

现代生物技术前沿

HANDBOOK OF GENOME RESEARCH

Genomics, Proteomics, Metabolomics,
Bioinformatics, Ethical and Legal
Issues

〔加拿大〕 C.W. 森森 主编

谢东等译

基因组研究手册

——基因组学、蛋白质组学、代谢组学、
生物信息学、伦理和法律问题



科学出版社

www.sciencep.com

中国科学院植物研究所

Handbook of Genomics Research

Genomics Research Handbook
Genomics Research Handbook
Genomics Research Handbook

基因组研究手册
基因组研究手册
基因组研究手册

基因组研究手册

基因组学、基因组学、基因组学
基因组学、基因组学、基因组学



中国科学院出版社

现代生物技术前沿

Handbook of Genome Research

Genomics, Proteomics, Metabolomics,
Bioinformatics, Ethical and Legal Issues

基因组研究手册

——基因组学、蛋白质组学、代谢组学、
生物信息学、伦理和法律问题

科学出版社

北京

图字：01-2006-7417 号

内 容 简 介

本书旨在为读者引入“组学”的各种基本概念，了解其基本内容，介绍各种常用的“组学”技术，并结合对实例的引用和阐述，展示“组学”的应用及前景。同时，就“组学”的迅猛发展所带来的社会伦理问题，进行了开放性的讨论。基于上述目的，本书分为四个部分：第一部分，关键的生物，对“组学”进行了概括介绍；第二部分，基因组和蛋白质组技术，结合应用实例，详细地介绍了基因组学和蛋白质组学的技术发展；第三部分，生物信息学，对这一在组学研究中的应用中不可或缺的工具进行了细致全面的阐述；第四部分，伦理、法律和社会问题，归纳和展示了由“组学”的迅速发展所带来的不可避免的社会问题，并开展了多方面、多视角的讨论。

本书可供分子生物学、生物技术、医药卫生，以及生命科学领域相关科研、技术人员，研究生和高年级本科生参考使用。

Originally published in the English language by WILEY-VCH Verlag GmbH & Co. KGaA, Boschstraße 12, D-69469 Weinheim, Federal Republic of Germany, under the title “Handbook of Genome Research. Genomics, Proteomics, Metabolomics, Bioinformatics, Ethical and Legal Issues”.

Copyright 2005 by WILEY-VCH Verlag GmbH & Co. KGaA.

图书在版编目(CIP)数据

基因组研究手册：基因组学、蛋白质组学、代谢组学、生物信息学、伦理和法律问题/（加拿大）森森（Sensen, C. W.）主编；谢东等译。—北京：科学出版社，2009

书名原文：Handbook of Genome Research. Genomics, Proteomics, Metabolomics, Bioinformatics, Ethical and Legal Issues

（现代生物技术前沿）

ISBN 978-7-03-023119-2

I. 基… II. ①森…②谢… III. 基因组-研究 IV. Q343.1

中国版本图书馆 CIP 数据核字（2008）第 152590 号

责任编辑：罗 静 莫结胜/责任校对：刘亚琦

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>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2009 年 3 月第 一 版 开本：787×1092 1/16

2009 年 3 月第一次印刷 印张：31 3/4 插页：4

印数：1—2 500 字数：717 000

定价：95.00 元

（如有印装质量问题，我社负责调换〈路通〉）

译者序

21 世纪的生物学研究，对每一位生物学工作者而言，既是机遇，又是挑战。基因组学是近十几年来的生物科学研究热点。基因组密码的破译，展开了生命科学研究的崭新篇章。

随着多个生物，特别是人类基因组测序工作的初步完成，基因组生物学的研究由最初的结构基因组学向功能基因组学转移。从单纯的对于基因结构和表达的研究，发展为对其在整个基因组中所扮演的角色和功能的关注，从整体水平上分析基因组的作用。这种对于功能的侧重，也引发了相应的新“组学”的产生和发展，如蛋白质组学、代谢组学。这些新兴的热点组学和初始的基因组学一起构建了“组学”这一体现生命科学新时代特色的研究系统。而这些新兴领域的发展，也孕育和促成了生物信息学的产生和发展。并且，从伦理道德的角度和层面上，引起了社会各界的关注和讨论。

