

新一代基因组测序技术

陈浩峰 主编

科学出版社

北京

内 容 简 介

新一代基因组测序技术是目前生命科学领域中发展非常快的新技术，它的出现极大地推动了生物学、农学和医学诊断等学科各方面的发展，其应用非常广泛。本书以新一代基因组测序技术 Illumina 平台和 PacBio RS 平台为代表，详细阐述了从实验样品处理到数据分析的整个过程，重点介绍了新一代测序技术实践过程中的文库构建与质检、测序仪器操作和测序数据的初步处理及分析。本书是国内第一本详细阐述新一代基因组、转录组等测序技术的中文书，包含了目前最新的技术资料，极具实用性和可操作性。

本书是高等院校师生和科技工作者学习新一代测序技术的最佳参考书。

图书在版编目 (CIP) 数据

新一代基因组测序技术/陈浩峰主编. —北京: 科学出版社, 2016
ISBN 978-7-03-048791-9

I. ①新… II. ①陈… III. ①基因组—序列—测试—研究
IV. ①Q343.1

中国版本图书馆 CIP 数据核字(2016)第 131838 号

责任编辑: 罗 静 田明霞 / 责任校对: ×××
责任印制: 肖 兴 / 封面设计: 北京铭轩堂广告设计有限公司

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>

中国科学院印刷 印刷

科学出版社发行 各地新华书店经销

*

2016 年 6 月第 一 版 开本: 720×1000 1/16

2016 年 6 月第一次印刷 印张: 22 1/2

字数: 450 000

定价: 128.00 元

(如有印装质量问题, 我社负责调换)

《新一代基因组测序技术》编者名单

主编 陈浩峰

编者 (按姓氏拼音排序)

曹英豪 陈旭 陈浩峰 高 强 韩 瑶
雷 猛 李 妍 李 珍 孟 菲 齐 洺
王 静 王剑峰 杨 鑫 于 莹 张 兵

前 言

自从 2005 年新一代基因组测序技术诞生以来，短短十余年间，其发展日新月异。今天，新一代测序技术已经被广泛应用到生命科学研究的各个方面，可以说，该技术的迅速发展标志着生命科学组学研究时代的到来。目前，学习、掌握和应用新一代测序技术已经成为广大生命科学工作者的迫切需求，对于国内开展新一代测序研究的实验室来说，手头有一本详细介绍新一代测序各方面技术的中文书籍非常必要。鉴于此，我们编写了本书，供生命科学有关专业的高等院校师生、科研院所的研究人员及其他生命科学从业人员参考。

本书以新一代基因组测序技术 Illumina 平台和 PacBio RS 平台为代表，详细阐述了从实验样品处理到数据分析的整个过程，重点介绍了测序过程中的文库构建与质检、测序仪器操作和测序数据的初步处理及分析。全书共分为五章：第一章为绪论，概述了测序技术发展简史，以及各个主要的新一代测序平台的技术特点；第二章到第四章详细介绍了目前新一代测序技术的主流平台——Illumina 测序技术的各个方面，第二章介绍了 Illumina 建库技术，第三章介绍了 Illumina 测序操作，第四章介绍了 Illumina 测序数据的初步分析；第五章则是针对单分子测序技术 PacBio RS 的建库、测序和数据分析过程的介绍。本书的编写人员都是工作在科研第一线，有着多年第一代和新一代测序实践经验和数据分析经验的中青年科研工作者，希望我们从科研工作中得来的经验和体会对广大读者学习新一代测序技术起到借鉴和参考作用。

我们编写本书的目的是，希望具有一定分子生物学与遗传学背景的读者，通过阅读本书了解新一代测序技术的概念及其发展历史，以及各种新一代测序技术平台所独有的特点；并且能够参照书中所述的建库流程指导，在自己的实验室成功完成测序文库的构建和质量检测。此外，我们还希望在有条件进行新一代测序工作的实验室，读者可以参照本书了解 Illumina 和 PacBio RS 测序仪的基本工作流程和测序数据的初步分析过程。这样一来，读者在进行此类科研项目时，就可以做到对研究规划心中有数，在实验设计上有的放矢；而不是把测序建库、测序操作和数据分析过程完全交给测序服务公司，把这一部分工作作为“黑箱”来对待。对于其他各种生物学、医学等实验室的研究人员和工作人员来说，即使没有机会上机操作，至少也可以通过阅读本书获得一些新一代测序的知识，了解各种

新一代测序方法的优点与局限性，并且有能力进行测序项目实验的追踪纠错。

感谢中国科学院遗传与发育生物学研究所和北京基因组研究所提供的良好科研工作条件；特别感谢遗传发育所基因组生物学研究中心与植物基因组学国家重点实验室对编写工作的支持；感谢 Illumina 中国公司、New England Biolabs(NEB)中国公司和凯杰(Qiagen)中国公司的大力赞助；感谢本书责任编辑罗静女士的出色工作；最后还要感谢我海内外的亲朋好友与同行，他们在本书的写作过程中提出了宝贵的修改意见。没有上述机构和人员的共同努力，本书不可能很快面世。

限于编者的知识水平，本书所列举的测序研究案例主要集中于植物学、农学与育种学等方面。目前，新一代测序在生命科学的各个领域，尤其是医学研究和疾病的临床诊断中的应用越来越广泛，由于我们缺少这方面的研究和诊断实例，本书对测序诊断方面的应用涉及不多，这是我个人感到遗憾的地方，希望将来有机会增添这部分的内容。

衷心希望广大读者对本书内容提出宝贵意见，以便于我们将来修订和提高。

陈浩峰

2016年1月5日于北京

目 录

第一章 测序技术发展概述.....	1
第一节 第一代基因测序方法简介.....	1
第二节 新一代测序技术概述.....	4
参考文献.....	21
第二章 Illumina 测序建库.....	25
第一节 DNA 测序建库.....	25
第二节 转录组测序 (RNA-seq) 建库.....	80
第三节 小 RNA 测序建库.....	163
第四节 简化基因组测序建库.....	184
第五节 目标序列捕获测序建库.....	204
第六节 单细胞测序建库.....	227
参考文献.....	244
第三章 Illumina 仪器操作.....	246
第一节 簇生成操作流程.....	246
第二节 测序仪 HiSeq 操作流程.....	254
第三节 测序仪 MiSeq 操作流程.....	268
第四节 测序仪 NextSeq500 操作流程.....	276
参考文献.....	283
第四章 Illumina 测序数据分析方法简介.....	284
第一节 下机数据的初步处理.....	284
第二节 DNA 测序数据分析简介.....	291
第三节 转录组测序标准信息分析.....	304
第四节 建造中等高性能计算机群系统.....	316
参考文献.....	323
第五章 PacBio RS 测序技术.....	326
第一节 PacBio RS 测序原理.....	326

第二节	PacBio RS 测序 DNA 样品准备及文库构建流程	328
第三节	SMRT Portal 二级分析软件的安装	333
第四节	SMRT Portal 数据分析流程	334
第五节	PacBio RS 测序应用简介	348
第六节	PacBio 测序案例	349
	参考文献	350
	常用英文简写列表	351