



研究生规划教材

全国高等医药院校教材·全国高等医药教材建设研究会规划教材

医学遗传学

供 研 究 生 用

主 编 夏家辉
副主编 刘德培



人民卫生出版社

全国高等医药院校教材

供研究生用

医学遗传学

主 编 夏家辉

副主编 刘德培

编 者 (按章节顺序)

- | | |
|-------------|---------------------------------|
| 夏家辉 | (中南大学医学遗传学国家重点实验室) |
| 孙开来、赵彦燕 | (中国医科大学医学遗传学教研室) |
| 夏 昆、郑 多、蔡 芳 | (中南大学医学遗传学国家重点实验室) |
| 余 龙 | (复旦大学生命科学院) |
| 张灼华、刘建湘、夏 昆 | (中南大学医学遗传学国家重点实验室) |
| 邬玲仟、戴和平、夏家辉 | (中南大学医学遗传学国家重点实验室) |
| 邓汉湘、梁德生、蔡 芳 | (中南大学医学遗传学国家重点实验室) |
| 顾东风 | (中国医学科学院心血管病研究所) |
| 傅松滨 | (哈尔滨医科大学基础医学院) |
| 傅松滨、张 学 | (哈尔滨医科大学基础医学院、中国医学科学院协和医科大学) |
| 唐北沙、邬玲仟、龙志高 | (中南大学湘雅医院神经内科、中南大学医学遗传学国家重点实验室) |
| 刘德培 | (中国医学科学院协和医科大学) |
| 刘春宇、潘 乾、梁德生 | (中南大学医学遗传学国家重点实验室) |

学 术 秘 书

- | | |
|-----|--------------------|
| 夏 昆 | (中南大学医学遗传学国家重点实验室) |
| 戴和平 | (中南大学医学遗传学国家重点实验室) |

人 民 卫 生 出 版 社

图书在版编目 (CIP) 数据

医学遗传学/夏家辉主编. —北京:
人民卫生出版社, 2004. 4

ISBN 7-117-06002-6

I. 医… II. 夏… III. 医学遗传学-研究生-教材 IV. R394

中国版本图书馆 CIP 数据核字(2004)第 016893 号

医学遗传学

主 编: 夏家辉

出版发行: 人民卫生出版社(中继线 67616688)

地 址: (100078)北京市丰台区方庄芳群园 3 区 3 号楼

网 址: [http://www. pmph. com](http://www.pmph.com)

E - mail: [pmph @ pmph. com](mailto:pmph@pmph.com)

印 刷: 北京人卫印刷厂

经 销: 新华书店

开 本: 850×1168 1/16 印张: 28

字 数: 690 千字

版 次: 2004 年 5 月第 1 版 2004 年 5 月第 1 版第 1 次印刷

标准书号: ISBN 7-117-06002-6/R·6003

定 价: 50.00 元

著作权所有, 请勿擅自用本书制作各类出版物, 违者必究

(凡属质量问题请与本社发行部联系退换)

全国高等医药院校研究生规划教材出版说明

《中国医学教育改革和发展纲要》明确指出，在今后的 5~15 年我国医学教育要加速发展研究生教育，到 2005 年，本专科教育（含高等职业技术教育）和研究生教育年招生总量占总体的比例要达到 60% 以上，到 2015 年增长到 70% 以上。为适应这一要求，经全国高等医药教材建设研究会和卫生部教材办公室研究决定，自 2001 年 8 月起组织编写一套供研究生使用的规划教材。此套教材较五年制和七年制教材要体现“更高”、“更新”、“更深”的特点；在教材的“三基”（基础理论、基本知识、基本技能）、“五性”（思想性、科学性、先进性、启发性、适用性）方面要更强调启发性，以培养善于思考、勇于探索、敢于创新的临床型和科研型人才。与以课程教育为主的本科学历教育不同，研究生学历教育是课题教育，研究生可根据自己的课题方向选择性地研修相关课程。这就要求我们除了考虑整套教材的一定系统性和交叉内容外，还要指出每种课题中有争论的问题，以及其前沿和发展的方向，以启发研究生在学习中的兴趣，甚至产生科学灵感。