然而，无论是对“组学”技术的阐述，或者对与其密切相关的生物信息学的介绍，还是对“组学”时代伦理道德问题的讨论，在国内的出版物中，仅散见于各种相关的专著和教材，或者杂志刊物中，并没有系统地整合。本书以简明通俗的文字，不仅详细地介绍了在各“组学”中的多种概念、技术、实验模型，以及它们对于各“组学”发展的影响，还引入了对于生物信息工具和分析策略的介绍，同时，对生命科学研究的迅速发展所引发的一些社会问题展开了探讨。因此，它不仅显示出科学性，而且对于从事生命科学的工作者而言，还具有学术参考价值。同时，也具有通俗性，对于非生物专业的读者，也不失为一本可读性极强的读物。

本书的翻译工作是由中国科学院上海生命科学研究院营养科学研究所谢东研究组的师生共同完成的。参加本书翻译工作的有谢东研究员、佟向军博士、王岩博士、刘冬平博士，以及研究组博士生孙志坚、李晶晶、冯宇雄、邓跃臻、季小丹、李果、石硕、蔡震、赵江沙，及其他人员彭丹妮、陈刚和熊渊。全书由谢东研究员负责组织和审校工作。希望本书的翻译和出版，对于我国的后基因组学研究的发展，起到良好的促进作用。

在翻译中若有疏漏和不当之处，欢迎指正。

中国科学院上海生命科学研究院营养科学研究所

谢 东

2008 年 8 月 8 日

前 言

在近十五年里，生命科学研究，尤其在生物学和医学领域，经历了剧变。1995年第一个微生物基因组的测序完成，引发了一系列变革。如今，我们将完成数百个基因组的测序，包括人类的和其他更多待测的物种。虽然基因组测序几乎已经成为一项常规技术，而在最初对这一工作的美好憧憬，包括相信这一生命的蓝图能够提供给我们大量信息这一信念，已渐变黯淡。

现在，人们已经意识到，要获得关于生命的基因组构建方面的知识，就必须要了解基因表达模式和生物体所有蛋白质的详细功能。因此，需要花费相当长的时间，才能在对生物体的全面认识上取得重要进展。这就引发了多种“组学”的发展，包括基因组学、蛋白质组学、代谢物组学和代谢组学。四年前，一本以“组学”命名的杂志面世，成为了“组学”时代的特征。

对于“组学”技术的全面介绍是本书的主要内容，其中，包括对研究基因组结构、基因表达模式、蛋白质组分和蛋白质翻译后修饰的专用技术的详细描述。主要的模式生物和对它们的生物研究是本书的另一个重点。此外，还有一些章节介绍了在大规模实验中必不可少的生物信息学工具和分析策略。

公众对于生物科学研究最新进展的关注日益增加，在如何面对这一新的研究领域的问题上，引发了激烈的讨论。这个涉及多个社会团体的讨论，亦在本书中得以反映。一些章节展示了研究的社会效应，以及利用新技术或利用大规模实验得到的新数据所带来的发展。我们必须明确，没有人可以忽视这个问题，因为它已经对每个人的日常生活都产生了直接或间接的影响。

大规模研究的新浪潮将在未来使人类受益匪浅，虽然在很多情况下，我们离这一理想还有着相当的距离。我们前行的每一步，都需慎重衡量这一领域的前途和危险。而本书，努力引入和展示这些不仅为科学家，也为社会领导者所探讨和关注的问题。

我们诚挚地感谢 Andrea Pillmann 博士和德国 Weinheim 的 Wiley-VCH 的工作人员，他们在此书的筹备工作中，显示出极大的耐心。有赖于他们的有益建议，本书才得以顺利出版。

