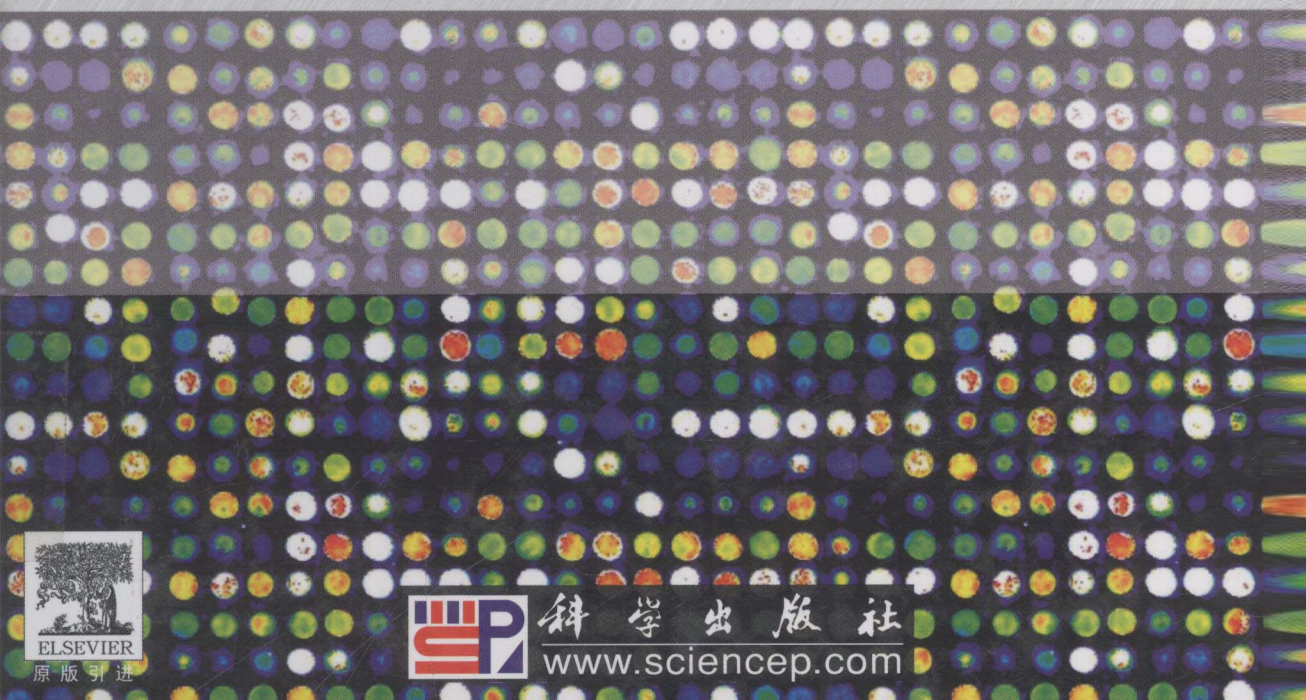




New High Throughput Technologies for
DNA Sequencing and Genomics

高通量DNA测序和 基因组学新型技术

Keith R. Mitchelson



科学出版社

www.sciencep.com

New High Throughput Technologies for DNA Sequencing and Genomics

高通量 DNA 测序和 基因组学新型技术

Editor

Keith R. Mitchelson

Capitalbio Corporation;
National Engineering Research Center
for Beijing Biochip Technology
18 Life Science Parkway
Changping District
Beijing 102206
China

AND

Medical Systems Biology Research Center
Tsinghua University School of Medicine
Beijing 100084

科学出版社

北京

图字：01-2007-4865 号

This is an annotated version of

New High Throughput Technologies for DNA Sequencing and Genomics

Keith R. Mitchelson

Copyright © 2007, Elsevier B. V.

