



Precision
Medicine

精准医学出版工程

精准医学基础系列

总主编 詹启敏

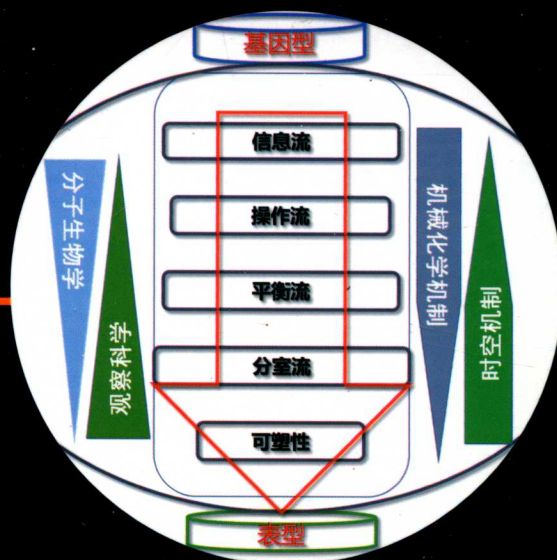
“十三五”国家重点图书出版规划项目

基因组学与精准医学

Genomics and Precision Medicine

于军 等

编著



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



“十三五”国家重点图书出版规划项目

Precision
Medicine

精准医学出版工程

精准医学基础系列

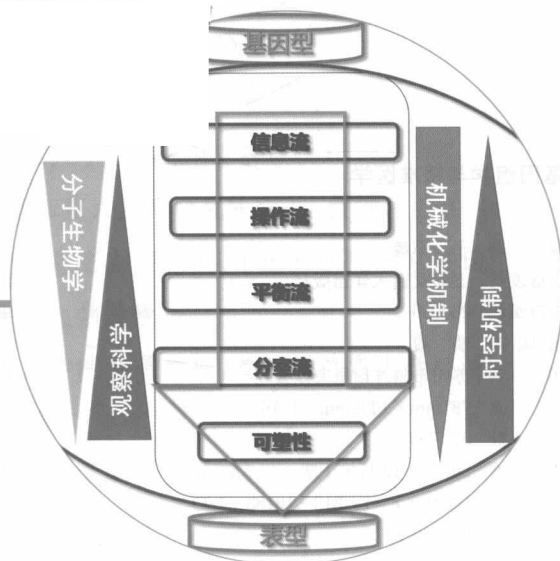
总主编 詹启敏

基因组学与精准医学

Genomics and
Precision Medicine

于军 等

编著



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

内容提要

基因组是指细胞或者其生命体所携带的全部 DNA 序列总和。因为 DNA 分子通常与蛋白质等形成复合体,所以也指细胞染色质或染色体的总和。基因组学或基因组生物学是研究基因组结构、动态和功能的科学。基因组学研究为疾病分类、诊断、治疗、预后和药物研发等提供理论和技术支持。本书系统介绍了基因组学及其核心技术——DNA 测序和生物信息学的发展及其在精准医学方面的应用实例与最新研究成果,为相关科研人员、医务工作者和专业学生提供最新知识信息的同时,也启迪人们对新兴的精准医学研究和发展的思考。

图书在版编目(CIP)数据

基因组学与精准医学/于军等编著. —上海:上海交通大学出版社,2017
精准医学出版工程
ISBN 978-7-313-18183-1

I. ①基… II. ①于… III. ①人类基因—基因组 IV. ①Q987

中国版本图书馆 CIP 数据核字(2017)第 236731 号

基因组学与精准医学

编 著: 于 军等	地 址: 上海市番禺路 951 号
出版发行: 上海交通大学出版社	电 话: 021-64071208
邮政编码: 200030	
出 版 人: 谈 毅	
印 制: 苏州市越洋印刷有限公司	经 销: 全国新华书店
开 本: 787mm×1092mm 1/16	印 张: 22.25
字 数: 373 千字	
版 次: 2017 年 12 月第 1 版	印 次: 2017 年 12 月第 1 次印刷
书 号: ISBN 978-7-313-18183-1/Q	
定 价: 228.00 元	

