



Supported by the National Basic Research Program of China (973 Program)

国家重点基础研究发展计划(973)项目资助

蚕的基因组

The Genome Of Silkworm

主编 夏庆友 向仲怀



科学出版社

国家科学技术学术著作出版基金资助出版

国家重点基础研究发展计划(973)项目资助

The Genome Of Silkworm

蚕的基因组

主编 夏庆友 向仲怀



科学出版社

北京

内 容 简 介

家蚕是完成了基因组框架图、精细图和遗传变异图的少数物种之一。本书是作者团队历经十年研究所完成的家蚕基因组计划的总结,系统阐述了家蚕基因组的结构与进化;家蚕基因表达及蛋白质组;家蚕重要蛋白质编码基因家族及小RNA;家蚕变态发育、性别决定、免疫防御、丝蛋白生物合成,以及食性与化学感受的分子基础;家蚕基因功能研究平台技术;家蚕基因组数据库等研究进展。

本书既可作为蚕学及昆虫基因组和功能基因组研究的参考书,也可作为特种动物学、昆虫学和生物学等学科研究生的教材。

图书在版编目(CIP)数据

蚕的基因组 / 夏庆友, 向仲怀主编. —北京: 科学出版社, 2013. 3

ISBN 978-7-03-036408-1

I. ①蚕… II. ①夏…②向… III. ①家蚕—基因组—研究 IV. ①S881.2

中国版本图书馆CIP数据核字(2012)第319156号

责任编辑: 夏 梁 童安齐 刘 晶 / 责任校对: 刘小梅

责任印制: 钱玉芬 / 封面设计: 王 煤

版式设计: 梅木子

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>

中国科学院印刷厂 印刷

科学出版社发行

各地新华书店经销

*

2013年3月第 一 版 开本: 889 × 1194 1/16

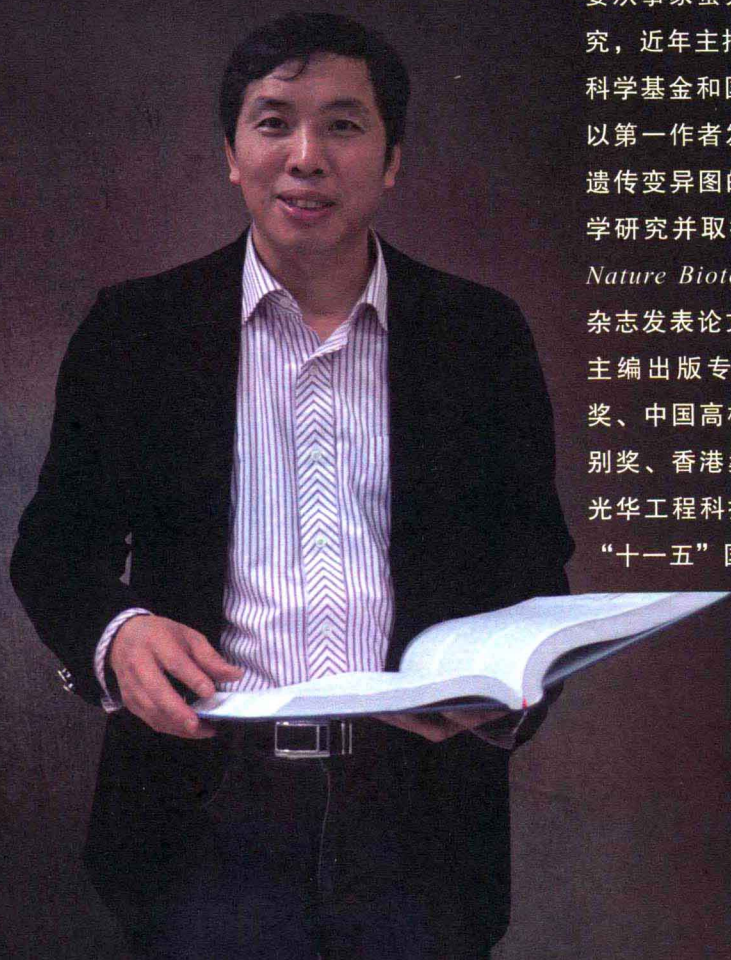
2013年3月第一次印刷 印张: 28 1/4 插页: 2

字数: 723 000

定价: 298.00元

(如有印装质量问题, 我社负责调换)

