

遗传统计学

基于连锁和关联分析的基因定位

Statistical Genetics: Gene Mapping through Linkage and Association

主编 Benjamin M Neale
Manuel AR Ferreira
Sarah E Medland
Danielle Posthuma

主译 尹先勇

遗传统计学

基于连锁和关联分析的基因定位

Statistical Genetics

Gene Mapping through Linkage and Association

主 编 Benjamin M. Neale Manuel A.R. Ferreira
Sarah E. Medland Danielle Posthuma

主 译 尹先勇

译 者 (以姓氏笔画为序)

尹先勇 (安徽医科大学)

左先波 (安徽医科大学皮肤病研究所)

左晓宇 (中山大学肿瘤防治中心)

任韵清 (浙江大学医学院附属第二医院)

刘稳升 (中山大学附属肿瘤医院)

孙勇虎 (山东省皮肤病性病防治研究所)

李志强 (上海交通大学 Bio-X 研究院)

李妙竹 (杜克大学)

陈 刚 (安徽医科大学皮肤病研究所)

陈 超 (中南大学)

陈小伟 (中国科学院生物物理研究所)

林 媛 (北卡罗莱纳大学教堂山分校)

学术秘书

沈长兵 (安徽医科大学皮肤病研究所)

易洪刚 (南京医科大学公共卫生学院)

周伏圣 (安徽医科大学皮肤病研究所)

郑厚峰 (杭州师范大学衰老研究所)

郑晓冬 (安徽医科大学皮肤病研究所)

姚 莹 (奥斯陆大学)

高金平 (安徽医科大学皮肤病研究所)

郭 辉 (中南大学生命科学学院)

郭允苗 (中山大学肿瘤防治中心)

郭梦彪 (香港大学李嘉诚医学院)

曹鹏博 (军事医学科学院放射与辐射医学研究所)

盛宇俊 (安徽医科大学皮肤病研究所)

戴俊程 (南京医科大学公共卫生学院)

杨 超 (安徽医科大学皮肤病研究所)

人民卫生出版社

Statistical Genetics: gene mapping through linkage and association / by Benjamin M. Neale, et al

ISBN: 9780415410403

Copyright © 2008 by Taylor & Francis Group.

Authorized translation from English language edition published by CRC Press, part of Taylor & Francis Group LLC; All rights reserved. 本书原版由 Taylor & Francis 出版集团旗下 CRC 出版公司出版, 并经其授权翻译出版。版权所有, 侵权必究。

People's Medical Publishing House is authorized to publish and distribute exclusively the Chinese (Simplified Characters) language edition. This edition is authorized for sale throughout Mainland of China. No part of the publication may be reproduced or distributed by any means, or stored in a database or retrieval system, without the prior written permission of the publisher. 本书中文简体翻译版授权由人民卫生出版社独家出版并限在中国大陆地区销售。未经出版者书面许可, 不得以任何方式复制或发行本书的任何部分。

Copies of this book sold without a Taylor & Francis sticker on the cover are unauthorized and illegal. 本书封面贴有 Taylor & Francis 公司防伪标签, 无标签者不得销售。

图字: 01-2016-2222

图书在版编目(CIP)数据

遗传统计学: 基于连锁和关联分析的基因定位 / (美) 本杰明·M. 尼尔(Benjamin M. Neale)主编; 尹先勇主译. —北京: 人民卫生出版社, 2017

ISBN 978-7-117-24231-8

I. ①遗… II. ①本… ②尹… III. ①遗传学—应用统计学
IV. ①Q3-32

中国版本图书馆 CIP 数据核字(2017)第 042968 号

人卫智网	www.ipmph.com	医学教育、学术、考试、健康, 购书智慧智能综合服务平台
人卫官网	www.pmph.com	人卫官方资讯发布平台

版权所有, 侵权必究!