Christoph W. Sensen
Calgary, May 2005

参编者名单

Lothar Altschmied
Institute of Plant Genetics and Crop Plant
Research (IPK)
Corrensstr. 3
06466 Gatersleben
Germany

Mélanie Arbour
MicroArray Laboratory
National Research Council of Canada
Biotechnology Research Institute
6100 Royalmount Avenue
Montreal
Quebec, H4P 2R2
Canada

Mayi Arcellana-Panlilio
Southern Alberta Microarray Facility
University of Calgary
HM 393b
3330 Hospital Drive, N.W.
Calgary
Alberta, T2N 4N1
Canada

Gary D. Bader
Computational Biology Center
Memorial Sloan-Kettering Cancer Center
Box 460
New York, 10021
USA

François Benoit
MicroArray Laboratory
National Research Council of Canada
Biotechnology Research Institute
6100 Royalmount Avenue
Montreal
Quebec, H4P 2R2
Canada

Ernst M. Bergmann
Alberta Synchrotron Institute
University of Alberta
Edmonton
Alberta, T6G 2E1
Canada

Richard Bourgault
Department of Biological Sciences
University of Calgary
2500 University Drive N.W.
Calgary
Alberta, T2N 1N4
Canada

Detlev Buttgerit
Fachbereich Biologie
Entwicklungsbiologie
Philipps-Universität Marburg
Karl-von-Frisch-Straße 8b
35043 Marburg
Germany

Jian Chen
Bristol Myers Squibb Pharmaceutical
Research Institute
311 Pennington-Rocky Hill Road
Pennington
New Jersey, 08534
USA

Ming Chen
Department of Bioinformatics /
Medical Informatics
Faculty of Technology
University of Bielefeld
33501 Bielefeld
Germany

Antoine Danchin
Institut Pasteur
Unité de Génétique des Génomes
Bactériens
Département Structure et
Dynamique des Génomes
28 rue du Docteur Roux
75724 PARIS Cedex 15
France

Daniel B. Davison
Bristol Myers Squibb
Pharmaceutical Research Institute
311 Pennington-Rocky Hill Road
Pennington
New Jersey, 08534
USA

Norman J. Dovichi
Department of Chemistry
University of Washington
Seattle
Washington, 98195-1700
USA

Edna Einsiedel
University of Calgary
2500 University Drive N.W., S318
Calgary
Alberta, T2N 1N4
Canada

Peter J. Facchini
Department of Biological Sciences
University of Calgary
2500 University Drive N.W.
Calgary
Alberta, T2N 1N4
Canada

Abul Fazal
Department of Chemistry
University of Washington
Seattle
Washington, 98195-1700
USA

Horst Feldmann
Adolf-Butenandt-Institut für
Physiologische Chemie der
Ludwig-Maximilians-Universität
Schillerstraße 44
80336 München
Germany

Greg Finak
Department of Biochemistry
McGill University
3775 University St
Montreal
Quebeck, H3A 2B4
Canada

Andreas Freier
Department of Bioinformatics /
Medical Informatics
Faculty of Technology
University of Bielefeld
33501 Bielefeld
Germany

His Excellency Dr. Gebhard Fürst
Bischof von Rottenburg-Stuttgart
Postfach 9
72101 Rottenburg a. N.
Germany

Paul Gordon
University of Calgary
Department of Biochemistry and
Molecular Biology
3330 Hospital Drive N.W.
Calgary
Alberta, T2N 4N1
Canada

Roger C. Green
Faculty of Medicine
Memorial University of Newfoundland
St. Johns
Newfoundland, A1B3Y1
Canada

Michael Hallett
Department of Biochemistry
3775 University St
McGill University
Montreal, H3A 2B4
Canada

Ralf Hofestädt
Department of Bioinformatics /
Medical Informatics
Faculty of Technology
University of Bielefeld
33501, Bielefeld
Germany

Christopher W.V. Hogue
Dept. Biochemistry
University of Toronto and the
Samuel Lunenfeld Research Institute
Mt. Sinai Hospital
600 University Avenue
Toronto, ON M5G 1X5
Canada