ISBN-13: 978-0-444-52223-8

ISBN-10: 0-444-52223-9

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

高通量 DNA 测序和基因组学新型技术 = New High Throughput Technologies for DNA Sequencing and Genomics: 英文 / (澳) 米切尔森 (Mitchelson, K. R.) 编著. 北京: 科学出版社, 2007

ISBN: 978-7-03-020346-5

I. 高… II. 米… III. ①脱氧核糖核酸-英文②基因组-英文
IV. Q523 Q343.1

中国版本图书馆 CIP 数据核字(2007)第 155609 号

责任编辑: 孙红梅 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.sciencecp.com>

中国科学院印刷厂 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: 26 插页: 6

印数: 1—1 500 字数: 617 000

定价: 98.00 元

(如有印装质量问题, 我社负责调换〈科印〉)

导 读

20 世纪 70 年代，DNA 双脱氧测序和化学测序方法分别由 Sanger 和 Gilbert 建立，随后即迅速在基因组学和功能基因组学研究中得到了广泛的应用，现在已经成为分子生物学实验室的常规研究手段和方法。这项技术获得了 1980 年的诺贝尔奖。DNA 测序技术是人类探知大自然生命奥秘的基本工具，所有生物基因组序列的数据都来自于 DNA 测序。上世纪 80 年代启动的由多个国家参加的人类基因组计划，对测序技术提出了高通量、自动化和低成本的要求。这一伟大计划的完成，大大地推动了 DNA 测序技术的发展，开发出了一系列自动化和半自动化的测序仪，真正实现了高通量、高质量、低成本的 DNA 测序。世界上各大测序中心和相关的实验室建立了规模化的 DNA 测序技术平台，DNA 测序的能力大大提高，每天可以完成成千上万条 DNA 序列的精确测定，加快了人类基因组计划的进程，从而使人类基因组测序工作得以提前完成。

在人类基因组测序完成以后，许多人认为基因组测序工作可以告一段落，大的测序中心的测序仪将派不上用场。但实际情况是，人类基因组测序完成以后，为我们提供了更多的基因序列并开辟了新的研究领域。继人类基因组计划之后，科学家们又相继完成了人类单倍体计划、水稻基因组计划以及其他模式生物的基因组测序，这些生命体的 DNA 测序量远远超过了预期。去年美国提出了人类癌症基因组计划，为 50 种最常见癌症中的每一种测出所有的致癌突变目录，建立包含上万个肿瘤样本的基因库，把肿瘤细胞与正常细胞的基因图进行比较。他们力图通过这一计划攻克癌症，寻找到癌症的致病基因，为药物开发和发展治疗癌症的新手段提供靶点。这一计划预计将持续 10 年左右。这些计划主要依赖于测序技术，所有测序工作的总规模预计是人类基因组计划的 100 倍。

尽管 DNA 测序技术得到了非常大的发展，自动化程度在不断的改善，技术日见成熟，但是目前的大部分测序技术一次只能读取 1000 个左右的碱基，需要繁琐的 DNA 样品的制备和分离步骤，在测序的通量和成本方面还不能完全满足大规模测序的要求。人类癌症基因组计划的策划和实施以及个体化医疗，也对测序技术有巨大的远景需求。人们对更快、更好的基因组测序技术的期待空前强烈。因而科学家们一直在寻求革命性的技术革新：高速测序技术。各大生物技术公司都不惜重金来开发高速测序技术，相互之间展开了激烈角逐。从 2000 年开始，DNA 高速测序方面陆续取得了可喜的进展，如 454 生命科学开发了 454 的纳米测序仪，已于去年上市，并被世界上几个大的测序中心所应用；最近公布了用 454 测序仪测定 Watson 的基因组序列，整个基因组测序工作历时不到 2 年，仅耗资 200 万美元。Illumina 公司收购了 Solexa，推出了快速测序仪器，测序成本更为低廉。