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)

《基因组学与精准医学》

编委会

主 编

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)

副主编

胡松年(中国科学院基因组科学与信息重点实验室主任、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)

肖景发(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)

吴双秀(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副研究员)

编 委

(按姓氏拼音排序)

顾明亮(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高级工程师)

华 沙(上海交通大学医学院附属瑞金医院卢湾分院主治医师)

黄绮梦(中国科学院重庆绿色智能技术研究院助理研究员)

康 禹(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副研究员)

孟庆姝(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助理研究员)

舒 畅(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助理研究员)

王德强(中国科学院重庆绿色智能技术研究院精准医疗单分子诊断技术研究中心主任,研究员)

吴 浩(瑞典哥德堡大学萨尔格林斯卡医院助理研究员)

张永彪(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助理研究员)

张哲文(中国科学院北京基因组研究所生命与健康大数据中心助理研究员)

孟庆姝(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助理研究员)

舒 畅(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助理研究员)

王德强(中国科学院重庆绿色智能技术研究院精准医疗单分子诊断技术研究中心主任,研究员)

吴 浩(瑞典哥德堡大学萨尔格林斯卡医院助理研究员)

张永彪(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助理研究员)

张哲文(中国科学院北京基因组研究所生命与健康大数据中心助理研究员)

主编简介



于军，1956 年出生。美国纽约大学医学院生物医学专业博士，现为中国科学院特聘研究员，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博士生导师，中国科学院大学教授。主要研究方向包括基因组学、生物信息学和精准医学仪器设备研发。1983 年毕业于吉林大学化学系生物化学专业；同年，在中国科学院生物物理研究所攻读硕士学位并考取 CUSBEA 奖学金；次年赴美国纽约大学(New York University)医学院攻读生物医学博士学位，师从生物化学与细胞生物学家孙同天教授，从事表皮细胞分化分子生物学研究；1990

年毕业，获得 AFUD 博士研究者奖并留校任研究助理教授。1993—1998 年加入华盛顿大学(西雅图)基因组科学系 Maynard V. Olson 教授主持的人类基因组研究中心，参与国际人类基因组计划，创建了“多酶全消化基因组物理图谱制作法”，从事规模化人类基因组物理图谱制作、基因测序和生物信息数据分析。1998 年回国后，先后主导创立了中国科学院遗传所“人类基因组中心”、“华大基因(BGI)”、中国科学院北京基因组研究所(BIG)等。策划并参与主持了国际人类基因组计划(中国部分)及超级杂交水稻、家蚕、枣椰、鲤鱼等动植物基因组研究计划，主持了国家重大科学研究计划“小型猪和小鼠等医学实验哺乳动物模型建立与基础数据集成”和“以细胞为单元的人类基因转录组与蛋白质组的关联性研究”。曾获香港求是科技基金会“求是杰出科技成就奖”(2002 年)、中

国科学院杰出科技成就奖(2003 年)、第三世界科学院农业科学贡献奖(TWAS, 2012 年)等。现任中国遗传学会常务理事、基因组学分会主任委员, *Genomics* 和 *Proteomics and Bioinformatics* 杂志主编。迄今已在包括 *Nature*、*Science* 等同行评议科学期刊发表科学论文 300 余篇。



“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防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、刘昌效院士、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。

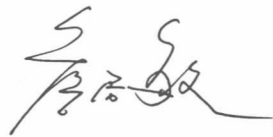
“精准医学出版工程”是国内外第一套系统总结精准医学前沿性研究成果的系列专著,内容包括“精准医学基础”“精准预防”“精准诊断”“精准治疗”“精准医学药物研发”以及“精准医学的疾病诊疗共识、标准与指南”等多个系列,旨在服务于全生命周期、全人群、健康全过程的国家大健康战略。