夏庆友，教授，博士生导师。第十一届全国政协委员。教育部“长江学者”特聘教授，“新世纪百千万人才工程”国家级人选，教育部“新世纪优秀人才培养计划”人选，国家“973”家蚕项目首席科学家和中国烟草基因组计划首席科学家。1985年获西南农业大学学士学位，1996获西南农业大学博士学位，1996~2000年留学日本九州大学并任日本文部省特别研究员。现任家蚕基因组生物学国家重点实验室主任，西南大学生物技术学院院长。主要从事家蚕分子生物学、基因组及功能基因组学研究，近年主持完成了“973”、“863”、国家自然科学基金和国际科技合作项目等10余项重大课题，以第一作者发表了家蚕基因组框架图、精细图和蚕遗传变异图的绘制成果，主持开展家蚕功能基因组学研究并取得了重要进展。累计在包括*Science*、*Nature Biotechnology*、*PNAS*等在内的国内外学术杂志发表论文200余篇（其中SCI论文100余篇），主编出版专著2部。先后获重庆市自然科学一等奖、中国高校十大科技进展、日本蚕丝科学进步特别奖、香港桑麻纺织科技大奖、中国青年科技奖、光华工程科技奖青年奖、全国“五一”劳动奖章、“十一五”国家科技计划执行突出贡献奖等奖励。





向仲怀，教授，博士生导师，中国工程院院士。第四、第五、第六届国务院学位委员会委员，中国蚕学会名誉理事长，中国丝绸协会副会长，原西南农业大学校长。1958年西南农学院毕业，1982年留学日本信州大学并获工学博士学位。现任西南大学蚕学与系统生物学研究所所长，家蚕基因组生物学国家重点实验室学术委员会主任。长期从事蚕桑教育与科学研究，先后发表研究论文300余篇（SCI论文50余篇），主编出版学术专著6部。先后获国家自然科学基金四等奖、四川省科技进步一等奖、重庆市自然科学一等奖、重庆市科技进步一等奖及全国农业科技先进工作者、中华农业科教奖、日本蚕丝学会特别奖、香港何梁何利基金科技进步奖、香港桑麻基金会纺织科技大奖、全国茧丝绸行业终身成就奖等奖励。