Hervé Hogues
MicroArray Laboratory
National Research Council of Canada
Biotechnology Research Institute
6100 Royalmount Avenue
Montreal
Quebec, H4P 2R2
Canada

Patrick G. Humphrey
LI-COR Inc.
4308 Progressive Ave.
P.O. Box 4000
Lincoln
Nebraska, 68504
USA

Gregor Jansen
Department of Biochemistry
McGill University
3655 Promenade Sir William Osler
Montreal
Quebec, H3G 1Y6
Canada

Doris Jording
Fakulät für Biologie
Lehrstuhl für Genetik
Universität Bielefeld
33594 Bielefeld
Germany

Jörn Kalinowski
Fakulät für Biologie
Lehrstuhl für Genetik
Universität Bielefeld
33594 Bielefeld
Germany

Hans-Peter Klenk
e.gene Biotechnologie GmbH
Pöckinger Fußweg 7a
82340 Feldafing
Germany

Bartha Maria Knoppers
University of Montreal
3101, Chemin de la Tour
Montreal
Quebeck, H3C 3J7
Canada

James Kraly
Department of Chemistry
University of Washington
Seattle
Washington, 98195-1700
USA

Manfred Kröger
Institut für Mikro- und Molekularbiologie
Justus-Liebig-Universität
Heinrich-Buff-Ring 26-32
35392 Giessen
Germany

Sheena Lambert
Department of Biochemistry and
Molecular Biology
University of Calgary
3330 Hospital Drive N.W.
Calgary
Alberta, T2N 4N1
Canada

David Michels
Department of Chemistry
University of Washington
Seattle
Washington, 98195-1700
USA

Lyle R. Middendorf
LI-COR Inc.
4308 Progressive Ave.
P.O. Box 4000
Lincoln
Nebraska, 68504
USA

Narasimhachari Narayanan
VisEn Medical, Inc.
12B Cabot Road
Woburn
Massachusetts, 01801
USA

Morag Park
Department of Biochemistry
McGill University
3775 University St.
Montreal
Quebec, H3A 2B4
Canada

François Pepin
Department of Biochemistry
McGill University
3775 University St.
Montreal
Quebec, H3A 2B4
Canada

Stephanie Pollock
Department of Biochemistry
McGill University
3655 Promenade Sir William Osler
Montreal
Quebec, H3G 1Y6
Canada

Alfred Pühler
Fakulät für Biologie
Lehrstuhl für Genetik
Universität Bielefeld
33594 Bielefeld
Germany

Renate Renkawitz-Pohl
Fachbereich Biologie,
Entwicklungsbiologie
Philipps-Universität Marburg
Karl-von-Frisch-Straße 8b
35043 Marburg
Germany

Peter Rice
European Bioinformatics Institute
Wellcome Trust Genome Campus
Hinxton
Cambridge, CB10 1SD
UK

Tracey Rigby
MicroArray Laboratory
National Research Council of Canada
Biotechnology Research Institute
6100 Royalmount Avenue
Montreal
Quebec, H4P 2R2
Canada

Stephen C. Roemer
Fisher Scientific
Chemical Division
One Reagent Lane
Fairlawn
New Jersey, 07410
USA

Clémentine Sallée
University of Montreal
3101, chemin de la tour
Montreal
Quebeck, H3C 3J7
Canada

Christa Schleper
Department of Biology
University of Bergen
Jahnebakken 5
Box 7800
5020 Bergen
Norway

David C. Schriemer
Dept. of Biochemistry and
Molecular Biology
University of Calgary
3330 Hospital Drive N.W.
Calgary
Alberta, T2N 4N1
Canada

Agnieszka Sekowska
Institut Pasteur
Unité de Génétique des
Génomés Bactériens
Département Structure et
Dynamique des Génomés
28 rue du Docteur Roux
75724 Paris Cedex 15
France

Christoph W. Sensen
Faculty of Medicine
Sun Center of Excellence for
Visual Genome Research
University of Calgary
3330 Hospital Drive NW
Calgary
Alberta, T2N 4N1
Canada

