

李成涛 赵书民 柳 燕 编著

DNA 鉴定前沿



科学出版社

DNA 鉴定前沿

李成涛 赵书民 柳 燕 编著

科学出版社

北 京

内 容 简 介

本书是一本专门介绍 DNA 鉴定技术前沿的专著。全书分为 5 部分。第一部分为绪论,详细介绍了 DNA 鉴定的发展历程、基因、遗传信息和遗传标记的现代概念。第二部分为 DNA 鉴定传统遗传标记的发展动态和研究进展,包括短串联重复序列、单核苷酸多态性,以及其他二等位基因遗传标记、线粒体 DNA 等。第三部分为 DNA 鉴定新型遗传标记的最新研究成果,包括表观遗传标记、拷贝数变异及 RNA 分子等。第四部分为血缘关系判定,介绍了依据 STR 分型结果进行各类血缘关系鉴定的方法,包括统计学基础、亲权分析、各种亲权评估参数的计算、复杂的隔代亲权鉴定,以及突变时亲权指数的计算方法和亲权评价,特别对亲权鉴定结果的判断标准进行了系统探讨。第五部分为 DNA 鉴定技术的新应用,如 DNA 鉴定技术在法医植物学、法医昆虫学和法医微生物学中的应用。书的最后还收集了国内外的一些经典案例。全书参考了大量的国内外文献,内容丰富,对 DNA 鉴定的理论研究和实践有较强的指导作用。

本书可作为 DNA 鉴定工作者、科研工作者及高校教师、研究生、本科生的参考用书。

图书在版编目 (CIP) 数据

DNA 鉴定前沿/李成涛,赵书民,柳燕编著. —北京:科学出版社, 2011

ISBN 978-7-03-031846-6

I. ①D… II. ①李…②赵…③柳… III. ①脱氧核糖核酸-生物鉴定
IV. ①Q523

中国版本图书馆 CIP 数据核字 (2011) 第 137940 号

责任编辑:李悦 刘晶/责任校对:张怡君

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.sciencep.com>

新蕾印刷厂印刷

科学出版社发行 各地新华书店经销

*

2011 年 8 月第 一 版 开本: B5 (720×1000)

2011 年 8 月第一次印刷 印张: 22 1/4 插页: 2

印数: 1—2 500 字数: 533 000

定价: 75.00 元

(如有印装质量问题, 我社负责调换)

序

在过去的 25 年间，DNA 鉴定技术获得了长足的发展，已经逐渐发展成为一门独立的学科，为各国政府接受并采用。DNA 双螺旋模型的发现，是生命科学研究历程中的一个具有划时代意义的里程碑。这一模型对遗传学发展具有深远影响，它不仅使遗传学研究从此深入到分子水平，而且奠定了现代遗传学的基础，进一步推动和影响着生命科学各个学科的飞速发展。在遗传学不断发展、取得各项研究成果的同时，法医遗传学也有了很大的发展，使得法医工作者们能够运用和依靠遗传学的这些新技术、原理来为有关的法律问题服务，如对个体中的组织、细胞、体液的同一认定及个体之间的血缘关系鉴定等。

随着基因组时代的到来，DNA 鉴定技术的发展日新月异，并呈现出多学科的交叉化和综合化。近些年来，DNA 鉴定技术已经成为司法鉴定领域的研究热点和发展前沿，一大批崭新的技术、方法得以开发并应用。对这些研究成果的及时总结、归纳与提炼有助于促进 DNA 鉴定行业的整体技术进步与统一。

该书的几位作者都是近年在国内外司法鉴定学术界非常活跃、多有建树的年轻学者。近些年来，在国内外杂志上陆续刊登了他们的系列研究论文。这一团队在 DNA 鉴定研究，特别是在复杂血缘关系鉴定、新型遗传标记的开发方面积累了较为丰富的实践经验。

该书的实战性很强，是一本专门介绍 DNA 鉴定技术前沿的专著，系统地介绍了传统遗传标记与新型遗传标记的最新研究成果、各类血缘关系鉴定的最新技术、新型 DNA 鉴定技术，以及国内外的经典案例。各类读者可以直接切入感兴趣的问题，提供解决实际问题的参考。

