废弃物,或者增强自身的抗氧化、抗应激(抗药物)能力。总之,癌基因调控的目的明确——要生存、增殖、从周边掠夺资源并向周边扩散。如果我们能够这样看问题,就可以在寻找共性的变化中把复杂问题简单化。因此,代谢组学将是疾病分子表型和功能研究中

的一门核心技术!

再举个例子说明代谢组学的重要性。现在,肠道菌群研究已经成为科研界最热门的领域之一。但是,这个领域目前采用的只有一个技术——测序,不是 16S rRNA 测序就是宏基因组测序。测序告诉我们的是什么呢?是肠道细菌的种类信息,从门到属和种(有时甚至能到株)的分类和丰度值。如同你要研究一个城市的安全问题,这个检测技术可以帮你得到一本覆盖全市大多数居民的花名册,仅此而已。两个肠道菌群组成相差很大的健康人站在一起,我们无法判断他们的肠道菌群结构差异意味着什么,如果两个人用同一种饮食,这些菌群差异会给两个人的代谢和生理带来什么功能性变化我们还难以预测。当然,菌群研究者们说,他们可以通过检索数据库获得功能信息,但怎么利用这些功能信息建模预测呢?将每一种代谢功能下的各种细菌进行相加或是加权后相加?那么互相抵消、互相干扰的怎么算呢?最简单明了的代谢功能表征方法就是测代谢组,由此获得的数据是各种细菌集成的功能及各种细菌与宿主共同作用的最终结果。

但是,代谢组学目前尚无法全面进入精准医学和相关健康领域提供产业化服务。其主要瓶颈有两个:一是标准化的问题,二是通量的问题。代谢组学向前发展的一个必经之路是定量化和标准化。基因测序技术目前已经成为转化研究和技术产业化的首选工具,其中一个重要原因是这种高通量检测技术的标准化已日渐成熟并正在行业内逐步普及。目前,国内测序行业的多家企业在基因组数据分析处理方面(包括测序采样与分析、碱基读出、载体标识与去除、拼接与组装、间歇填补、重复序列标识等)逐步建立了统一的标准和流程。我们可以把华大基因比喻成统一六国的秦国,它积极参与国际领域内大数据管理、整合和共享标准的建立,利用自身庞大的测序平台体量和实力,在技术标准方面成为行业内的执牛耳者。而代谢组学还没有发展成熟,还处于“春秋战国诸侯争霸时期”。

目前除了核磁共振仪外,代谢组学的主要分析仪器为质谱仪。而生产飞行时间质谱(TOFMS)、三重四极杆串联质谱(TQMS)、四极杆-飞行时间串联质谱(Q-TOFMS)、离子回旋共振质谱(ICRMS)、轨道离子阱(Orbitrap)等高分辨质谱仪的厂家有 10 家以上。这些厂家都有自己独特的仪器配置、数据处理软件及数据库。不同厂家用的工作软件和数据库之间都无法对话(cross-talk),因此一旦购买了某一个厂家的设备进行代谢组学研究,研究者往往只有照搬该厂家提供的全套分析工具,因而整个行业

缺乏包括数据处理标准、数据分析途径、生物描述规范及报告标准在内的统一的代谢组学标准流程或标准协议。对于代谢物的鉴定,各个实验室的做法也是参差不齐,有的完全依赖国际数据资源库,有的用厂家自带数据库,有的用自己的标准品进行鉴定,以致数据的质量良莠不齐。

代谢组学如果要全面进入临床医学和健康产业的服务领域,就需要花大力气解决技术平台的行业标准化问题。从代谢组学设备生产厂家到各个实验室之间都需要逐步改变工作模式,从各自为战到相互合作、统一标准,共同建立行业内的技术规范,不同平台产生的数据可以相互验证(cross-validation),最终建立起一个行业内可以共享的代谢组学数据库。

也只有行业普遍接受的技术标准下,才可能扩大检测规模。如果没有一定的检测通量,如一次检测数万或数十万样本的能力,代谢组学技术就很难在大型研究项目和精准医学领域扮演一个有意义的角色。前面说到复旦大学开展的基于大型队列的表型组学研究,目前已经纳入计划的队列达到 20 万人,以每个人在 6 个时间点采集样本计,总样本数就达到 120 万份,随着计划的推进,样本数将持续上升。可以想象,只有采用统一的技术标准和具备足够检测通量的代谢组学实验室才可能承担这类项目的研究工作。

我记得 20 世纪 80 年代读书的时候看过一部纪录片《话说长江》,当时这部电视纪录片播出后举国轰动,中国观众在信息闭塞了几十年后,通过对一条流淌了数千万年全长六千多公里河流的介绍,第一次直观地、全景地在电视中看到了自己国家广袤的大地、多彩的文化以及长江流域美丽的自然风光。同样地,尽管今天科学高度发达,人类对于自身几乎所有重大疾病发病机制的认知水平还处于黑暗时代(Dark Ages)阶段。随着基因组学的日趋成熟和表型研究工具如代谢组学的广泛应用,我们将会把基因和表型信息连接起来,有可能逐步打开一些疾病的“黑箱”,像了解一条古老的河流一样逐步认识我们的生命,一步一步地逼近疾病和生命的本质!

本书由美国夏威夷大学癌症研究中心副主任贾伟教授主持著述,上海交通大学系统生物医学研究院王晓艳副研究员任副主编。编写工作得到诸多科研院所、高等院校和临床医院的大力支持和帮助。衷心感谢刘昌孝院士为本书作序!编写组由美国夏威夷大学、上海交通大学、上海交通大学附属第六人民医院、国际和平妇幼保健院、上海中医药大学、浙江大学医学院附属儿童医院、麦特绘谱生物科技(上海)有限公司等单位的