遗传统计学: 基于连锁和关联分析的基因定位

主 译: 尹先勇

出版发行: 人民卫生出版社(中继线 010-59780011)

地 址: 北京市朝阳区潘家园南里 19 号

邮 编: 100021

E - mail: pmph@pmph.com

购书热线: 010-59787592 010-59787584 010-65264830

印 刷: 中国农业出版社印刷厂

经 销: 新华书店

开 本: 787 × 1092 1/16 印张: 27 插页: 2

字 数: 620 千字

版 次: 2017 年 5 月第 1 版 2017 年 5 月第 1 版第 1 次印刷

标准书号: ISBN 978-7-117-24231-8/R · 24232

定 价: 90.00 元

打击盗版举报电话: 010-59787491 E-mail: WQ@pmph.com

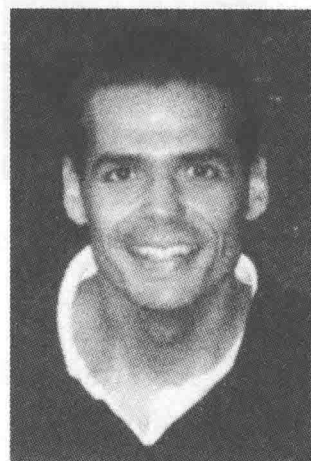
(凡属印装质量问题请与本社市场营销中心联系退换)

主编简介

Benjamin M. Neale 于 2006 年毕业于弗吉尼亚联邦大学 (Virginia Commonwealth University), 获得精神疾病遗传学学士学位。在本科学习期间, 他与 Patrick Sullivan 博士、Cynthia Bulik 博士、Kenneth Kendler 博士等一起从事结构方程建模和精神疾病的连锁分析研究工作。2004 年, 他搬到精神疾病研究所 (IOP) 与 Pak Sham 博士共同开发关联分析的方法。在精神疾病研究所, 他开始与 Philip Asherson 博士进行多动症的遗传学研究, 主要是关联分析研究。目前, 他与 Mark Daly 博士正在麻省总医院 (Massachusetts General Hospital) 的人类遗传研究中心和麻省理工学院 - 哈佛大学 Broad 研究所做访问研究, 主要是开发全基因组关联研究的分析方法。



Manuel A. R. Ferreira 于 2000 年毕业于葡萄牙里斯本大学 (University of Lisbon, Portugal) 的生物科学专业。在本科学习期间, 特别是作为一名访问学生在英国牛津大学学习期间, 他主要从事行为生态学方面的研究。2001 年, 他移居到澳大利亚的昆士兰大学 (University of Queensland) 的昆士兰医学研究所和医学院跟随 Nick Martin 教授和 David Duffy 博士攻读博士学位。他的论文是通过连锁分析鉴定哮喘的遗传风险因素。他曾获得多项学术奖励, 包括 2004 年度的欧洲人类遗传学学会 Lodewijk-Sandkuijl 奖。2006 年, 他获得了澳大利亚国家健康和医学研究理事会 Sidney Sax 奖学金, 加入哈佛大学医学院人类遗传研究中心 Shaun Purcell 博士领导的研究组。目前, 他在那里参与全基因组关联研究分析方法的开发。





Sarah E. Medland 于 2000 年毕业于澳大利亚昆士兰大学 (University of Queensland), 获得神经心理学学士学位 (荣誉)。她本科期间主要针对大脑语言中枢的偏侧性进行研究。她在昆士兰医学研究所的遗传流行病部攻读博士学位时, 指导老师为 David Duffy 博士、Margie Wright 博士和 Gina Geffen 教授。在博士学习期间, 她与 Nick Martin 教授开展合作。她的博士课题主要针对行为的偏侧性进行遗传流行病学研究。她还开展了多个其他领域的研究, 包括方法学的开发、药物滥用、政治和社会态度、婴儿生长发育和肥胖。2006 年, 她获得了澳大利亚国家健康和医学研究理事会 Sidney Sax 奖学金, 加入弗吉尼亚州精神病学和行为遗传学研究所 Mike Neale 博士领导的研究组。