王根洪, 1980年生, 理学博士, 副教授, 现工作单位为西南大学蚕学与系统生物学研究所, 主要从事动植物RNAi及转基因研究

王菲, 1978年生, 理学博士, 副教授, 现工作单位为西南大学蚕学与系统生物学研究所, 主要从事家蚕免疫应答机制研究

刘春, 1975年生, 农学博士, 副教授, 现工作单位为西南大学蚕学与系统生物学研究所, 主要从事家蚕丝蛋白合成机制及重要突变基因定位克隆研究

刘仕平, 1970年生, 理学博士, 副教授, 现工作单位为西南大学蚕学与系统生物学研究所, 主要从事miRNA功能研究

李东, 1985年生, 2011年获西南大学理学博士学位, 现为美国费城儿童医院博士后, 主要从事动植物基因组学研究

何宁佳, 1971年生, 农学博士, 教授, 博士生导师, 现工作单位为西南大学蚕学与系统生物学研究所, 主要从事家蚕表皮蛋白基因功能和桑树基因组研究

柴春利, 1979年生, 农学博士, 副教授, 现工作单位为西南大学生物技术学院, 主要从事家蚕Hox基因功能研究

徐汉福, 1978年生, 理学博士, 副教授, 现工作单位为西南大学蚕学与系统生物学研究所, 主要从事家蚕转基因研究

张辉洁, 1983年生, 2011年获西南大学理学博士学位, 现为西南大学蚕学与系统生物学研究所博士后, 主要从事家蚕味觉嗅觉信号转导机制研究

余泉友，1979年生，农学博士，讲师，现工作单位为重庆大学农学及生命科学研究院，主要从事家蚕抗性解毒机制研究

林 英，1972年生，农学博士，副教授，现工作单位为西南大学蚕学与系统生物学研究所，主要从事家蚕发育相关基因及其功能研究

赵 萍，1963年生，农学博士，教授，博士生导师，现工作单位为西南大学蚕学与系统生物学研究所，主要从事家蚕蛋白质组学及丝蛋白合成机制研究

段 军，1978年生，理学博士，副教授，现工作单位为西南大学蚕学与系统生物学研究所，主要从事动植物基因组学研究

查幸福，1978年生，农学博士，副教授，现工作单位为西南大学蚕学与系统生物学研究所，主要从事家蚕性别决定机制研究

侯 勇，1980年生，理学博士，副教授，现工作单位为西南大学生物技术学院，主要从事家蚕蛋白质组学研究

程廷才，1978年生，理学博士，副教授，现工作单位为西南大学蚕学与系统生物学研究所，主要从事家蚕免疫抗性机制研究

程道军，1973年生，农学博士，副教授，现工作单位为西南大学蚕学与系统生物学研究所，主要从事家蚕生长发育激素调控机理及基因组学研究



长江后浪推前浪。任何一种学科的发生发展，都是一个过程，都有连续性与继承性。研究生物体遗传与变异规律的遗传学（genetics）也是如此。

1900年，孟德尔定律被重新发现，标志着遗传学的诞生。过去的一个世纪里，遗传学获得了空前辉煌的成就。

基因（gene）一词，是1909年丹麦遗传学家约翰逊提出的一个术语，其含义不过是代表生物遗传性状的一个抽象符号。后来美国著名遗传学家摩尔根（Thomas Morgan, 1866~1945）在染色体遗传学说基础上实验证明基因在染色体上呈线性排列，占有一定空间，而且基因传递与其所在的染色体传递相连锁，这样便赋予了基因物质的内涵。在1926年出版的《基因论》中，他更明确地指出基因是“有机的化学实体”。1953年，英国物理学家克里克（Francis H. C. Crick）与美国遗传学家沃森（James D. Watson）共同发现DNA分子的双螺旋结构，以及其后的分子遗传学进展，使得人们今天更清楚地认识到基因这个“有机的化学实体”，就是DNA分子的片段，就是遗传信息的载体（它传递着支配一切生命活动的指令），就是控制生物性状的遗传单位；同时，基因又是可以人工操纵，并用于改造生物性状的元件。

排列在同一染色体上的基因群，称为连锁群（linkage group）。家蚕单倍体细胞有28条染色体（ $n=28$ ），故有28个连锁群。自1913年日本家蚕遗传学泰斗田中义麿（Yoshimoro Tanaka, 1884~1978）发现素蚕（plain）与黄血（yellow blood）基因连锁以来，家蚕连锁群分析已有百年历史。据1994年土井良宏绘制的基因连锁图，已研究过的家蚕基因400余个，发现27个连锁群、177个基因座，第28个连锁群的标志基因已确定，即tr过刺肢E-tr基因。另有报道称已经记录的形态、生理、生化等家蚕遗传性状达到440个以上（日本九州大学基因资源研究中心家蚕遗传部门，1998），确定的基因座已增至230个（鲁成等，2003）。基因连锁群研究不仅在经典遗传学研究中占有重要地位，而且对家蚕分子遗传学和家蚕基因组学研究亦很有启示，因为基因连锁图都要给出基因在染色体上的排列顺序与距离。不过，遗憾的是，家蚕染色体形状小，呈颗粒状，最大与最小之间不超过3倍；加之着丝粒不明显，呈弥散状，要按形态鉴定比较困难。所以家蚕连锁群序号与细胞学上的特定染色体暂时还对不上号。

基因组（genome）是指一种生物或个体细胞所含有的一套完整基因及其调控序列，是生物单倍体细胞内染色体的总和，20世纪90年代开始，形成了以基因组为研究对象的基因组学（genomics），这个名词是1986年由洛德力克（Tpmas Roderick）提出的。以后又出现了结构基因组学与功能基因组学，分别研究基因组的结构与功能。因为基因组本身就是基因与非基因DNA序列的总和，所以基因组除基因外，还有大量非基因DNA序列，这些非基因序列的功能，大部分还是未知的。若这些非基因序列确有一定生物功能的话，估计它们实现这些功能，要么通过形成新的基因，要么通过调控（或改变）原有基因的表达。因此，有人认为，只有弄清每

个单独基因的功能及实现功能必需的环境条件（包括基因与基因之间、基因与细胞生物大分子之间的相互关系，以及外界信息在细胞传递的分子途径等），才能阐明基因组的功能。当然，这里并无排除非孟德尔遗传规律（非DNA突变引起的遗传变异，epigenetics）存在的可能之意。反过来，也有人认为只有把基因组的活动与功能作为一个整体，并在此背景下才能真正理解每个单独基因的功能及实现其功能所需的条件。现在生命科学界将越来越多的注意力集中到基因组与基因组学上，这是很正常的，是遗传学发展历史上必然而且已经出现了的崭新亮点。

该书由西南大学夏庆友教授、向仲怀院士主编，总结了近10年来由他们组织领导的科研团队所获成果，首次对家蚕基因组作了最全面、最权威的论述。

2004年美国《科学》杂志刊登了西南大学家蚕基因组研究团队2003年完成的家蚕基因组框架图。2008年12月该团队宣布世界第一张家蚕基因组精细图诞生。2009年8月《科学》杂志又发表了西南大学研究论文“40个蚕类基因组重测序揭示家蚕的驯化事件与驯化相关基因”。这样，他们出色地完成了家蚕基因组框架图、精细图与遗传变异图。这些成果标志着家蚕基因组计划圆满结束，同时标志着家蚕功能基因组研究的开始。现在的任务是：发现基因、研究基因、利用基因。