DNA 测序技术一直以来备受生命科学研究者的重视，有众多的综述介绍 DNA 测序的原理和实验操作，如由 Elsevier 集团编辑出版的 *Automated DNA Sequencing and Analysis*，Jones and Bartlett Publishers, Inc 编辑出版的 *DNA Sequencing: Optimi-*

zing the Process and Analysis, CRC 编辑出版的 *Analytical Techniques in DNA Sequencing* 等。这些著作主要介绍 DNA 测序技术的原理和实验操作,最新进展涉及的不多,难以全面和系统地反映 DNA 测序技术的最新成果。

让我们感到欣慰的是 Elsevier 集团在 2007 年出版了 *New High Throughput Technologies for DNA Sequencing and Genomics* 一书,本书全面地综述了 DNA 测序技术的最新进展。

本书的主编 Keith R. Mitchelson 博士是中国博奥生物有限公司和生物芯片北京国家工程研究中心的研究员,也是清华大学医学院系统医学研究中心的访问学者。生物芯片北京国家工程研究中心是由国家发展与改革委员会支持建立的国家级生物芯片技术研发中心之一,主要从事生物芯片技术、仪器设备的研究开发与产业化。该中心聚集了一批生物芯片技术和基因组学研究领域的知名学者。本书的作者之一程京教授,是国际上微流体芯片的技术专家,作为中国生物芯片的发起人之一,领导和开创了生物芯片技术的研究与开发,对中国生物芯片技术的发展和产业化做出了巨大的贡献。

本书共分为 12 章,四大部分。

第一部分:现有的技术(第 1~4 章)。概述了 DNA 测序的进展和已有的成熟技术。

第二部分:基于 DNA 合成的测序技术平台(第 5~6 章)。介绍了 454 生命科学皮升测序系统,包括测序的基本原理、样品的制备、数据的读取、分析,以及其在基因组学中的应用;第 6 章介绍了通过 DNA 合成测序的集成系统。

第三部分:单分子测序技术(第 7~9 章)。这是今后测序的一个发展方向,该部分介绍了目前最新的几种单分子测序的方法,主要有基于单分子荧光显微技术的循环合成单分子测序,基于单个 DNA 链的纳米级核苷酸序列快速测序和全基因组水平的单分子测序系统。

第四部分:序列的验证与分析(第 10~12 章)。序列读取以后的拼接是一个难度较大的工作,这部分介绍了通过诱导突变与测序相结合的方法,以利于短读长获得全新百万级的 DNA 片段序列的原理以及目标序列的拼接方法、全基因组测序和拼接的方法、存在的问题和解决的方案。本书的最后介绍了如何测定具有高污染风险性样品的古 DNA 及环境 DNA 的序列。

本书的主要特色有以下几个方面:

(1) 前沿性

前沿性是本书最大的一个特点。编者站在 DNA 测序技术发展的最前沿,总结了近 3~5 年测序技术方面的新进展,对新技术的介绍比较翔实,配以彩图加以说明,易于理解,引用了最新的文献和高水平的 DNA 测序工作。而对基本的测序技术原理和操作的介绍比较少,避免了与其他书籍的雷同。

(2) 系统性

本书的系统性较强,第一部分简要介绍了测序技术的进展和已经具有的成熟技术,随后重点突出了新建立的基于 DNA 合成和单分子测序的技术以及测序以后的序列的拼接。每部分的基本原理介绍得非常清楚,配合了大量的图解加以说明,对技术的基本特点以及应用和应用过程中需要注意的问题也都进行了比较详细的阐述,便于读者掌握比