Danielle Posthuma 目前是荷兰阿姆斯特丹自由大学 (VU University, Amsterdam) 生物心理学系的副教授。她于 1996 年获得了三合一的硕士学位 (荣誉), 并于 2002 年获得了行为遗传和优生学博士学位, 还获得了一些国家和国际的荣誉和奖励, 比如因其早年在行为遗传学关联研究中的突出贡献而获得富勒和斯科特奖。她现在是认知基因鉴定项目和遗传与环境相互作用 (GxE) 遗传统计学方法开发项目的主要负责人, 也是全基因组分析专属的计算机集群的科学总监, 同时, 还是一年一度的欧洲“基因鉴定和遗传流行病方法”论坛的共同策划者。

编者

Gonçalo R. Abecasis

Center for Statistical Genetics, Dept. of Biostatistics, University of Michigan, USA

David Bender

Broad Institute of MIT and Harvard University, USA
Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, USA

Stéphanie M. van den Berg

Department of Biological Psychology, Vrije Universiteit Amsterdam, The Netherlands

Dorret I. Boomsma

Department of Biological Psychology, Vrije Universiteit Amsterdam, The Netherlands

Stacey S. Cherny

Department of Psychiatry and Genome Research Centre, The University of Hong Kong, China

Danielle M. Dick

Departments of Psychiatry and Psychology, Washington University in St. Louis, USA

Conor V. Dolan

Department of Psychology, University of Amsterdam, The Netherlands

David M. Evans

Wellcome Trust Centre for Human Genetics, University of Oxford, UK

Jesen Fageress

Broad Institute of MIT and Harvard University, USA
Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, USA

Manuel A.R. Ferreira

Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, USA
Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Australia

Jouke Jan Hottenga

Department of Biological Psychology, Vrije Universiteit
Amsterdam, The Netherlands

Jo Knight

Social, Genetic, and Developmental Psychiatry Centre, Institute of
Psychiatry, King's College, UK

Johnny S.H. Kwan

Genome Research Centre, The University of Hong Kong, China

Hermine H. Maes

Virginia Institute of Psychiatric and Behavioral Genetics,
Department Human Genetics and Massey Cancer Center, Virginia
Commonwealth University, USA

Julian B. Maller

Broad Institute of MIT and Harvard University, USA
Center for Human Genetic Research, Massachusetts General
Hospital, Harvard Medical School, USA

Nicholas G. Martin

Genetic Epidemiology Laboratory, Queensland Institute of
Medical Research, Australia

Sarah E. Medland

Virginia Institute of Psychiatric and Behavioral Genetics, Virginia
Commonwealth University, USA
Genetic Epidemiology Laboratory, Queensland Institute of
Medical Research, Australia

Andrew P. Morris

Wellcome Trust Centre for Human Genetics, University of Oxford,
UK

Benjamin M. Neale

Social, Genetic, and Developmental Psychiatry Centre, Institute of
Psychiatry, King's College, UK
Broad Institute of MIT and Harvard University, USA
Center for Human Genetic Research, Massachusetts General
Hospital, Harvard Medical School, USA

Michael C. Neale

Virginia Institute of Psychiatric and Behavioral Genetics,
Departments of Human Genetics, Psychiatry and Psychology,
Virginia
Commonwealth University, USA
Department of Biological Psychology, Vrije Universiteit
Amsterdam, The Netherlands

Dale R. Nyholt

Genetic Epidemiology Laboratory, Queensland Institute of
Medical Research, Australia

Danielle Posthuma

Department of Biological Psychology, Vrije Universiteit
Amsterdam, The Netherlands

Shaun Purcell

Psychiatric and Neurodevelopmental Genetics Unit,
Massachusetts General Hospital, Harvard Medical School, USA
Broad Institute of MIT and Harvard University, USA