家蚕基因组研究使整个蚕业科学进入了以分子生物学为主导的新时代。我认为，中国蚕遗传育种学、蚕种制造学、蚕体生理学、蚕体生化学、蚕体病理学的科研工作者都可在家蚕基因组全序测定已经完成、功能基因组学与蛋白质组学研究方兴未艾之际，充分利用家蚕基因组研究成果，拓展本学科的深度与广度，提高本学科的科研效率与水平，并推动整个蚕业科学的进步与蚕业生产的发展。深信随着大量重要基因的发现与应用，不仅会给家蚕育种工作、蚕体生理研究及蚕体病理研究带来革命性变化，而且将会促进家蚕模式生物学研究的进展。

果蝇是当代生命科学界国际公认的遗传学研究模型。但在20世纪20年代以前，家蚕与果蝇遗传学研究曾经几乎是并驾齐驱的。例如，摩尔根1912年发现果蝇染色体交换只限于雌蝇，而田中1913年便报道家蚕染色体交换只发生于雄蚕，并且证明雄蚕性染色体结构呈同配子型（ZZ），而雌蚕为异配子型（WZ）。21世纪伊始，西南大学家蚕基因组研究或许可以作为向果蝇急起直追的一个信号。

如果说1913年田中发现素蚕与黄血基因连锁是计划地进行家蚕经典遗传学研究开始的标志，那么2003年向仲怀院士完成的家蚕基因组全序列测定研究，则是当今家蚕分子遗传学发展史上一块重要里程碑，必将促进家蚕分子遗传学迅速发展。

在有生之年，我有幸作为该书第一位读者，学习到不少新的知识。据我所知，国内外还没有一位作者像本书第8章那样，对家蚕分子免疫学作过如此系统而深入的综述。同样，我相信看过该书第3章~第12章的蚕学界读者，一定会承认下面的事实：当你深入研究家蚕胚胎发生、幼虫成长蜕皮、蛹变态蜕皮、成虫交尾产卵、蚕体摄食行为、营养代谢、性别决定、抗病免疫、丝蛋白合成等生命活动（现象）的本质或机制时，都离不开家蚕基因组提供的信息。换句话说，家蚕一切生命活动（现象）的本质或机制，都涉及基因（全基因组有关基因及基因家族）的结构、功能。我们应该在全基因组或基因层面上探讨家蚕各种生命活动（现象）的本质和发生的原因。该书除前几章详细论述家蚕基因组计划的实施和业绩外，最后两章还专门介绍了功能基因组平台技术、家蚕基因组生物信息数据库SilkDB的建立和应用，给读者提供了更多的方便。使我特别激动的是，我国有志于家蚕分子免疫学研究的科学工作者，能及时获得《蚕的基因组》这样一本参考书，真是难得的机遇。

受益之余，深感这个世界先进水平的重大成果来之不易。已故蒋同庆教授毕生心血，艰苦经营，创立了西南农业大学家蚕遗传研究室，坚持科研与教育，在我国蚕学界独树一帜，名师出高徒，他培养的向仲怀院士，人才出众；周泽扬、鲁成、秦俭诸教授，都是当代蚕学英才；在他们的影响下，夏庆友等一批后起之秀，正茁壮成长。一项重大科研成果的获得，少不了深厚的科学储备、长远科研方向正确选择、长期科研资料积累、丰富遗传资源保存，但最重要的还是人才的储备。出成果可以带动出人才，但正确地讲，还是有了适当的人才，才能做出相应的成果。

家蚕基因组研究是一项巨大的工程，向仲怀院士独具慧眼，由此胆识与魄力，决心启动这项工程。从获得的激动人心的辉煌成果看，他所领导的这支家蚕基因组科研团队实力很强、素质较高。

中国是世界蚕业的发祥地，但近代实验科学的理论应用于蚕业科研与生产，则从欧洲开始，日本明治维新后向西欧学习，20世纪上半叶蚕业科研生产上的重大发明创造，几乎全被日本垄断，一跃而为世界蚕业生产与出口大国，而且亦是蚕业科技发展强国；70年代开始我国蚕茧产量与蚕丝产量都远远超过日本，而占世界第一。现在世界各国蚕业同行都善意地希望中国在蚕业科技发展中作出更大的贡献。发展世界蚕业科技的任务，又历史地落到了中国科技工作者的肩上。西南大学家蚕基因组研究成果，是中国对继续发展21世纪世界蚕业科技历史任务的积极回答，也是对我国蚕业科技力量的一次实地检阅。