将该书推荐给大家，希望能对大家的工作有所裨益，并使其成为一本很有实用价值的参考书籍。更希望它的出版，能够进一步推动 DNA 鉴定技术在法庭科学中的应用。

薛京伦

2011 年 3 月

前 言

DNA 鉴定也称为法医 DNA 分析，同其他科学一样是在长期的社会实践中，特别是在司法鉴定实践中逐步形成并发展起来的。DNA 鉴定包括亲权鉴定和个体识别，它是指应用生物学和遗传学等理论知识与技术，通过检测分析人类的遗传学标记，判断个体之间的亲缘关系和现场检材的来源。近年来，人们越来越深刻地认识到 DNA 鉴定在各类刑事案件侦破和民事案件解决中所发挥的重要作用。俄国末代沙皇遗骸的识别、萨达姆·侯赛因身份的确认、美国总统克林顿的性丑闻、美国“9·11”事件大量遗体身份的认定、蚊子帮助芬兰警察确定盗车贼、曹操墓主人身份的 DNA 鉴定等，让人们对于 DNA 鉴定应用的巨大前景和威力充满了无数的期待。

随着基因组时代的到来和成果的产出，DNA 鉴定技术的发展日新月异，并呈现出多学科的交叉化和综合化，DNA 鉴定的内容也发生了深刻变化，传统的 DNA 鉴定技术已经不能胜任这一快速变化。近些年来，DNA 鉴定技术的开发得到了足够的重视，并成为司法鉴定领域的研究热点和发展前沿，一大批崭新的技术、方法得以开发并应用，但是，关于 DNA 鉴定技术发展前沿的专著甚少。

我们编写这本《DNA 鉴定前沿》，旨在介绍近年来 DNA 鉴定领域的现状和发展趋势，从而为推动司法鉴定领域的 DNA 鉴定贡献微薄之力。全书共分 5 部分，包括 16 章，系统地介绍了传统遗传标记与新型遗传标记的最新研究成果、各类血缘关系鉴定的最新技术和新型 DNA 鉴定技术等。

限于篇幅，本书并没有包括 DNA 鉴定涉及的全部内容，如 DNA 鉴定的基本理论和方法、质量控制、DNA 数据库，以及目前已经很少使用的技术和方法。

在本书即将出版之际，著名的医学与人类遗传学家薛京伦教授欣然为本书作序，使本书增色不少。

本书作者均系在这一研究领域中比较活跃的学者。应该说，本书系统地、广泛地反映了该领域的研究热点与成就。我们热切期望本书能为本领域的 DNA 鉴定人员、科研工作者、相关教师和学生提供有价值的新知识和参考信息，这将是

给我们的最大安慰。由于这一领域的发展极为迅速，本书肯定存在诸多不足之处，敬请读者批评指正。同时，我们将以更高的热情投入该学科的研究并跟踪学科进展，为读者提供新的信息。

编 者

2011 年 3 月 10 日

目 录

序

前言

绪论	1
第一节 法医遗传学的发展	1
第二节 基因的现代概念	2
第三节 遗传信息的现代概念	4
第四节 遗传标记的现代概念	6
参考文献	8

第一篇 遗传标记与分型技术

第一章 STR 遗传标记在 DNA 鉴定中的应用现状与新进展	13
第一节 STR 基因座及商业化分型试剂盒的历史回顾	13
第二节 常用 STR 基因座的生物信息学分析	15
第三节 常用 STR 基因座在中国人群中的基因频率及法医学参数	18
第四节 常用 STR 基因座的突变率分析	19
第五节 STR 分型技术	26
第六节 STR 基因座研究的新进展	35
参考文献	53
第二章 单核苷酸多态性在 DNA 鉴定中的应用现状与新进展	57
第一节 SNP 在法医学中的应用前景	57
第二节 SNP 分型方法和技术	60
第三节 SNP 在法医学中应用的思考	72
参考文献	73
第三章 其他二等位基因遗传标记在法医学中的应用	76
第一节 Alu 插入多态性遗传标记在法医学中的应用	76
第二节 插入缺失多态性遗传标记在法医学中的应用	81
参考文献	89
第四章 线粒体 DNA 在法医学中的应用	91
第一节 线粒体的分子遗传学	91