Frühling V. Rijsdijk

Social, Genetic, and Developmental Psychiatry Centre, Institute of
Psychiatry, King's College, UK

Pak C. Sham

Department of Psychiatry and Genome Research Centre, The
University of Hong Kong, China

Sophie van der Sluis

Department of Biological Psychology, Vrije Universiteit
Amsterdam, The Netherlands

Patrick F. Sullivan

Department of Genetics, University of North Carolina, USA

Peter M. Visscher

Genetic Epidemiology Laboratory, Queensland Institute of
Medical Research, Australia

Naomi R. Wray

Genetic Epidemiology Laboratory, Queensland Institute of
Medical Research, Australia

序 一

人类已经迈入精准医学时代。遗传和基因组学是实现精准医学的基础。近年来,高通量基因测序技术突飞猛进,个体基因组的获取不再是个别人的专利。基因组测序的广泛应用,使得遗传大数据如潮水般涌现。如何在海量的遗传和基因组学数据中,抽丝剥茧、去伪存真、乃至格物致知,都离不开一门工具——遗传统计学。遗传统计学是一门新兴交叉学科,是遗传学的一个重要分支。遗传统计学之于遗传学,犹如猎枪之于猎人。理解、掌握并应用遗传统计学的理论、方法,已成为每一位遗传学科技工作者所必须掌握的一项基本技能。

遗传统计学科的发展,人才是关键;而人才的培养,离不开优秀的书籍。*Statistical Genetics: Gene Mapping through Linkage and Association* 由活跃在当代遗传统计学领域的一批优秀科学家编写而成。编委中既有功成名就的科学大师,也有小荷才露尖尖角的后起之秀。该书紧紧抓住了遗传统计学在人类性状研究中的“牛鼻子”——基因定位,系统介绍了遗传统计学的理论知识、数理基础,以及基因定位的连锁和关联研究常用方法和实践,理论与实践交融,是一本从事基因定位研究难得一见并值得阅读的参考书。

《遗传统计学:基于连锁和关联分析的基因定位》译著的出版,汇聚了海内外26位华人青年才俊的集体智慧。他们是遗传统计学在中国的践行者,亲身经历和参与了我国疾病遗传学和基因组学研究的伟大实践,对遗传统计学具有科学而客观的认识和思考。现在,他们又勇敢地担起传播遗传统计科学的社会责任,充分展现了我国青年学者扎实的科学素养和良好的精神风貌。

我国遗传统计学起步较晚,相信此译著的出版,定将进一步推动我国遗传统计学的发展。



中国科学院院士

上海交通大学 Bio-X 研究院教授

2016年6月10日

序 二

遗传统计学集遗传学、数学、计算机科学、信息科学和统计学等学科知识与理论于一身，是一门新兴的前沿学科。在遗传学研究的历史长河中，尤其是在近年来遗传学海量数据的挖掘中，遗传统计学已经展现了其无限的魅力。

由我组建和领导的安徽医科大学皮肤病遗传学研究创新团队，开展疾病遗传病因学研究超 20 年。在 20 年的科学实践中，我们运用了一系列遗传统计学方法，开展了几十种疾病，尤其是皮肤病的全基因组连锁分析和全基因组关联研究，发现了疾病 100 多个致病或易感基因，推动了疾病病因学研究。因此，我常常说，我们 20 年的研究实践，就是一部遗传统计学的发展史和应用史。在当前科学大数据的时代背景下，遗传统计学比以往任何时候都显得尤为关键；其科学和广泛地应用，更将加速度推进遗传学研究前进的步伐。

实践出真知。20 年的实践，成就了我团队中一批遗传统计学专业人才。他们通过学习、摸索，逐渐理解并掌握了基因定位研究的常用统计分析方法，并在科学研究实践中成功运用。他们再次升华，由实践回到理论，主导并参与《遗传统计学：基于连锁和关联分析的基因定位》译著的翻译工作，集中展示了他们扎实的理论功底和丰富的遗传统计学实践感悟。