Lorraine Sheremeta
Health Law Institute at the
University of Alberta
University of Alberta
402 Law Centre
Edmonton
Alberta, T6G 2H5
Canada

Julie Stromer
University of Calgary
Department of Biochemistry and
Molecular Biology
3330 Hospital Drive NW
Calgary
Alberta, T2N 4N1
Canada

Daniel C. Tessier
IatroQuest Corporation
1000 Chemin du Golf
Verdun
Quebec, H3E 1H4
Canada

David Y. Thomas
Department of Biochemistry
McGill University
3655 Promenade Sir William Osler
Montreal
Quebec, H3G 1Y6
Canada

Theerayut Toojinda
Rice Gene Discovery
National Center for Genetic Engineering
and Biotechnology
Kasetsart University
Kamphangsaen
Nakorn Pathom, 73140
Thailand

Somvong Tragoonrungrung
Rice Gene Discovery
National Center for Genetic Engineering
and Biotechnology
Kasetsart University
Kamphangsaen
Nakorn Pathom, 73140
Thailand

Alexander H. Treusch
University of Bergen
Department of Biology
Jahnebakken 5
Box 7800
5020 Bergen
Norway

Andrei Turinsky
University of Calgary
Department of Biochemistry and
Molecular Biology
3330 Hospital Drive NW
Calgary
Alberta, T2N 4N1
Canada

Apichart Vanavichit
Center of Excellence for Rice Molecular
Breeding and Product Development
National Center for
Agricultural Biotechnology
Kasetsart University
Kamphangsaen
Nakorn Pathom, 73140
Thailand

Hans J. Vogel
Department of Biological Sciences
University of Calgary
Calgary
Alberta, T2N 1N4
Canada

Aalim M. Weljie
Chenomx Inc.
#800, 10050 - 112 St.
Edmonton
Alberta, T5K 2J1
Canada

David S. Wishart
Departments of Biological Sciences and
Computing Science
University of Alberta
Edmonton
Alberta, T6G 2E8
Canada

Shu Jian Wu
Bristol Myers Squibb Pharmaceutical
Research Institute
311 Pennington-Rocky Hill Road
Pennington
New Jersey, 08534
USA

Katherine G. Zulak
Department of Biological Sciences
University of Calgary
2500 University Drive N.W.
Calgary
Alberta, T2N 1N4
Canada

Emily Xu
Department of Biochemistry and
Molecular Biology
University of Calgary
3330 Hospital Drive N.W.
Calgary
Alberta, T2N 4N1
Canada

目 录

译者序

前言

参编者名单

第一部分 关键的生物

1 模式生物的基因组工程	3
1.1 绪论	3
1.2 特定模式原核生物的基因组工程	4
1.3 真核模式生物的基因组计划	17
1.4 总结	30
参考文献	32
2 环境基因组：研究未获培养微生物的新手段	36
2.1 绪论：为什么需要用新的方法来研究微生物的基因组	36
2.2 环境基因组学：方法学	37
2.3 首试：海洋环境基因组学	38
2.4 确定群体的环境基因组学：生物层和微生物层	39
2.5 用环境基因组学研究土壤微生物	40
2.6 生物技术方面	42
2.7 总结和展望	43
参考文献	43
3 基因组学在植物学中的应用	46
3.1 绪论	46
3.2 植物基因组	47
3.3 表达序列标签	50
3.4 基于 DNA 芯片的基因表达谱	51
3.5 蛋白质组学	53
3.6 代谢物组学	54
3.7 功能基因组学	56
3.8 总结	59
参考文献	60
4 人类遗传疾病	63
4.1 绪论	63
4.2 遗传对人类健康的影响	65

4.3 基因组和单基因缺陷·····	65
4.4 基因组学和多基因疾病·····	68
4.5 癌症的遗传基础·····	71
4.6 心血管疾病的遗传学·····	74
4.7 总结·····	76
参考文献·····	77