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

前言

本书是“精准医学出版工程·精准医学基础系列”图书的一个分册。应邀撰写本分册各章的作者均为基因组学、生物信息学和临床研究中经验丰富的资深科研人员,大家精心写作,以展示国际、国内该领域最新和最激动人心的研究成果、研究方法和应用案例为基本目标。虽然有些概念、技术和方法还不能完全解决精准医学将要面对的种种问题,尤其是诸多复杂疾病的精准诊疗问题,但是作者们希望这些知识和线索能引导相关科研和医务人员开拓思路、大胆尝试。

走进精准医学是生物医学发展的必然,是基因组生物学与生命科学要素系统细分、精炼,以及攀登知识新高峰历程中的又一个里程碑。从基因组结构到基因组生物学,最后到疾病生物学和健康保障科学与技术——这就是人类基因组计划和精准医学计划延续走来的路线图。在这个路线图下,精准医学纵向整合各类技术,横向整合各类“组学”数据,将“大数据”推向新的疾病分类和基础科学研究,形成整合数据、信息和知识的回路,如此循环。

精准医学始于科学项目,通过学科建设和成果应用,进一步拓展生物医学研究的视野,增强科学发现对社会的直接影响力。首先是研究对象的前移,从模式生物系统到真正的患者群体,从原理和机制的研究到临床实际应用;其次是研究范围的拓宽,从罕见病到常见病,从诊断到治疗和药物研发;最后是数据收集在时间轴上的拓展,从医院收集到患者生活圈收集,从染病时段收集到正常生活收集。这样,精准医学就必然走向精准医疗和精准健康。最终,生物医学研究和其他生命科学相关研究将汇聚到健康保障体系里面。

精准医学成为基因组学的新时代标签。30 年来的基因组学研究,为这一时代打下

了坚实的基础,使人们可以直接面对恶性肿瘤、神经系统疾病、心血管疾病、代谢性疾病等复杂疾病——也就是多基因疾病,提出科学问题,为这些疾病的诊断和治疗提供科学依据。从2011年“精准医学”概念的提出到各国启动类似大研究项目,充分肯定了这个思路的正确性。我国在2015年也启动了精准医学的多个项目,体现了国家的高度重视。这些项目的短期目标是提高恶性肿瘤的基因诊断和个体化治疗的准确性和疗效,长远目标是在这一过程中推动对各类复杂疾病起因、发病机制、预防 and 治疗的深入理解,探索各种疾病的新型临床诊断技术和标准,建立疾病和健康知识库,制定诊疗规范和指南等。这些任务和目标的实现,必将为我国的医学科学和医疗体系的发展带来新的变革,同时也在相关知识的理论储备、学科交叉、技术革新、发展趋势和实际应用等方面给相关科研人员和医务工作者提出了更高的要求。

在这种大背景下,由詹启敏院士和上海交通大学出版社一起推动了“精准医学出版工程”系列图书的出版,以期为科研人员、医务工作者和相关专业学生提供系统而全面的最新理论知识、技术方法、信息资源和应用实例,成为他们开展科学研究和指导临床诊断治疗的得力工具书,推动精准医学发展和早日实现应用。首批出版的丛书以支撑性和基础性为主,涵盖精准医学总论、生物样本库、多个组学分册、大数据及新技术等方面的内容。

本书共分9章。在前两章里,我们系统地介绍了基因组学和基因测序技术的发展及从基因组学到精准医学的转折。后面七章逐一介绍基因组学各方面的最新进展和部分研究成果及其在精准医学中的应用,内容涵盖基因组科学与信息的主要分支领域(转录组研究部分将在《转录组学与精准医学》分册中阐述),并举例说明主要成就和事件、方法和技术细节、国际国内主要基因组研究计划等,深入浅出地描述了基因组学在精准医学中的可能应用。本书还分章叙述了基因组生物信息学、群体遗传学、肿瘤及其他复杂疾病、药物基因组学、微生物宏与泛基因组学、比较基因组学在精准医学研究框架下的理论基础、研究方法和应用实例。其中,许多章节都提到了相同的国内外主要相关基因组研究计划,部分章节还提到了相同的公共数据库资源信息,本书并没有对这些看似重复的内容做删减,因为每一部分对这些项目和数据有着不同的写作侧重,尤其是在不同案例中的实际应用。这些信息无疑对加深读者印象和拓展思路是有益的。