窥一斑而知全豹。安徽医科大学皮肤病遗传学研究创新团队的成长和经历，是我国疾病遗传学研究和人才建设的一个缩影。20 年的实践，我们的感悟是用好遗传统计学！

此译著的出版，凝聚了 26 位青年才干的心血。他们经过了一系列遗传学研究的洗礼和锤炼，基础扎实；他们年轻，朝气蓬勃，斗志昂扬，充满希望！

愿他们在遗传统计学研究路上越飞越高！



安徽医科大学教授
中华医学会皮肤性病学分会名誉主任委员
国际皮肤科学会联盟常务理事
2017 年 3 月 10 日

译者前言

2014年,我在美国参加学术会议期间第一次见到 *Statistical Genetics: Gene Mapping through Linkage and Association*, 就被书中的内容深深吸引, 猛然间萌生出翻译此著作的念头。

回国后,我尝试着与该著作的主编 Benjamin Neale 博士取得联系, 询问出版中文译著的可能。出乎我的意料, Neale 博士及时回应并给予了极大支持, 细致地指导我如何获取版权。Neale 博士的肯定和鼓励, 更加坚定了我翻译此著作的想法。在人民卫生出版社的支持下, 译著出版计划于2015年11月正式获得立项。

该书以基因定位研究为主线, 内容覆盖了遗传学、统计学和数学的基础知识, 并分别从连锁分析和关联分析的角度, 介绍了基因定位研究的常用方法、应用范畴及其数理基础; 书中引入了大量研究实例, 语言通俗易懂, 内容新颖生动, 适合于有志于或正在从事遗传学研究的研究生及科技工作者阅读。

此译著的26位青年译者全部工作在遗传统计学研究的一线, 在各自的领域都有着突出的成绩; 长期参与遗传统计学研究的实践, 使得他们对遗传统计学有着客观而专业的认识。其中部分人我虽素未谋面, 但在译著翻译出版过程中, 他们专业的素养、专注的态度和执著的精神都给我留下了深刻的印象。

感谢上海交通大学贺林院士、安徽医科大学张学军教授的鼓励和指导。译著翻译出版过程中, 得到了安徽医科大学叶冬青教授、上海交通大学师咏勇教授、中山大学贝锦新教授、中南大学夏昆教授、军事医学科学院周钢桥教授、南京医科大学胡志斌教授和香港大学杨万岭教授等的帮助和指导, 在此一并谢过。

感谢人民卫生出版社给予的帮助和支持。感谢我的爱人一直默默地支持我的工作。

由于译者水平有限, 书中某些翻译难免存在不尽如人意之处, 真诚地希望读者提出宝贵意见和建议。

尹先勇

2017年4月

原著序

这本书是双生子和家系研究方法研讨会的成果。该研讨会自 1987 年在比利时的鲁汶市 (Leuven) 举行第一次会议以来 (由 Robert Derom 主持), 已经举行了 20 余次。此研讨会的发起源于一个国际性协作网的成员对相同研究领域的热情, 这些研究者主要致力于结构方程建模方法的开发及其在行为遗传学和遗传流行病学新兴领域的应用。在当时, 虽然这些方法的检验效能正在变得广为人知, 但是很少有地方能够正式讲授这些方法, 因此, 很多国家不同学科的学生对学习这种方法的场所有极大需求。把这种新兴学科的领衔者聚集在舒适的环境中, 以期为更广泛的听众传授这种神秘的艺术, 这似乎是一个不错的主意。鲁汶 (Leuven)、博尔德 (Boulder)、科罗拉多 (Colorado)、赫尔辛基 (Helsinki) 和最近的荷兰海岸的伊哥芒特 (Egmond) 都提供了非正式但适合密集培训的理想场所, 研讨会由此而变得出名。根据最新统计, 超过 800 个成员已从 1 周的集训中受益, 研讨会的内容包括结构方程建模、矩阵代数、心理统计学和生物统计遗传学等。