较系统和全面的知识。

(3) 实用性

本书是一本 DNA 测序前沿技术的参考书，结合以前出版的 DNA 测序技术的书籍，读者可以对测序的原理、操作步骤和最新的技术发展有比较全面的了解，实用性非常强。

我国在基因组学研究中一直非常活跃，积极参与国际项目，如参与了国际人类基因组计划，圆满完成了其中 1% 的任务，承担了国际 HapMap 计划 10% 的任务。此外还进行了以我国科学家为主的测序工作，如国际水稻基因组计划、家猪基因组计划、鸡基因组计划、家蚕基因组测序、血吸虫表达标签序列测序和基因组测序、钩端螺旋体基因组测序等，这些工作在国际上都产生了较大的影响。在“十一五”国家高技术研究发展计划（863 计划）和相关项目的支持下，中国的肿瘤基因组项目和元基因组项目得以启动。在 DNA 测序技术方面，我国科学家也积极开展高通量和快速测序技术的研究，如东南大学陆祖宏教授领衔的课题组正在开展快速低成本人类基因组测序技术和装备的研究，上海交通大学也在进行 DNA 自动化测序技术的研究。这些具有自主知识产权的技术取得突破后，对于提高我国在基因组学研究中的地位，增强国际竞争能力具有重要的意义。

《高通量 DNA 测序和基因组学新型技术》(*New High Throughput Technologies for DNA Sequencing and Genomics*) 一书比较全面和系统地总结了最新的测序技术的进展，因此本书的出版发行，对于高等院校和科研院所从事基因组学研究的科技人员和在读的研究生，都是一本具有系统性和前沿性的参考书。

肖华胜

生物芯片上海国家工程研究中心

中国科学院上海生命科学研究院系统生物学研究中心
huasheng_xiao@shbiochip.com, xiaohs@sibs.ac.cn

序

三十年前，由 Sanger 和 Gilbert 分别建立了 DNA 双脱氧测序和化学测序技术。现在 DNA 测序的规模逐渐扩大，从每天能读几千个碱基对序列的小规模测序，发展到如今每天能够进行成千上万个序列的精确测定，成为数十亿美元的“产业”。世界上几个大的测序中心建立了大规模的基因组测序平台，支持几乎所有大学和大型医院的 DNA 测序工作。国际生物信息学中心拥有存储约一千五百万碱基对序列信息的能力，我们正处于一个“后基因组”时代。然而 DNA 测序在生命科学和医学的许多领域，仍然是一个主要的诊断和研究手段。这种强劲的需求，推动了新的测序技术的发展。这些新技术既不用电泳，也不用 Sanger 测序方法，而是通过 DNA 合成的方法进行测序，包括在固相表面上进行的多重平行测序反应，每一个循环的数据通过成像的方法获得。最近有一项由 454 生命科学和马普研究所制定的计划，使用合成法测序技术测定穴居人基因组，该计划显示了这些新的 DNA 测序技术在后基因组学研究中的强大作用。这将成为第二代全人类基因组项目，而且预计将进一步快速提升我们对人类基因组生物学知识的认识，加深我们对人类遗传性疾病机制的了解。

本书汇集了 DNA 测序技术的一些新的进展，全书的框架由一系列综述组成，主要包括 Sanger 和 SBS 测序方法的进展 (Kumar and Fuller)、毛细管电泳和微型集成系统 (Xiong and Cheng)、MS 测序 (Ehrich *et al.*) 和应用 (Mitchelson *et al.*)。还包括了一些最新的进展，如 SBS 技术的新进展 (Margulies *et al.*)、核酸染料 SBS 测序方法 (Edwards *et al.*)、通过循环合成实现单分子测序 (Hebert and Braslavsky)、纳米孔测序 (Lee and Meller)，以及阵列 DNA 的光标测 (Schwartz *et al.*)。在书的最后，讨论了基因组拼接 (McGrath) 中的生物信息学工具、古代和环境 DNA 样品 (Kowalchuk *et al.*) 的测序和 SBS 序列拼接工具 (Keith *et al.*) 等 SBS 测序的新应用。作者真诚希望此书的出版和发行能够引起学生和研究人员对测序这个关键基因组学技术的兴趣，并引发相关的技术革新。

Keith Mitchelson

(肖华胜 译)

Contributors

Numbers in parentheses indicate the pages where the authors' contributions can be found.