在经典遗传理论指导下，20世纪家蚕育种工作取得了长足进步，促进了蚕业生产的发展，其中包括一代杂种的推广普及。另外，值得一提的是，田岛弥太郎博士1940年用辐射诱变，育成了雌蚕有斑纹、雄蚕无斑纹的限性遗传蚕品种，为一代杂种制造时利用幼虫斑纹鉴别雌雄开辟了道路；日本普及的现行品种中斑纹限性蚕已占40%以上，还有发展的趋势。50年代田岛和司德隆尼科夫（V. A. Strunnikov, 1914~2005）分别独立研究并获得了卵色限性的蚕品种（黑白卵）。60年代司氏应用辐射遗传与人工单性生殖的手段，育成了家蚕性连锁平衡致死系，使专养雄蚕的研究达到实用化阶段；我国浙江省蚕桑研究所引进司氏原系加以转育，现山东青州蚕种场已开始繁育，浙江部分农村示范推广。司氏从40年代开始，长期研究家蚕人工单性生殖，发现减数分裂单雌生殖，产生全部同基因型纯合雄蚕，开家蚕性别调控之先河，这在他性连锁平衡致死系的培育中也是关键技术。

毫无疑问，家蚕基因组研究，肯定对促进蚕业科学与蚕业生产有不可估量的作用；分子育种，也肯定能提高家蚕品种改良的效率，并创造自然界不存在的全新有用种质资源。不过，任何一种理论，从基础研究到实际应用往往有一个过程。无条件地都要求立竿见影，是不合理的，也不公平。任何浮躁情绪，都不是科学态度。改良家蚕品种，是家蚕基因组研究的目的之一，现在还是万里长征第一步，途中有许多艰巨的工作要完成。

如上所述，家蚕基因组研究除直接影响家蚕遗传育种外，对从基因层面、分子水平分析研究家蚕生理、生化及病理机制亦有指导作用。这里就需要动员不同学科力量，分工协作，从不同角度在基因与基因组层面上深入研究，共创蚕业科学未来，推动蚕业生产发展。相信该书的出版发行，能起到这个作用，达到这个目的。是为序。

八五奥 吕鸿毅谨识

二〇一一年十一月二十五

北京·中国科学院



第 1 章 家蚕基因组计划 / 001

- 1.1 基因组计划的产生和发展
- 1.2 家蚕基因组计划
- 1.3 家蚕基因组框架图
- 1.4 家蚕基因组精细图
- 1.5 家蚕基因组遗传变异图

第 2 章 家蚕基因组结构和进化 / 029

- 2.1 家蚕核基因组
- 2.2 家蚕核基因组中的重复序列
- 2.3 家蚕线粒体基因组
- 2.4 家蚕起源进化研究
- 2.5 家蚕的驯化

第 3 章 蚕的基因及其表达 / 065

- 3.1 家蚕的基因
- 3.2 家蚕基因的功能注释
- 3.3 家蚕全基因组表达图谱
- 3.4 家蚕mRNA的选择性剪接

第 4 章 蚕的基因家族 / 093

- 4.1 基因家族的形成和进化
- 4.2 家蚕锌指蛋白基因家族
- 4.3 家蚕30K蛋白基因家族
- 4.4 家蚕同源盒子 (Homeobox) 基因家族
- 4.5 家蚕P450基因家族
- 4.6 家蚕表皮蛋白基因家族
- 4.7 家蚕神经肽基因家族

第 5 章 家蚕丝氨酸蛋白酶及其抑制剂 / 141

- 5.1 家蚕丝氨酸蛋白酶
- 5.2 家蚕丝氨酸蛋白酶抑制剂

第 6 章 家蚕变态发育的激素调控 / 169

- 6.1 家蚕变态发育与相关激素
- 6.2 家蚕蜕皮激素
- 6.3 家蚕保幼激素
- 6.4 家蚕变态发育相关的神经肽激素

第 7 章 家蚕性别决定基因及分子机制 / 197

- 7.1 家蚕的性决定
- 7.2 家蚕性别决定网络
- 7.3 家蚕性别决定的分子机制

第 8 章 家蚕免疫防御的分子机制 / 211

- 8.1 家蚕基因组中的免疫相关基因
- 8.2 家蚕抗菌肽
- 8.3 家蚕抗菌肽基因表达调控机制
- 8.4 家蚕全基因组免疫应答

第 9 章 家蚕丝蛋白的生物合成 / 243

- 9.1 家蚕丝腺发生和生长
- 9.2 家蚕丝蛋白基因表达及调控
- 9.3 家蚕茧丝突变品种
- 9.4 家蚕丝腺组织基因组表达谱
- 9.5 家蚕丝腺的改造与应用