无论是传授知识还是促进研究的应用和开发新方法, 研讨会无疑取得了成功。由 Nick Martin、Dorret Boomsma 和 Michael Neale (1989 年 1 月) 所编辑的《行为遗传学》一期特刊, 出版了来自于第一次研讨会的十篇论文, 设定了遗传学中的结构方程建模的通用框架, 以及处理诸如如何对年龄和性别影响进行建模的实际问题, 并创新性地用于双生子研究, 同时, 首次利用双生子进行连锁分析。

1991 年的研讨会, 由鲁汶 (Leuven) 的 Hermine Maes (曾是第一期研讨会的学生) 组织, 研讨会获得了 NATO 的资助。由于要求出版一个论文集, 这激发了本书的编写, 参加人员耗费了数周的周末时间进行了粗加工, 然后再充实和完善, 而成为著名的“Neale 和 Cardon”方法 [M.C. Neale and L.R. Cardon (1992) *Methodology for Genetic Studies of Twins and Families*, NATO ASI Series D: Behavioral and Social. Kluwer Academic, Dordrecht, The Netherlands], 已经被引用 1000 多次。值得一提的是, Neale 和 Cardon 影印版本在 eBay 的售价为 \$200, 精装版在亚马逊的售价超过 \$1000, 市场证明了其价值所在。

Neale 和 Cardon 是围绕 LISREL 软件编程而写成的, 课程前 5 年左右的时间依赖该软件。然而, 1990 年, Michael Neale 开发了他自己的程序, 即 Mx, 概念上基于 LISREL, 但便于书写, 更加灵活, 从而适合遗传学应用, 现已变为被行为遗传学家和遗传流行病学家使用的标准工具。Mike 和他的妻子 Hermine Maes 基本完成 Neale 和 Cardon 的修订 (已经被称为 Neale 和 Maes), 使其适于 Mx 语言。

1990 年, John Hewitt [2001 年成为博尔德 (Boulder) 科罗拉多大学 (University of Colorado) 行为遗传学研究所主任] 获得了 NIMH 培训基金来资助博尔德 (Boulder) 研讨会。从此以后, 研讨会已经在博尔德 (Boulder) 举行了 15 次, 通常在三月份。从 1999 年开始, 入门 (偶数年) 和进阶 (奇数年) 研讨会交替进行, 这

两种类型的研讨会通常会被超额预定, 平均每期研讨会会有 55 位学生, 一些学生实际上是高级研究员(包括 NIH 的所长), 来参与该领域的最新进展研究。研讨会的一个重要特色是理论介绍后, 通常紧跟着进行练习应用, 这对研讨会的成功至关重要, 这些特色通常可将这些研讨会与其他类似的研讨会区分开。考虑到研讨会的成功、流行和美国之外需要培训的人, 这种研讨会的模式被欧洲一系列研讨会所模仿, 如 2003 年之后的九月, 在荷兰伊哥芒特(Egmond-aan-Zee)由 Danielle Posthuma 组织举办的研讨会, 她获得了荷兰科学组织和 GenomEUtwin 项目的基金支持。

Neale 和 Cardon 包含了入门级研讨会的教学大纲、生物统计遗传学和遗传学中的结构方程建模。然而, 随着高通量微卫星和 SNP 分型技术的突破, 复杂性状基因定位的连锁和关联统计遗传学方法同时发展起来。大规模的表型和基因型数据分析方法的提高, 并通过研讨会不断地传播这些方法, 对各种复杂表型的病因学理解取得了大量成果。这些统计方法(和每年研讨会上相关的培训)的影响在快速扩张, 这很容易通过成千上万的出版物来判断。例如, 主要用于遗传研究的 Mx 软件, 自 1991 年以来已经被超过 1250 次引用, 在一系列研究领域出版物的数量每年都在持续增长。同样, 用于连锁和关联分析的 Merlin 软件(Abecasis 等)自 2002 年以来已被引用 547 次。参加研讨会的许多教师和学生都已经利用这些方法发表了论文, 有些人还教会了他们的同事。因此, 研讨会对世界各地大学研究议程的设置有着相当大的影响。曾参加研讨会的学生和青年教师有些已成为该领域的新领军人。