P. Adams (303), Department of Mathematics, University of Queensland, St. Lucia, Queensland 4072, Australia

J. J. Austin (357), Department of Environmental Biology, University of Adelaide, North Terrace, Adelaide 5005, South Australia, Australia

Dirk van den Boom (97), Sequenom Corporation, 3595 John Hopkins Court, San Diego, CA 92121, USA

I. Braslavsky (209), Department of Physics and Astronomy, Clippinger 251B, Ohio University, Athens, OH 45701, USA

D. E. Bryant (303), Department of Mathematics, University of Queensland, St. Lucia, Queensland 4072, Australia

J. C. Carter (303), Leukaemia Foundation Queensland Laboratories, Queensland Institute of Medical Research, Herston, Queensland 4006, Australia

J. Cheng (45), Capitalbio Corporation: National Engineering Research Center for Beijing Biochip Technology, Beijing, 18 Life Science Parkway, Changping District, Beijing 102206, China; and the Medical Systems Biology Research Center, Tsinghua University School of Medicine, Beijing 100084, China

D. A. E. Cochran (303), Agen Biomedical Limited, Durbell Street, Acacia Ridge, Queensland 4110, Australia

J. R. Edwards (187), Columbia Genome Center, Columbia University College of Physicians and Surgeons, Russ Berrie Medical Science Pavilion, 1150 St. Nicholas Avenue, New York, NY 10032, USA

M. Ehrich (97), Sequenom Corporation, 3595 John Hopkins Court, San Diego, CA 92121, USA

C. W. Fuller (119), GE Healthcare, 800 Centennial Avenue, Piscataway, NJ 08855, USA

P. S. Gooding (357), Agricultural Division – AGRF, Plant Genomics Centre, University of Adelaide, Hartley Grove, Waite Campus PMB1, Glen Osmond, SA 5064, Australia

D. B. Hawkes (3,303), AGRF, Institute of Molecular Bioscience, University of Queensland, St. Lucia, Queensland 4072, Australia

B. Hebert (209), Department of Physics, McGill University, Rutherford Physics Building 228, 3600 University Street, Montreal, Quebec H3A 2T8, Canada

J. Herschleb (265), Laboratory for Molecular and Computational Genomics, UW Biotechnology Center, Laboratory of Genetics and Department of Chemistry, University of Wisconsin, 425 Henry Mall, Madison 53706, USA

- F. Hillenkamp** (97), Institute for Medical Physics and Biophysics, University of Münster, Robert-Koch-Str. 31, Münster D-48149, Germany
- T. P. Jarvie** (153), 454 Life Sciences Corporation, 20 Commercial Street, Branford, CT 06405, USA
- J. Ju** (187), Columbia Genome Center, Columbia University College of Physicians and Surgeons, Russ Berrie Medical Science Pavilion, 1150 St. Nicholas Avenue, New York, NY 10032, USA
- J. M. Keith** (303), Institute of Molecular Bioscience and Department of Mathematics, University of Queensland, St. Lucia, Queensland 4072, Australia
- D. H. Kim** (187), Columbia Genome Center, Columbia University College of Physicians and Surgeons, Russ Berrie Medical Science Pavilion, 1150 St. Nicholas Avenue, New York, NY 10032, USA
- J. R. Knight** (153), 454 Life Sciences Corporation, 20 Commercial Street, Branford, CT 06405, USA
- G. A. Kowalchuk** (357), Netherlands Institute of Ecology, PO Box 40, 6666 ZG Heteren, The Netherlands
- S. Kumar** (119), GE Healthcare, 800 Centennial Avenue, Piscataway, NJ 08855, USA. Present Address: 21 Muirhead Court, Belle Mead, NJ 08502, USA
- J. W. Lee** (245), Oak Ridge National Laboratory, Oak Ridge, TN 37831-6194, USA
- M. Margulies** (153), 454 Life Sciences Corporation, 20 Commercial Street, Branford, CT 06405, USA
- A. McGrath** (327), Australian Genome Research Facility, University of Queensland, St. Lucia, Queensland 4072, Australia
- A. Meller** (245), The Department of Physics and Biomedical Engineering, Boston University, 44 Cummington Street, Boston, MA 02215, USA
- A. E. Men** (3), AGRF, Institute of Molecular Bioscience, University of Queensland, St. Lucia, Queensland 4072, Australia
- K. R. Mitchelson** (3,303), Capitalbio Corporation: National Engineering Research Centre for Beijing Biochip Technology, 18 Life Science Parkway, Changping District, Beijing 102206, China; and Medical Systems Biology Research Center, Tsinghua University School of Medicine, Beijing 100084, China
- D. C. Schwartz** (265), Laboratory for Molecular and Computational Genomics, UW Biotechnology Center, Laboratory of Genetics and Department of Chemistry, University of Wisconsin, 425 Henry Mall, Madison 53706, WI, USA
- J. F. Simons** (153), 454 Life Sciences Corporation, 20 Commercial Street, Branford, CT 06405, USA
- J. R. Stephen** (357), Agricultural Division – AGRF, Plant Genomics Centre, University of Adelaide, Hartley Grove, Waite Campus PMB1, Glen Osmond, SA 5064, Australia
- R. Turakulov** (3), AGRF, Walter and Eliza Hall Institute for Medical Research, 1G Royal Parade, Parkville, Victoria 3050, Australia