第二节 线粒体 DNA 在法医学中的应用进展	92
第三节 线粒体 DNA 在人类起源和进化研究中的应用	97
第四节 线粒体 DNA 的异质性	99
参考文献	103

第二篇 DNA 鉴定中的新型遗传标记

第五章 表观遗传标记在法医学中的应用	107
第一节 表观遗传学的发展	107
第二节 表观遗传现象	110
第三节 表观遗传学与遗传学的关系	115
第四节 表观遗传学对医学的影响	118
第五节 表观遗传学在法医学中的应用	121
第六节 DNA 甲基化在法医学中的应用前景及其检测方法新进展	127
参考文献	133
第六章 拷贝数变异研究新进展	136
第一节 拷贝数变异的概念与发展	136
第二节 拷贝数变异的研究进展	138
第三节 拷贝数变异在法医学中的应用前景	141
参考文献	142
第七章 RNA 分子在法医学中的应用	144
第一节 RNA 的概念与功能	144
第二节 RNA 的分类与意义	146
第三节 RNA 在法医学中的应用研究	148
参考文献	155

第三篇 血缘关系判定

第八章 统计学基础	159
第一节 统计学中的几个基本概念	159
第二节 常用的统计指标	160
第三节 频率与概率	164
第四节 几种常用的概率分布	165
第五节 参数估计与假设检验	171
第六节 条件概率与贝叶斯定理	174

第九章 三联体和二联体亲权鉴定	176
第一节 父权后概率和父权指数.....	176
第二节 父权指数的计算公式及其简化运算.....	179
第三节 三联体亲权鉴定中的父权认定与排除.....	189
第四节 二联体亲权鉴定中的亲权认定与排除.....	202
参考文献.....	211
第十章 其他类型血缘关系指数的计算与分布规律模拟研究	212
第一节 ITO 法计算其他血缘关系指数	212
第二节 不同血缘关系指数计算公式的简化运算.....	216
第三节 不同血缘关系指数的通用计算方法.....	220
第四节 其他血缘关系指数分布规律的模拟研究.....	222
参考文献.....	224
第十一章 基于常染色体 STR 进行全同胞关系判定的实验研究	225
第一节 概述.....	225
第二节 累积全同胞指数在全同胞对和无关个体对人群中的分布规律 研究.....	227
第三节 全同胞关系判定中的判别分析方法.....	230
第四节 全同胞关系判定准则的比较.....	235
结语.....	246
参考文献.....	247

第四篇 DNA 鉴定中的其他问题

第十二章 肿瘤组织身源鉴定新进展	251
第一节 肿瘤组织身源鉴定现状.....	251
第二节 常用 STR 基因座在消化系统肿瘤组织中的变异分析	254
第三节 共有基因座数和共有等位基因数在肿瘤组织身源鉴定中的应用	274
第四节 插入缺失多态性遗传标记在肿瘤组织身源鉴定中的应用.....	280
参考文献.....	287
第十三章 甲醛固定石蜡包埋组织及 H. E 染色切片的 DNA 检测	288
第一节 甲醛固定石蜡包埋组织及 H. E 染色切片的检验现状	288
第二节 甲醛固定与石蜡包埋组织和 H. E 染色切片对 DNA 检测的影响	289
第三节 甲醛固定和石蜡包埋组织及病理切片 DNA 提取	294