分子生物学、统计学、群体遗传学和结构方程建模等领域的交叉已成为高级课程的共同点, 特别令人欣喜的是, 一批非常杰出的青年科学家被吸引加入这一领域, 他们在该课程上讲授的内容已经被设定成为领域的标准。高级课程中所提供的内容是独特而有价值的。每一章都是由高级课程教师组的不同的成员所编写, 最终由四位最年轻和最敏锐的教师编辑并整理成卷。我们希望它能够找到市场并和其前身(Neale 和 Cardon)一样美好, 最终能够发挥它的作用。

Nicholas G. Martin

布里斯班昆士兰医学研究所(Queensland Institute of Medical Research)

Dorret I. Boomsma

阿姆斯特丹自由大学(Vrije Universiteit, Amsterdam)

Michael C. Neale

里士满弗吉尼亚联邦大学(Virginia Commonwealth University, Richmond)

Hermine H. Maes

里士满弗吉尼亚联邦大学(Virginia Commonwealth University, Richmond)

2007 年 3 月, 科罗拉多州博尔德市(Boulder, Colorado)

原著前言

编写本书的想法是于 2005 年 3 月在美国科罗拉多州博尔德市 (Boulder, Colorado) 参加第十八届国际双生子及家系学术研讨会时萌生的。很多当时参会的学生询问是否有相关的学习手册,他们希望通过手册了解现今关于人类复杂性状的遗传连锁分析及关联分析的知识。因此,作为教师,我们决定承担起这一任务,以促进研讨会相关知识的有效传播。

当决定接受这一挑战性的任务时,我们将目标定位于:针对人类基因定位感兴趣的研究者,编写一本在基因连锁和关联分析领域起指引作用的手册。我们邀请了相关领域的领头研究人员,其多数也是研讨会的成员,让他们来编写理论部分和(或)实践操作部分的章节。我们希望理论部分可以涵盖基因定位的基础理论知识,而实践操作部分讲述基因连锁及关联分析是如何实施的、实践中可能遇到的问题,以及如何解决这些问题。示例的脚本和教程中使用的数据可以通过本书的主页 (<http://www.genemapping.org>) 获取。

此书大多数的章节是由一些编写常用遗传学软件的研究人员负责执笔的。因此,此书仅涉及许多优秀软件中的一部分。此书获得的肯定应该给予所有负责编写研究者。非常感谢所有为此付出的人员,也非常荣幸与你们共事。

Benjamin M. Neale Manuel A.R. Ferreira

Sarah E. Medland Danielle Posthuma

2007 年 4 月 3 日, 荷兰阿姆斯特丹 / 美国波士顿 / 美国弗吉尼亚州里士满
(Amsterdam/Boston/Richmond)