Q. Xiong (45), Capitalbio Corporation: National Engineering Research Center for Beijing Biochip Technology, 18 Life Science Parkway, Beijing 102206, China

S. Zhou (265), Laboratory for Molecular and Computational Genomics, UW Biotechnology Center, Laboratory of Genetics and Department of Chemistry, University of Wisconsin, 425 Henry Mall, Madison 53706, WI, USA

Preface

Since the independent invention of DNA sequencing by Sanger and by Gilbert 30 years ago, it has grown from a small-scale technique capable of reading several kilobase-pair of sequence per day into today's multibillion dollar 'industry', with large Sequencing Centers for large-scale delineation of entire genomes, and supporting DNA sequencing activity at some level at virtually all Universities and larger hospitals throughout the world. We are now in a "post-genomic era" with possibly more than 150 billion base-pair of sequence information held at international Bioinformatics Centers, yet DNA sequencing continues as a major diagnostic and research activity in many areas of life science and medicine. This growth has spurred the development of new sequencing technologies that do not involve either electrophoresis or Sanger sequencing chemistries. Sequencing by synthesis (SBS) involves multiple parallel micro-sequencing addition events occurring on a surface, where data from each round is detected by imaging. The recent plan to sequence a complete Neandertal genome (*Homo sapiens neanderthalensis*) by the 454 Life Sciences and the Max Planck Institute using SBS underlines the relevance of these new sequencing technologies to post-genomic science. This will be the second entire human genome project, and it is expected to radically advance knowledge of the human genomic biology and provide profound new insights into genetic diseases in man.

This volume brings together some of the new developments in DNA sequencing technology. Reviews of complementary developments in Sanger and SBS sequencing chemistries (Kumar and Fuller), capillary electrophoresis and micro-device integration (Xiong and Cheng), MS sequencing (Ehrich *et al.*) and applications (Mitchelson *et al.*) set the framework. Reviews of new developments in SBS technology (Margulies *et al.*), the chemistry of nucleotide-dye SBS sequencing (Edwards *et al.*), and steps toward realizing the sequencing of single DNA molecules by cyclic synthesis (Hebert and Braslavsky), nano-pore sequencing (Lee and Meller) and optical mapping of arrayed DNA (Schwartz *et al.*) indicate the latest advances. Finally, bioinformatics tools for genome assembly (McGrath), sequencing ancient and environmental DNA samples (Kowalchuk *et al.*) and support for SBS sequence assembly (Keith *et al.*) discuss many issues relevant to new applications using SBS sequencing. The authors hope this volume will provide stimulus to both students and researchers interested in this vital field of chemistry and technological innovation.