第四节	DNA 质量的测定	296
第五节	甲醛固定石蜡包埋组织及 H. E 染色切片 DNA 检测的最新技术 进展	303
	参考文献	305
第十四章	非人类 DNA 在法医遗传学中的应用	307
第一节	非人类 DNA 在案件侦破中的应用	307
第二节	如何用动物 DNA 作为证据	308
	参考文献	310
第十五章	AFLP 分子标记技术的新进展及其在法医植物学中的应用	312
第一节	AFLP 技术的新进展	312
第二节	AFLP 技术在法医植物学中的应用	314
	参考文献	314
第十六章	法医昆虫学研究新进展	316
第一节	现代分子生物学技术在昆虫种类鉴定中的应用	316
第二节	应用于昆虫种类鉴定的常用分子标记检测技术	321
第三节	常见嗜尸性昆虫 DNA 提取方法及部位	324
第四节	从取食人体组织的昆虫嗉囊内容物中检测人类遗传标记的应用	330
	参考文献	333
附录	法医 DNA 分析技术在经典案例中的应用	336
案例 1	法医 DNA 分析的首次应用	336
案例 2	DNA 证据和莱温斯基的蓝裙子	336
案例 3	认真采集 DNA 证据的重要性: 辛普森案	337
案例 4	天然的混合样本及嵌合个体	338
案例 5	预测生物样本的种族来源帮助犯罪调查	339
案例 6	成吉思汗的遗传遗产	339
案例 7	无名士兵墓中遗骸的个体识别	340
案例 8	俄国末代沙皇遗骸的识别	341
案例 9	DNA 结果的重要性	342
案例 10	证实萨达姆·侯赛因的身份	343

绪 论

第一节 法医遗传学的发展

法医遗传学同其他科学一样是在长期的社会实践中，特别是在司法鉴定实践中逐步形成并发展起来的学科。公元 1247 年宋慈所著《洗冤集录》一书中就讲到“滴血之法，孙亦可验祖”。这种滴血认亲的方法，虽不很完备，但是它包含着法医遗传学的萌芽，或者说是应用血型遗传原理鉴定亲子关系的最早尝试。1865 年，孟德尔的研究揭开了遗传研究的序幕，创立了遗传学的基本规律，即分离定律和自由组合定律。1893 年，奥地利的汉斯·格劳斯所著《检验官手册》已将运用科学技术办案写入书中。1900 年，Landsteiner 发现 ABO 血型以后，人类红细胞血型应用于检案，法医物证检验步入了科学时代。1910 年，法国刑事犯罪学家爱德蒙·洛卡德提出了接触与物质交换的原理，表述为“任何接触都可以留下痕迹”，这一观点奠定了现代法庭科学的基础。1926 年，摩尔根基因学说的发表，为法医遗传学的发展奠定了基础，成为法医遗传学的基本理论。

在以后的岁月中，遗传学的研究突飞猛进，尤其是 20 世纪 50 年代，遗传物质 DNA 双螺旋模型的发现，是生命科学研究历程中一个具有划时代意义的里程碑。这一模型对遗传学发展具有深远影响，它不仅使遗传学研究从此深入到分子水平，而且奠定了现代遗传学的基础，进一步推动和影响着生命科学各个分支的飞速发展。在遗传学不断发展、取得各项研究成果的同时，法医遗传学也取得了很大的进展，使得法医工作者们能够运用和依靠遗传学的这些新技术、新原理来为有关的法律问题服务，如对个体中的组织、细胞、体液的同一认定等。1958 年白细胞抗原系统的发现以及 60 年代应用电泳检测血清型和酶型等，为法医物证检验与鉴定提供了更多的技术手段。70 年代，应用等电聚焦发现了多种血清型及酶型的亚型，进一步提高了个人识别概率。1985 年，英国科学家 Jeffreys 研究人类肌红蛋白基因结构时，在第一内含子中发现了一段由 33bp 串联重复的小卫星序列。以 33bp 为核心序列串联重复的单链 DNA 作为 RFLP 分析的探针，杂交结果表明，其可在 4~23kb 的范围内检出 20~30 条多态片段，多态性信息量极大，个体的条带模式独一无二，类似经典的指纹，故称 DNA 指纹。DNA 指纹的高度个体特异性克服了传统法医遗传标记鉴别能力低的缺陷，使法医个体识别和亲子鉴定实现了从仅能排除到高概率认定的飞跃，被誉为法医物证分析的里程碑。1993 年，国际法医遗传学会推广了以 STR (short tandem repeat) 为核