第 10 章 家蚕的食性与化学感受受体 / 269

- 10.1 家蚕的食性遗传
- 10.2 家蚕食性决定机制
- 10.3 家蚕化学感受信号转导途径中相关蛋白
- 10.4 家蚕对桑叶的适应机制

第 11 章 家蚕小RNA / 311

- 11.1 microRNA概述
- 11.2 家蚕microRNA的基因组鉴定
- 11.3 家蚕microRNA的时序表达
- 11.4 家蚕microRNA的空间分布
- 11.5 家蚕*let-7*的表达和功能研究
- 11.6 家蚕的其他小RNA

第 12 章 蚕的蛋白质组 / 355

- 12.1 蛋白质组学及其分析技术
- 12.2 家蚕蛋白质组学研究
- 12.3 家蚕重要功能蛋白质的鉴定
- 12.4 家蚕蛋白质组数据库

第 13 章 家蚕功能基因组平台技术 / 377

- 13.1 家蚕功能基因组学研究平台
- 13.2 家蚕转基因技术
- 13.3 家蚕RNA干扰技术
- 13.4 家蚕突变基因定位克隆技术

第 14 章 家蚕基因组生物信息数据库 / 413

- 14.1 家蚕相关生物信息学资源
- 14.2 家蚕基因组数据库SilkDB的访问与使用
- 14.3 家蚕功能基因的信息发掘

第 1 章

家蚕基因组计划

1.1 基因组计划的产生和发展 \ 001

1.1.1 人类基因组计划 \ 001

1.1.2 重要模式生物基因组计划 \ 002

1.2 家蚕基因组计划 \ 004

1.2.1 家蚕基因组计划的提出 \ 004

1.2.2 家蚕基因组框架图 \ 004

1.2.3 家蚕基因组精细图 \ 005

1.2.4 家蚕基因组遗传变异图 \ 006

1.2.5 家蚕功能基因组学研究 \ 006

1.3 家蚕基因组框架图 \ 006

1.3.1 家蚕基因组框架图测序的策略 \ 007

1.3.2 家蚕EST测序 \ 007

1.3.3 家蚕基因组文库构建及测序 \ 008

1.3.4 家蚕基因组框架图序列的组装 \ 008

1.3.5 家蚕基因组框架图的注释 \ 010

1.4 家蚕基因组精细图 \ 013

1.4.1 家蚕基因组精细图组装策略 \ 013

1.4.2 家蚕基因组精细图的组装 \ 014

1.4.3 家蚕基因组精细图序列的染色体定位 \ 015

1.4.4 家蚕基因组精细图的注释 \ 016

1.5 家蚕基因组遗传变异图 \ 018

1.5.1 基因组遗传变异图 \ 018

1.5.2 家蚕SNP研究 \ 019

1.5.3 家蚕高精度遗传变异图构建策略 \ 019

1.5.4 家蚕高精度遗传变异图的特征 \ 020

1.5.5 家蚕高精度遗传变异图的应用前景 \ 023

程道军 李 东

基因组 (genome) 一词是1920年Winkles从Gene和Chromosome两个词缩合而成的, 用于描述生物的全部基因和染色体组成。20世纪末, 以美国为首的各国科学家联合启动了“人类基因组计划 (Human Genome Project, HGP)”, 而人类基因组计划的实施催生了相关模式生物基因组研究的开展。家蚕基因组计划也应运而生。

1.1 基因组计划的产生和发展

1.1.1 人类基因组计划

解开生命之谜是人类多少年来的期盼。1971年, 美国投巨资启动的“向癌症开战 (War on Cancer)”计划, 应该是人类认识生命的第一个大规模行动, 虽然该计划未能如愿而终, 但使人们意识到, 仅靠单一学科, 如细胞学、发育学、肿瘤学、人类遗传学或分子生物学来阐明生物学的某个机制会显得苍白无力。1986年3月, 诺贝尔奖获得者、美国著名生物学家Dulbecco在*Science*杂志上发表了题为《癌症研究的转折点——人类基因组的全序列分析》的短文 (Dulbecco, 1986), 指出要真正认识人类疾病发生的本质, 必须从整体上研究和分析人类基因组及其序列。Dulbecco的这一倡议引起了生物界和医学界的广泛讨论。1986年6月, 美国能源部提出并宣布了“人类基因组计划”草案。1987年初, 美国能源部与美国国立卫生研究院为人类基因组计划下拨启动经费550万美元。1989年, 美国成立“国家人类基因组研究中心”, 由诺贝尔奖获得者、DNA分子双螺旋结构模型的提出者Watson任第一任主任。1990年, 美国国会正式批准能源部提出的人类基因组计划在1990年10月1日启动, 总体计划是在15年内至少投入30亿美元, 对人类全基因组进行测序分析。随着美国人类基因组计划的提出与实施, 英国、德国、法国、日本和中国等国的科学家相继参与了这一宏大的基因组计划, 形成了国际基因组测序联盟, 该计划也因而扩展为国际合作的国际人类基因组计划。