July 2006

Keith Mitchelson

目 录

撰稿人
序

现有技术

第一章 概述：DNA 测序技术进展	3
第二章 芯片毛细管电泳与系统遗传分析系统	45
第三章 应用 MALDI-TOF 质谱法比较序列分析——利用已知序列寻找新的序列	97
第四章 基于核苷酸偶联染料的 DNA 测序进展	119

合成测序技术平台

第五章 454 生命科学（454 Life Sciences）皮升级测序系统	153
第六章 合成法 DNA 测序的集成系统	187

单分子测序

第七章 单分子荧光显微镜及其在基于循环合成的单分子测序中的应用	209
第八章 基于单个 DNA 链的纳米级核苷酸序列快速测序	245
第九章 全基因组水平的单分子测序系统	265

序列验证和分析

第十章 测序和诱导突变相结合有利于短读长获得全新百万级的 DNA 片段序列	303
第十一章 基因组测序和拼接	327
第十二章 具有高污染风险的样品——古 DNA 和环境 DNA 核酸序列信息的有效测定	357

内容索引	373
------------	-----

（肖华胜 译）

详细目录

第一章 概述：DNA 测序技术进展	3
Keith R. Mitchelson, David B. Hawkes, Rustam Turakulov and Artem E. Men	
1 引言	4
1.1 超高通量测序能力的生物技术意义	6
2 测序前沿技术	9
2.1 毛细管电泳和 Sanger 测序法	9
2.2 高通量毛细管微阵列测序	9
2.3 染料和检测器	10
2.4 微型芯片电泳	11
2.5 基于微型毛细管芯片的毛细管电泳测序	11
2.6 质谱测序	14
3 固相阵列测序装置	15
3.1 超灵敏的检测器和测序仪	15
3.2 合成测序	15
3.3 DNA 单分子测序	19
3.4 杂交重组测序	22
4 技术展望	22
4.1 纳米孔膜	23
4.2 DNA 合成的直接电学检测	25
5 基因组短片段测序应用	25
5.1 基于重测序的基因分型	25
5.1.1 polony 基因分型	26
5.1.2 焦磷酸测序的基因分型	26
5.1.3 多态性比率测序	26
5.1.4 BEAMing	27
5.2 古生物基因组	27
5.3 洞人基因组	28
5.4 元基因组	29
5.5 重复 DNA 序列的 SAM 测序	30
5.6 转录组和表达 RNA 序列分析	31
5.7 MPSS 和基因组分析	34
5.8 光学定位	34
6 小结	35
参考文献	36

第二章 芯片毛细管电泳与系统遗传分析系统	45
Qiang Xiong and Jing Cheng	
摘要	46
1 引言	46
1.1 各种基于芯片的毛细管电泳系统	46
2 芯片设计和流体操作	48
2.1 芯片设计	48
2.2 流体操作	50
3 材料和制作	51
3.1 材料	51
3.2 制作	53
3.2.1 玻璃材料的制作步骤	53
3.2.2 高分子材料的制作步骤	53
4 检测	57
4.1 光学检测	58
4.1.1 激光诱导的荧光检测	58
4.1.2 吸光度检测	60
4.1.3 化学发光检测	60
4.2 电化学检测	60
4.2.1 电流检测	61
4.2.2 电导检测	63
4.2.3 电位检测	63
4.3 质谱	63
5 表面修饰	65
5.1 动态包被	65
5.1.1 玻璃/石英基片的包被	66
5.1.2 PMMA 基片的包被	66
5.1.3 PDMS 基片的包被	66
5.2 永久性包被	67
5.2.1 玻璃/石英基片的永久包被	67
5.2.2 PMMA 基片的永久包被	67
5.2.3 PDMS 基片的永久包被	67
6 应用	68
6.1 核酸分析	68
6.1.1 筛分介质	69
6.1.2 依大小排列 DNA 片段	71
6.1.3 基因分型	71
7 DNA 测序	74

7.1	微芯片 DNA 测序	77
7.2	系统遗传分析系统	79
7.2.1	微芯片的样品准备	80
7.2.2	微芯片的生物反应	80
7.2.3	集成系统	82
7.3	DNA 测序芯片实验室	85
7.3.1	其他的 DNA 测序技术	86
	参考文献	87