心的第二代 DNA 指纹或 DNA 纹印技术。不仅实现了法医物证检验高概率的认定,而且为法医 DNA 分型技术的标准化铺平了道路。

法医学 (forensic medicine, legal medicine) 是应用医学及其他自然科学的理论与方法,研究并解决立法、侦查、审判实践中涉及的医学问题的一门科学。法医学既是一门应用医学,又是法学的一个分支。法医学为制定法律提供依据,为侦查、审判提供科学证据。法医学的诞生和发展,与社会经济的发展、法的出现,以及医学和其他自然科学的进步有着密切的关系。因此,法医学是联结医学与法学的一门交叉科学。现代法医学分为基础法医学和应用法医学两部分,前者研究法医学的原理和基础;后者则运用法医学的理论和方法,解决司法、立法和行政上的有关问题。很长一段时间内,亲权鉴定、血型检验、性别鉴定等一系列问题都笼而统之被包括在法医学这门科学的研究范围之内。但是,随着社会的进步、科学的发展,各学科的分类也越来越细,新的边缘学科不断形成,各门新学科先后诞生,包括法医毒物化学、临床法医学、法医精神病学和法医遗传学。

随着科学技术的发展,法医遗传学在司法鉴定中的应用越来越广泛,而不只认为其与亲权鉴定、血型检验有关,如今遗传学的各项研究纷纷进入分子水平,从而使法医遗传学的研究与应用也进入了分子水平。在司法精神疾病鉴定中,除对被鉴定人进行临床观察鉴定外,有的则需要分析是否由遗传因素决定,这种分析往往用家谱分析法,目前普遍认为精神分裂症、精神发育不全、人格变态等疾病均与遗传因素有关,但这些病的具体遗传规律还不是很清楚,它是不遵循孟德尔遗传规律的。如今,法医遗传学普遍采用 DNA 遗传标记,因为它有足够的多态性,理论上可以通过 DNA 分型,而不必通过测定全基因组序列来进行个人识别。DNA 分型的优点还在于能从任何含有细胞的体液或组织中得到相同的结果,能够对陈旧斑痕和极微量的检验材料进行分型,分析结果能够成为计算机可查询的数据库形式。

当前,法医遗传学的发展迅速,miniSTR 技术、X 染色体遗传标记的研究及表观遗传学在法医中的应用等,使得法医 DNA 分析技术日臻完善,理论知识日趋丰富,解决检案问题的能力已经得到极大的提高。

第二节 基因的现代概念

20 世纪 70 年代后,基因的概念随着多学科渗透和实验手段日新月异,取得了突飞猛进的发展。大量的成果无疑给基因的概念注入了鲜活科学的内容,帮助人们揭开层层面纱去更加全面地了解基因的真面目。时代在发展,科学在进步,基因概念的深入发展必将对人类的文明进步产生强大的推动作用。

(一) “基因”概念的提出

1865年,孟德尔报道了性状遗传的分离定律和自由组合定律,为解释这些遗传现象,提出了决定性状遗传的遗传因子说。1910年后,摩尔根等发现了伴性遗传和连锁现象,第一次证明基因在染色体上呈线性排列,彼此间有连锁遗传的倾向,而不同染色体间基因的遗传则遵循孟德尔规律。可见,摩尔根的理论丰富和发展了孟德尔的遗传因子理论。

基因是遗传学最基本的概念,1909年由遗传学家约翰森(W. L. Johannsen, 1857~1927)首次提出,用来表示遗传的独立单位,相当于孟德尔在豌豆试验中提出的遗传因子,泛指控制生物性状且按孟德尔规律传递的遗传因子。随着生命科学研究的不断深入,基因的概念不断地被修正和发展。20世纪50年代以后,随着分子遗传学的发展,1953年在沃森和克里克提出DNA的双螺旋结构以后,人们普遍认为基因是DNA的片段,阐明了基因的化学本质,因此基因又被定义为“具有特定遗传效应的DNA片段”。60年代,本泽(S. Benzer, 1921~)又提出了基因内部具有一定的结构,可以区分为突变子、互换子和顺反子三个不同的单位。