尽管国际人类基因组计划一开始就表明是“国际参与、免费分享”的“公益计划”, 但它不言而喻的巨大的经济价值, 使得它还没完成时, 基因的专利之争就已开始。人类基因组的专利之争分两大阵营: 一是由各国公共资金资助的科学家, 通称“人类基因组计划”; 另一方是私营生物公司, 美国赛莱拉 (Celera) 公司是其主要代表。1998年, 赛莱拉公司成立时声称在三年内, 以一种新的策略完成“人类基因组计划”的全部序列, 但他们声称要对序列垄断, 在发布时间上拖延; 待把数据分析完毕, 筛选重要的基因为己所用, 然后再待价而沽。

不管怎样, 以人类基因组全序列测定为核心内容的序列图谱作图计划则由于全世界的共同努力以及私人公司的介入而提前完成。2000年6月26日, 人类基因组遗传密码的第一份完整草图正式宣告绘制完成, 人类基因组计划国际协作组织代表Francis Collins和塞莱拉公司代表Craig Venter共同出席了由美国总统克林顿在白宫举行的新闻发布会。人类基因组框架图的完成也罕

有地引起了各国领导人的强烈关注。中国前国家主席江泽民表示，人类基因组计划是人类科学史上的伟大科学工程，希望中国科学家再接再厉，为人类基因组最终序列图的完成，为我国在功能基因组学中的创新研究作出更大的贡献。美国前总统克林顿则表示，这份工作草图是“人类历来最神奇的一份图谱”，人类因此得以了解“上帝创造生命所用的语言”。英国前首相布莱尔说：“这是医学史上的重大革命，其影响效应将远胜于抗生素的发现，这也是人类在21世纪的第一项重大科技成就”。随后，人类基因组计划国际协作组织和赛莱拉公司先后在*Nature*和*Science*上发表了各自完成的人类基因组序列框架图（Lander et al, 2001; Venter, 2003）。

1.1.2 重要模式生物基因组计划

人类基因组计划在实际执行中形成了两个概念：一是以解析人类全基因组序列为目标的狭义的人类基因组计划；另外一个则是广义的人类基因组计划，是指除了完成人类基因组序列分析外，还要对一些重要的模式生物及具有经济价值的生物开展基因组研究。

目前，用于基因功能研究的许多技术方法首先是在大肠杆菌、酵母、果蝇、线虫和小鼠等模式生物的研究中发展起来的。这不仅是因为模式生物的基因组相对较小，在研究基因功能时容易操作；由于基因功能研究最好在活体水平进行，但许多研究在没有明确的定论前，不宜直接在人类自身开展，必须通过对模式生物的研究来获得有价值的线索。因此，开展模式生物的基因组研究就成了人类基因组计划最有力的补充。如今，许多模式生物的基因组序列也在人类基因组计划执行的同时得到了解析，这些模式生物不仅包括传统的模式生物如大肠杆菌、酵母、果蝇、线虫和小鼠等，还包括许多后来发展建立起来的模式生物，如涉及动物的秀丽隐杆线虫（*Caenorhabditis elegans*）、斑马鱼（*Brachydanio rerio*）、冈比亚按蚊（*Anopheles gambiae*）、埃及伊蚊（*Aedes aegypti*）、蜜蜂（*Apis mellifera*）和赤拟谷盗（*Tribolium castaneum*）等，植物中的拟南芥（*Arabidopsis thaliana*）、水稻（*Oryza sativa* L.）、玉米（*Zea mays* L.）、小麦（*Triticum aestivum* L.）、黄瓜（*Cucumis sativus* L.）等，微生物的流感嗜血杆菌（*Haemophilus influenzae*）、严重急性呼吸系统综合征冠状病毒（SARS）等。截至本书成稿为止，即2010年10月，GenBank上已经公布的包括模式生物在内的各种物种基因组多达6485种。模式生物基因组的完成，不仅是人类基因组计划的补充，有助于探索生命进化和揭示更多生物之间差别的奥秘，而且可以为动植物经济性状的分子改良提供巨大支撑。本章仅列举果蝇等部分昆虫基因组研究的概要。