缩 略 词

A	additive genetic variance component (累加性遗传方差组分)
ALS	amyotrophic lateral sclerosis (肌萎缩侧索硬化症)
C	shared environmental variance component (共享环境方差组分)
CDCV	common disease common variant (常见疾病常见变异)
CEPH	Centre d'Etude du Polymorphisme Humain collection (d'Etude du 人类多态性收集中心)
CNV	copy number variation (拷贝数变异)
DNA	deoxyribonucleic acid (脱氧核糖核酸)
E	nonshared environmental variance component (非共享环境方差组分)
EM	expectation maximization (最大期望值法)
FDR	false discovery rate (错误发现率)
GAIN	Genetic Association Information Network (遗传关联信息网络)
GPC	genetics power calculator (遗传把握度计算器)
GWAS	genome-wide association study (全基因组关联分析)
GxE	gene-environment interaction (基因环境交互作用)
Hap Map	haplotype map (单倍型图谱)
HE	Haseman-Elston (Haseman-Elston 回归)
HLOD	heterogeneity LOD (异质性 LOD 值)
HRRT	haplotype relative risk test (单倍型相对危险度检验)
HWE	Hardy-Weinberg equilibrium (哈迪 - 温伯格平衡)
IBD	identity-by-descent (血缘同源)
IBS	identity-by-state (状态同源)
LD	linkage disequilibrium (连锁不平衡)
LOD	log of odds (比值比对数值)
LRT	likelihood-ratio test (似然比检验)
LSE	least squares estimation (最小二乘估计)
MAF	minor allele frequency (最小等位基因频率)
MAR	missing at random (随机缺失)
MCMC	Markov-Chain Monte-Carlo (马尔可夫链蒙特卡罗算法)
MDS	multidimensional scaling (多维标度)
ML	maximum likelihood (最大似然法)
MLE	maximum-likelihood estimation (最大似然估计)
MRV	multiple rare variants (多发罕见变异)
NCP	noncentrality parameter (非中心参数)

OLS	ordinary least squares (普通最小二乘法)
PAR	pseudoautosomal region (拟常染色体区域)
PCA	principal components analysis (主成分分析)
PCR	polymerase chain reaction (聚合酶链反应)
POE	parent-of-origin effect (亲源效应)
Q	variance component attributed to QTL effect (归因于数量性状基因座效应的方差组分)
QC	quality control (质量控制)
QTD	quantitative transmission disequilibrium test (数量传递不平衡检验)
QTL	quantitative trait locus (数量性状基因座)
rGE	gene-environment correlation (基因 - 环境相关性)
RNA	ribonucleic acid (核糖核酸)
SEM	structural equation modeling (结构方程模型)
SNP	single nucleotide polymorphism (单核苷酸多态性)
TDT	transmission disequilibrium test (传递不平衡检验)
VC	variance components (方差组分)
WGAS	whole genome association study (全基因组关联研究)
WLS	weighted least squares (加权最小二乘法)
WTCCC	Welcome Trust Case Control Consortium (威康信托基金会病例对照协会)

统计学符号

一般统计学符号

术语	符号
Mean (平均值)	μ
Variance (方差)	σ^2
Degrees of freedom (自由度)	df
Chi-square (卡方值)	χ^2
Summation (求和)	Σ
Multiplicative combination (乘法组合)	Π
Normal distribution (正态分布)	$N(\mu, \sigma^2)$
Normal function (正态函数)	$\varphi(\chi)$
Probability of A (A 的概率)	$P(A)$
Probability of A given B (B 条件下 A 的条件概率)	$P(A B)$
P (type-I error) (第一类错误概率)	α
P (type-II error) (第二类错误概率)	β
Correlation coefficient (相关系数)	ρ
Squared correlation coefficient (平方相关系数)	ρ^2
Variance covariance matrix (协方差矩阵)	Σ
Likelihood of x (x 的似然值)	$L(x)$
Total sample size (总样本量)	N
Regression coefficients (回归系数)	β

遗传学符号

术语	符号
Additive genetic component of variance (累加性遗传方差组分)	σ_A^2
QTL component of variance (QTL 方差组分)	σ_D^2
Additive QTL component of variance (累加性 QTL 方差组分)	σ_{Qa}^2
Dominance QTL component of variance (显性 QTL 方差组分)	σ_{Qd}^2
Common environment component of variance (常见环境方差组分)	σ_c^2
Dominance component of variance (显性方差组分)	σ_D^2
Specific environment/error (特定的环境 / 错误)	σ_E^2
Total phenotypic variance (总表型变异)	σ^2 或 σ_p^2
Recombination fraction (重组率)	θ