第二部分 基因组和蛋白质组技术

5 基因组定位和定位克隆及其在植物科学中的应用 ·····	85
5.1 绪论·····	85
5.2 基因组作图·····	85
5.3 定位克隆·····	89
5.4 比较作图与定位克隆·····	94
5.5 后基因组时代的遗传作图·····	98
参考文献·····	100
6 DNA 测序技术 ·····	106
6.1 绪论·····	106
6.2 概观 Sanger 的双脱氧测序·····	107
6.3 荧光染料化学·····	107
6.4 DNA 测序的生物化学·····	113
6.5 荧光 DNA 测序仪器·····	118
6.6 DNA 序列分析·····	126
6.7 力求成本为 1000 美元的基因组 DNA 测序法·····	131
参考文献·····	135
7 用于生物研究的蛋白质组学和质谱 ·····	150
7.1 绪论·····	150
7.2 蛋白质组学研究中样品的定义·····	152
7.3 新的发展——临床蛋白质组学·····	156
7.4 质谱——基本的蛋白质组学技术·····	156
7.5 样品驱动的蛋白质组学进程·····	162
7.6 总结·····	170
参考文献·····	170
8 利用毛细管电泳技术进行蛋白质组分析 ·····	175
8.1 绪论·····	175
8.2 毛细管电泳技术·····	175
8.3 用于蛋白质分析的毛细管电泳·····	178
8.4 单细胞分析·····	181
8.5 二维分离·····	182

8.6 总结	183
参考文献	184
9 面向实验室的 DNA 芯片制作策略	185
9.1 绪论	185
9.2 数据库	189
9.3 高通量的 DNA 合成	191
9.4 扩增子产生	193
9.5 基因芯片	194
9.6 检测和扫描芯片	194
9.7 总结	195
参考文献	196
10 DNA 微阵列的应用原理	198
10.1 绪论	198
10.2 定义	199
10.3 阵列的类型	199
10.4 阵列的产生	199
10.5 阵列的检测	201
10.6 数据分析	208
10.7 微阵列的记录	210
10.8 微阵列技术在肿瘤研究中的应用	211
10.9 总结	212
参考文献	212
11 酵母双杂交技术	216
11.1 绪论	216
11.2 经典的酵母双杂交系统	217
11.3 双杂交系统的变型	217
11.4 膜酵母双杂交系统	220
11.5 对双杂交结果的解释	223
11.6 总结	224
参考文献	224
12 结构基因组学	226
12.1 绪论	226
12.2 蛋白质晶体学与结构基因组学	226
12.3 NMR 与结构基因组学	233
12.4 总结	239
参考文献	240

第三部分 生物信息学

13 DNA 技术的生物信息学工具	247
13.1 绪论	247
13.2 比对方法	247
13.3 序列比较方法	253
13.4 一致法	255
13.5 简单序列屏蔽	256
13.6 罕见的序列构成	256
13.7 重复鉴定	257
13.8 序列模式的检测	258
13.9 限制酶酶切位点和启动子共有序列	262
13.10 EMBOSS 的展望	268
参考文献	269
14 蛋白质组学技术中的软件工具	270
14.1 绪论	270
14.2 蛋白质鉴定	270
14.3 蛋白质性质预测	279
参考文献	290
15 生物信息学在药物发现和研发中的应用	295
15.1 绪论	295
15.2 数据库	295
15.3 药物靶标发现中的生物信息学	303
15.4 化合物扫描和毒理基因组学的支持	305
15.5 药物发展中的生物信息学	310
15.6 总结	314
参考文献	315
16 通过图形的基因组数据表示法: MAGPIE/Bluejay 系统	320
16.1 绪论	320
16.2 MAGPIE 绘图系统	321
16.3 分级的 MAGPIE 显示系统	323
16.4 概观图像	324
16.5 编码区显示	326
16.6 编码序列功能显示	330
16.7 二级基因组背景图像	333
16.8 Bluejay 数据可视系统	337
16.9 Bluejay 系统结构	337
16.10 Bluejay 显示和数据探测	339