果蝇基因组

最早完成基因组序列测定的模式昆虫是黑腹果蝇（*Drosophila melanogaster*）。黑腹果蝇属于双翅目，是遗传学研究的百年经典模式生物。2001年，Adams等采用鸟枪法（也称霰弹法，shotgun）完成了黑腹果蝇基因组的全序列测定（Adams et al., 2000）。全基因组序列组装分析显示，黑腹果蝇的基因组大小为120Mb（1Mb=100万碱基对），有13 600个蛋白质编码基因，而超过60%的人类疾病基因在黑腹果蝇中都能找到直系同源基因，包括与脑疾病、神经分化和癌症相关的基因。黑腹果蝇基因组序列的完成也开启了果蝇功能基因组研究的序幕。例如，黑腹果蝇全基因组表达芯片的设计和应用发现了大量的基因功能线索，转录组测序则获得了诸多基因的表达调控信息。

2007年, 多国研究人员组成的研究小组完成了果蝇12个种的全基因组测序分析。这是世界上首次进行如此大规模的基因组测序研究, 研究成果在*Nature*上公布, 并同期发表了一组相关的研究论文 (Clark et al., 2007a; Stark et al., 2007)。比较分析发现, 果蝇基因组不仅不同区段受到的选择压力不尽相同, 进化速度差异明显, 而且不同类型基因的进化速度也存在显著差异, 进化最快的是与免疫和防御、味嗅觉、解毒和代谢、性别和繁殖相关的基因。另外, *Genome Research*等国际期刊上也相继发表了40多篇基于这些数据所完成的相关研究报告 (Hahn et al., 2007; Heger and Ponting, 2007), 在基因组研究领域引起轰动。

蚊子基因组

蚊子属于双翅目, 能通过吸取人或畜的血来传播多种疾病。目前, 冈比亚按蚊和埃及伊蚊是研究较多的两种蚊子亚科昆虫。冈比亚按蚊是疟疾的传播媒介。2001年, 一个由多国公共及私营研究机构组成的国际联盟采用鸟枪法完成了冈比亚按蚊基因组序列的测定。分析发现, 冈比亚按蚊基因组大小为278Mb, 包含约14 000个蛋白质编码基因, 其中一些基因发生了物种特异性的扩增, 如与细胞黏附和免疫等相关的基因 (Holt et al., 2002)。Christophides等从冈比亚按蚊基因组中鉴定了与昆虫先天免疫相关的18个基因家族的242个基因 (Christophides et al., 2002), 与黑腹果蝇相比, 冈比亚按蚊中参与模式识别、信号调制和效应应答的免疫相关基因的数目发生了显著的扩张, 这可能是由于两种昆虫因不同病原物侵染而受到了不同的选择压力所造成的。

埃及伊蚊是传播登革热和黄热病的主要媒介。2007年, Nene等完成了对埃及伊蚊基因组的测序 (Nene et al., 2007)。分析发现, 埃及伊蚊的基因组中有50%的序列属于转座子, 表明其基因组受到了转座子的侵袭。在转座子侵袭后, 埃及伊蚊的基因组也显著增大, 其大小达到1.38Gb, 是冈比亚按蚊基因组的5倍。不仅如此, 基因的平均长度和基因间区的长度也比冈比亚按蚊和黑腹果蝇增加了4~6倍。与冈比亚按蚊相比, 埃及伊蚊的部分基因发生了物种特异性扩增, 如气味结合蛋白、细胞色素P450和表皮蛋白基因等。这些分析有助于了解不同类别蚊子间为什么会对血液具有偏好性, 以及对不同病原具有不同的传播能力等。

蜜蜂基因组

蜜蜂属于膜翅目, 其授粉作用对于农业生产是必需的, 而且也是研究昆虫社会行为的典型模式。2006年, 蜜蜂基因组测序和序列分析结果在*Nature*上发表 (Honeybee Genome Sequencing Consortium, 2006)。结果显示, 蜜蜂基因组大小为236Mb, 含10 157个蛋白质编码基因。同黑腹果蝇和冈比亚按蚊基因组的比较发现, 蜜蜂基因组中转座子数目较少, 且进化速度相对较慢。另外, 蜜蜂含有较少的先天免疫基因、解毒酶基因、表皮蛋白形成基因和味觉受体基因, 但气味受体基因及花蜜和花粉利用相关的新基因数目较多, 这与蜜蜂的生态学和社会组织等生物学特征一致。除此主要结果外, *Insect Molecular Biology*和*Genome Research*还分别于2006年和2007年专刊发表了蜜蜂基因组研究的一系列成果。

赤拟谷盗基因组

赤拟谷盗属于鞘翅目, 是农业上的甲虫类害虫之一, 主要破坏干燥的农产品及其制品, 如