精准医学不仅仅是一个新概念,也是基因组学研究路线图的终极目标。我们已经尽可能地收集和梳理当前与此相关的最新研究成果和研究方法,集中展示于本书中,达

到了构思本书时的既定目标。

本书由中国科学院北京基因组研究所于军研究员主持编著,编写组由中国科学院北京基因组研究所、中国科学院重庆绿色智能技术研究院、瑞典哥德堡大学萨尔格林斯卡医院、上海交通大学医学院附属瑞金医院卢湾分院的研究人员组成。其中第1章由于军执笔,第2章由王德强、黄绮梦、于军执笔,第3章由肖景发、张哲文、时硕、孟庆仁执笔,第4章由张永彪执笔,第5章由舒畅执笔,第6章由顾明亮执笔,第7章由吴双秀执笔,第8章由吴浩、华沙、康禹执笔,第9章由孟庆姝、胡松年执笔。吴双秀对本书文字进行了校对工作。

本书引用了一些作者的论著及其研究成果,在此向他们表示衷心的感谢!

由于作者水平有限,对于书中存在的缺点甚至错误,恳请读者批评指正。

编著者

2017年11月于北京

目 录

1	生物医学：从“人类基因组计划”到“精准医学计划”	001
1.1	生物医学：学科汇聚的必然产物	001
1.1.1	生物医学是生命科学的核心学科	004
1.1.2	大科学计划的成功范例与生物医学路线图	007
1.1.3	生物医学必须直接聚焦人类社会的健康保障需求	009
1.2	基因组学和精准医学研究技术的发展轨迹	010
1.2.1	基本基因操作技术的发展轨迹	011
1.2.2	生物医学规模化技术的兴与衰	012
1.2.3	未来生物医学技术的愿景	013
1.3	走向成熟：用历史的眼光看“人类基因组计划”	014
1.3.1	人类基因组计划的时代背景	014
1.3.2	人类基因组计划与其划时代意义	015
1.4	后“人类基因组计划”探索性项目的缘由和沿革	018
1.4.1	国际人类基因组单体型图计划和千人基因组计划	018
1.4.2	人类基因组 DNA 元件百科全书计划	019
1.4.3	医学测序的必要性	019

1.5 “精准医学计划”：“人类基因组计划”真正的续集	020
1.5.1 精准医学概念的提出	020
1.5.2 走向精准医学：构建生物医学和疾病新分类法的知识网络	021
1.5.3 当前精准医学的研究目标	022
1.6 生物医学的新思考	023
1.6.1 生命系统的复杂性	023
1.6.2 多系统生物学的新思考：多流生物学	024
1.6.3 多流生物学的应用	028
1.7 生命科学和生物医学的整体观问题	028
1.7.1 生物医学的研究对象	029
1.7.2 生物医学研究的模式生物	030
1.7.3 中国生物医学的学科发展	031
1.8 小结	033
参考文献	033

2 走向精准医学：DNA 测序技术的现状和未来

035

2.1 第一代 DNA 测序技术	036
2.1.1 化学降解法	036
2.1.2 双脱氧链终止法	037
2.1.3 荧光自动测序技术	038
2.1.4 杂交测序技术	039
2.2 第二代 DNA 测序技术	040
2.2.1 454 测序技术	040
2.2.2 Solexa 测序技术	042
2.2.3 SOLiD 测序技术	044
2.3 第三代 DNA 测序技术	045
2.3.1 Helicos 单分子测序技术	045
2.3.2 Pacific Biosciences 单分子测序技术	046