(二) 人类基因组计划的完成丰富了基因概念的内涵

人类基因组计划的完成和技术的发展,极大地丰富了近代基因概念的内涵,基因的定义已经不再局限于编码蛋白质的DNA序列这个概念,因为通过近年来的研究发现,编码蛋白质的DNA序列占全基因组序列的很小部分。在人类,蛋白质编码序列占全基因组序列的3%~5%。与此同时,越来越多的证据表明,许多RNA“基因”具有明确的生理功能,但却不编码任何蛋白质,它们仅是以RNA形式发挥功能,而且这些非编码RNA的数量似乎与物种的复杂性相关,如rRNA和tRNA就属于这种类型。另外,还有一类基因,如操纵基因,它们既没有转录作用,又没有翻译产物,仅仅起着控制和操纵基因活动的作用。特别是人类基因组全序列确定后发现,在DNA分子上有相当一部分片段只是某些碱基的简单重复,它们不编码蛋白质,但是在真核细胞生物中这些片段的数量很大,甚至占全基因组序列55%以上。这些重复碱基片段的功能目前还不十分了解,推测可能与某些基因活动调节和染色体结构的稳定有关。因此,应该把基因看做是DNA分子上具有特定功能的(或具有一定遗传效应的)核苷酸序列,而不仅仅是编码蛋白质的DNA序列。

(三) 蛋白质组学研究使基因的概念外延

与此同时,21世纪初,蛋白质组学研究的深入使人们对基因概念再度反思。

组成蛋白质的氨基酸有 20 种, 远远多于组成脱氧核糖核酸的 4 种碱基。蛋白质多种多样, 几乎执行着生物体的所有功能。因此, 几十年来, 人们一直认为蛋白质不仅参与生物体组织和器官的组成, 还可作为酶来催化调节生物体的各种代谢活动, 并与特异的 DNA 或 RNA 序列结合以调节基因的表达, 维持生命活动有条不紊地进行。在确定人类含有 27 000 条基因的同时却发现人类含有大约 10 万种蛋白质, 显示了蛋白质水平上基因表达的多样性。另外, 通过基因组比较发现, 人类的基因大约只有 27 000 个, 与其他脊椎动物相似, 这远远少于人们的估计。尽管由于 mRNA 的可变剪接可使蛋白质种类增加一些, 但不同物种编码的蛋白质相当保守, 从蛋白质水平讲, 它不足以说明物种的复杂性和个体表现出来的差异。例如, 大约 99% 的人类蛋白质在小鼠中能找到它的类似物, 甚至许多人类蛋白质在结构和功能上与一些无脊椎动物也相似; 人类基因组序列在个体水平上有 600 万个 (0.1%) 碱基差异, 而编码蛋白质的基因突变仅有 2 万个, 这些突变大多是不影响蛋白质氨基酸序列的无效突变。科学家们也发现, 即使是基因的表达也具有时间和空间的特异性。在基因的 DNA 序列没有发生改变的情况下, 基因功能发生了可遗传的变化, 并最终导致了表型的变化。例如, 基因沉默、X 染色体剂量补偿、DNA 甲基化、组蛋白乙酰化、基因组印记等表观遗传学现象, 都不符合孟德尔遗传规律的核内遗传, 但这些特性使得细胞和生物体在保持遗传稳定性的同时能够更好地适应环境。

上述基因组学研究的进展都说明表面上基因组 DNA 序列仅仅是一份信息模板, 如何从中读出丰富多彩的信息是生命科学迫切需要研究的新课题。因此, 基因可以定义为“不仅仅是遗传的基本功能单位, 更应该是遗传信息储存和加工的单元”。21 世纪, 基因概念的外延将有可能随“表观遗传学”的发展而进一步拓展, 其内涵也将随着纳米生物学 (nanobiology) 和量子生物学 (quantum biology) 的发展而在量子水平上充实完善, 人们也将能更准确、更全面地揭示生物遗传变异的本质规律。

第三节 遗传信息的现代概念

基因的本质确定后, 人们又把研究视线转移到基因传递遗传信息的过程上。在 20 世纪 50 年代初, 人们已懂得基因与蛋白质间似乎存在着相应的联系, 但基因中的信息怎样传递到蛋白质上这一基因功能的关键课题在 60~70 年代才得以解决。