第三章 应用 MALDI-TOF 质谱法比较序列分析——利用已知序列寻找新的序列 ... 97

Mathias Ehrich, Franz Hillenkamp and Dirk van den Boom

摘要	97
1 比较测序的概念	98
1.1 群体基因分型	98
2 基于 MALDI-TOF 质谱的核酸序列分析	99
2.1 样品准备	99
3 碱基特异性剪切测定	100
3.1 碱基特异性剪切方法	102
3.2 MassCLEAVE	103
4 比较测序的应用	103
4.1 特征序列鉴定/病原体鉴定	103
4.2 SNP 发现和突变检测	105
4.3 甲基化检测	109
5 小结	112
6 展望	112
6.1 仪器和过程的改进	113
6.2 高分辨率质谱和同位素耗尽的核苷酸	113
6.3 可剪切与不可剪切的核苷酸混合物的使用	114
致谢	115
参考文献	115

第四章 基于核苷酸偶联染料的 DNA 测序进展

Shiv Kumar and Carl W. Fuller

摘要	119
1 引言	119
2 荧光 DNA 测序	121
2.1 单色荧光标记的引物和终止剂	121
2.2 基于荧光共振能量转移 (FRET) 的引物和终止剂	124
3 能量转移染料终止剂	125

3.1	用于“直接上样”DNA测序的带电终止剂	132
3.2	带负电的终止剂	134
3.3	赖氨酸衍生物终止剂	136
3.4	三甲基赖氨酸衍生物终止剂	140
4	末端磷酸标记的核苷酸	144
5	结论	146
	参考文献	146
第五章	454 生命科学皮升级测序系统	153
	Marcel Margulies, Thomas P. Jarvie, James R. Knight and Jan Fredrik Simons	
	摘要	153
1	引言	154
2	454 生命科学皮升级测序系统	155
2.1	样品制备	155
2.2	测序	158
2.3	图像处理	163
2.4	测序准确性	165
2.5	碱基识别	165
2.6	序列排列	168
3	应用	170
3.1	全新序列拼接	170
3.2	测序结果	172
3.3	比较基因组学	173
3.4	PCR 扩增产物的深度测序	174
3.5	规模化能力	181
4	讨论	182
	致谢	184
	参考文献	184
第六章	合成法 DNA 测序的集成系统	187
	John R. Edwards, Dae Hyun Kim and Jingyue Ju	
	摘要	187
1	引言	187
2	基于合成方法的 DNA 测序	189
2.1	DNA 吸附表面化学物质	192
2.2	新型的报告核苷酸	193
2.2.1	SBS 核酸报告基团	195
2.3	3'端羟基的封闭	200
3	结论	203

致谢	203
参考文献	203
第七章 单分子荧光显微镜及其在基于循环合成的单分子测序中的应用	209
Benedict Hebert and Ido Braslavsky	
摘要	210
1 引言	210
2 背景	212
2.1 单分子检测	212
2.2 完全内反射	214
2.3 FRET 原理	218
3 基于循环合成的 DNA 测序	219
3.1 目的	219
3.2 表面处理	221
3.3 多聚酶动力学	222
3.4 测序策略	224
3.4.1 应用 FRET 的循环合成	224
3.4.2 实时成像	226
3.4.3 非 FRET 成像	227
3.4.4 可去除的接头	227
3.4.5 可去除的终止剂	229
3.4.6 多色与单色成像	229
4 数据分析	230
4.1 空间相关性	230
4.2 数据收集——碱基识别	231
4.2.1 强度踪迹	232
4.2.2 单图像数据收集	233
4.3 序列排列	234
5 碱基识别中的误差来源	234
6 性能	237
7 应用	238
8 结论	238
致谢	239
参考文献	239
第八章 基于单个 DNA 链的纳米级核苷酸序列快速测序	245
James Weifu Lee and Amit Meller	
摘要	245
1 引言	246