一、中心法则

中心法则主要是描述遗传信息是如何在 DNA、RNA 和蛋白质之间流动的。

从生物学中心法则出发,基因就是 DNA 分子上的许多个片段。以它们为模板经过转录和翻译能合成一条完整的多肽链。中心法则认为遗传信息流是单向的控制流,由 DNA 转录给 RNA,再由 RNA 进入细胞质传递、表达为蛋白质。随着科学的发展,人们越来越发觉事情并非如此。1970 年,巴尔的摩 (D. Baltimore) 和特明 (H. M. Temin) 在致癌的 RNA 病毒中发现了一种酶,它能以 RNA 为模板合成 DNA,他们称这种酶为依赖 RNA 的 DNA 聚合酶,称为逆转录酶。这一发现进一步完善了当时的中心法则,两位科学家也因此获得了诺贝尔奖。他们发现有的 RNA 可以作为 DNA 复制的模板进行逆转录和存在 RNA 逆转录酶的实际例,说明信息能够反向流动,即由 RNA $\rightarrow$ DNA,改变遗传信息。可以说,“DNA $\rightarrow$ RNA $\rightarrow$ 蛋白质”的遗传信息流,是生物繁衍后代保持物种稳定的转录信息流;但“蛋白质 $\rightarrow$ RNA $\rightarrow$ DNA”以及“蛋白质 $\rightarrow$ DNA $\rightarrow$ RNA”的信息流向,是促使生物不断地适应环境、不断地进化的另一个必不可少的重要的信息流。正是由于这双向的信息流,才使生物从低级到高级、从简单到复杂、从单一到丰富,从而形成丰富多彩的生命世界。

二、中心法则遇到的挑战

除了传统意义上的遗传信息外,近年来,科学家们发现了大量隐藏在 DNA 序列之中或之外更高层次的遗传信息,使遗传学家不得不重新审视以前“根深蒂固”的生命科学规律。多年来,分子生物学领域公认的中心法则遇到了前所未有的挑战。目前认为,这些高层次基因组信息主要包括非编码 RNA (non-coding RNA)、DNA 甲基化和组蛋白共价修饰等表观遗传学信息 (epigenetic information)。表观遗传学是指在基因的 DNA 序列没有发生改变的情况下,基因功能发生了可遗传的遗传信息变化,并最终导致了表型的变化。它是不符合孟德尔遗传规律的细胞核核内遗传。由此我们可以认为,基因组含有两类遗传信息,一类提供生命必需蛋白质的模板,称编码遗传信息;另一类是表观遗传学信息,它提供了何时、何地、以何种方式去应用遗传信息的指令。因此,准确地说,现在看来,中心法则不是精细、逐一地传递序列信息的通道,而是一个动态的分子遗传信息的加工流水线。

三、遗传编码信息和表观调控信息

21 世纪,生物遗传信息的范畴在扩大,遗传编码信息和表观遗传信息成为生物遗传信息的两个支柱。染色质 DNA 的甲基化是表观遗传信息的主要形式,据称人单倍体基因组有 5000 万个 CpG 位点,有“甲基化”和“未甲基化”两种形式,这样就存在巨大的 DNA 甲基化可能的组合,可储存大量的信息。此外,染色质另一组分组蛋白的氨基端发生多组合修饰,可调控其本身进入 DNA 的通

道,不同的组蛋白氨基端修饰,对染色体结合蛋白产生协同或拮抗作用,从而调控染色质转录活动或沉寂状态的动力学转换。

研究表明,一个细胞的甲基化形式大致代表了该细胞表达特征的蓝图,组蛋白氨基端修饰的组合方式构成“组蛋白密码”。这些信息表达和传递的交叉组合和相互调节,极大地扩大了遗传密码的信息储存。同时,随着功能基因组学和蛋白质组学的发展,人们逐渐认识到蛋白质空间结构的特异性和重要性,基于蛋白质和酶的空间结构特异性决定其作用物、产物的特异性,有人提出“空间密码”的理论,认为蛋白质分子的结构与修饰的模式也能体现并传递核遗传信息,谓之“蛋白质遗传”。此外,尚有“糖密码”(saccharide code or glycome)的报道等。这些都提示,遗传信息并非仅仅是中心法则所说的信息大分子中的一级序列,大分子“构型”(conformation)本身也是一种信息,而且可以是一种遗传信息。因此,基因概念的外延似乎就应随着遗传信息范畴的扩大而有所拓展。

第四节 遗传标记的现代概念

遗传标记(genetic marker)是基因型特殊的、易于识别的表现形式。它一般具有较强的多态性、表现的共显性、不影响重要的农艺性状、经济方便和易于观察记载等优点,在遗传学的建立和发展过程中起着重要作用。随着遗传学的发展,遗传标记也在不断地发展。从遗传学的建立到现在,遗传标记的发展主要经历了4个阶段,表现出4种类型。

一、形态标记

形态标记(morphological marker)是指生物的外部特征特性,包括质量性状作遗传标记和数量性状作遗传标记,如人的肤色、作物的株高、种子的粒形和动物的体重等。它是最早被使用和研究的一类遗传标记。在遗传的分离定律、自由组合定律和连锁交换定律,以及群体遗传学和数量遗传学的研究与应用等方面发挥着重要作用。它具有易于识别和掌握、简单直观等优点,但也具有标记数量少、多态性差和易受环境因素影响等不足。用人工诱变培育形态标记的材料所用周期长,许多形态标记对有益生物性状会产生不良的影响。

二、细胞标记

细胞标记(cytological marker)主要指染色体组型和带型。染色体组型是指生物体细胞所有可测定的染色体特征总称,包括染色体总数、染色体组数、每条染色体大小和形态、着丝点的位置等。它是物种特有的染色体信息之一,具有很高的稳定性和再现性,对鉴别真假杂种、研究染色体结构和数目的变异、物种起

源和生物的遗传进化等具有重要价值。带型是用染色体分带技术所产生的明显的染色带（暗带）和未染色带（明带）相间的带纹，从而使染色体呈现鲜明的个体性。因此，可以把染色体带型作为一种遗传标记，有效地识别不同物种之间及同一物种不同染色体之间的差异。带型有 Q 带、G 带、R 带、T 带等类型，但是该遗传标记也具有标记数目有限的缺点。

三、生化标记

生化标记（biochemical marker）是指生物的生化特征特性，主要包括同工酶和储藏蛋白两种标记。同工酶是指同一种属中功能相同但结构不同的一组酶，它是由不同基因位点或等位基因编码的多肽链单体、纯聚体或杂聚体。同工酶是生物体代谢的调节者，与特殊的生理功能和细胞分化相联系，也与基因进化和物种的演变有关，因此，同工酶既可作为生理指标又可作为遗传标记在生物群体的遗传进化和分类研究中广为应用。另外，同工酶研究也有助于探讨生物个体发育过程中基因的表达与调控。生化标记具有简便、经济、快速和准确等优点，但同样具有标记数目有限的缺点。储藏蛋白与种子的萌发等发育过程有关，也具有生物种属的特异性，可以作为遗传标记。例如，马铃薯蛋白是马铃薯块茎中的主要蛋白（储藏蛋白）；玉米醇溶蛋白是玉米种子内的储藏蛋白，按其结构和溶解度分为 A、B、C 和 D 四大类。虽然其他生物中无该储藏蛋白，但有其特异的储藏蛋白。

四、DNA 分子标记

DNA 分子标记（molecular marker）是能反映生物个体或种群间基因组中某种差异特征的 DNA 片段。此 DNA 片段是将基因组 DNA 经过限制性内切核酸酶切割和分子杂交或 PCR 扩增后在电泳凝胶上（或杂交膜上）检测到的。与前三种遗传标记相比，DNA 分子标记具有诸多优点，包括：①能对生物各发育时期的个体、组织、器官和细胞进行检测，直接以 DNA 的形式表现，不受环境影响，不存在基因表达与否的问题；②数量丰富，遍及整个基因组；③遗传稳定，多态性高；④对生物体的影响表现为“中性”，不影响目标性状的表现，与不良性状无必然连锁；⑤多为共显性，能为鉴别纯合基因型和杂合基因型提供完整的遗传信息；⑥操作简便等。这些优点使其广泛地应用于生物基因组研究、进化分类、遗传育种、医学和法学等方面，成为分子遗传学和分子生物学研究与应用的主流技术之一，并且对社会产生巨大冲击。目前，被广泛应用的 DNA 分子标记主要有 RFLP（限制性片段长度多态性）、RAPD（随机扩增多态性 DNA）、AFLP（扩增片段长度多态性）、STS（序列标记位点）、SSR 或 SSLP（简单重复序列）和 DNA 指纹技术等。