这次编写的 19 种为第一批研究生规划教材，今后将陆续编辑出版，以供广大读者使用。

第一批研究生教材目录

1. 医学科学技术哲学	主 编	冯显威
2. 医学计算机实用教程	主 审	王行言
	主 编	童隆正
3. 医学统计学	主 编	孙振球
4. 临床流行病学	主 审	李立明
	主 编	黄悦勤
5. 医学科研方法学	主 编	梁万年
6. 医学分子生物学	主 审	刘德培
	主 编	查锡良
7. 医学分子生物学实验技术	主 编	药立波
8. 医学细胞分子生物学	主 编	宋今丹
9. 组织和细胞培养技术	主 编	章静波
10. 分子病理学	主 编	李玉林
11. 组织病理技术	主 审	王伯沅
	主 编	李甘地
12. 医学遗传学	主 编	夏家辉
13. 神经生物学	主 编	鞠 躬
14. 分子病毒学	主 编	黄文林
15. 基础与临床药理学	主 编	姚明辉
16. 实验核医学	主 编	张永学
17. 肿瘤学（第二版）	主 编	曾益新
18. 外科学——前沿与争论	主 编	邹声泉 龚建平
19. 外科常用实验方法及动物模型的建立	主 编	陈孝平

前 言

自 1865 年孟德尔采用不同品种的豌豆杂交实验证明遗传基因的分离定律与自由组合定律以后,经过近 100 年的研究历程,才于上世纪中被广大临床医学工作者所认识,并应用于人类疾病的诊断、预防与治疗,从而产生了“医学遗传学”这一崭新的学科。现在遗传基因学说不但已深入到临床各个学科,而且已被人们普遍认识到在常规条件下,一个人的生、老、病、死与基因的组合密切相关……。

据此,作为一个当代从事人类与医学专业的硕士生、博士生如果对医学遗传学没有较为系统而深入的研讨,可以肯定他将看不懂现代医学的进展,必将被现代生命科学、医学发展的潮流所淘汰。

为此,本书以绪论、染色体、基因、人类基因组计划、蛋白质组、染色体病、单基因病、多基因病、线粒体遗传病、肿瘤遗传学、遗传病的诊断和预防、基因治疗、生物信息学技术资源及其利用等为题对“医学遗传学”的基础理论、基本知识、基本技能作了阐明,并对其前沿和发展方向问题作了论述。

所有参加编写的作者都是在相关领域从事研究和教学工作的有较深造诣的人员,他们所撰写的章节都是他们长期研究的领域,同时在编写中尽可能收集了本领域的最新进展。不少信息是国际一流杂志、互联网上信息资源库的实时动态。故本书不仅可作为医学研究生的教材,也是从事医学遗传学研究、教学和临床医生值得一读的工具书。

夏家辉

2003年 12 月 28 日

目 录

第一章 绪论	1
第一节 学科发展基础	1
第二节 展望	4
第二章 染色体	6
第一节 染色体数目与结构	6
第二节 性染色体	10
第三节 染色体复制和传递	13
第四节 染色体技术和命名	19
第三章 基因	25
第一节 基因的本质	25
第二节 基因的概念	31
第三节 基因表达及调控	41
第四节 DNA 的损伤与修复	60
第四章 人类基因组计划	66
第一节 人类基因组计划产生的科学背景	66
第二节 人类基因组作图、测序及相关技术	71
第三节 人类基因组计划的进程及主要成就	89
第四节 人类基因组计划的延伸	94
第五节 人类基因组研究成果衍生的科学问题	102
第五章 蛋白质组	110
第一节 蛋白质的结构和功能	110
第二节 蛋白质合成的调节	114
第三节 细胞之间的信号传递	119
第四节 基因功能研究方法	124
第五节 细胞周期	128
第六节 细胞凋亡	132
第六章 染色体病	137
第一节 染色体病研究史	137

第二节	染色体病的类型	137
第三节	染色体异常的遗传基础	142
第四节	染色体病的致病机理	170
第七章	单基因病	188
第一节	概述	188
第二节	遗传方式	188
第三节	基因定位	198
第四节	基因克隆	204
第五节	单基因病的发病机理	208
第六节	单基因病的基因克隆	211
第八章	多基因病	286
第一节	概述	286
第二节	多基因病易感基因定位策略	290
第三节	连锁分析	292
第四节	关联研究	301
第五节	表型和中间表型	307
第六节	多基因病研究趋势	308
第九章	线粒体遗传病	311
第一节	线粒体 DNA 的结构与遗传特征	311
第二节	线粒体基因突变与线粒体基因病	314
第三节	核 DNA 变异引起的线粒体疾病	319
第十章	肿瘤遗传学	323
第一节	染色体不稳定综合征与遗传性肿瘤综合征	324
第二节	染色体异常与肿瘤	328
第三节	癌基因	330
第四节	肿瘤抑制基因	334
第五节	肿瘤发生的遗传学说	338
第六节	肿瘤的侵袭与转移	340
第七节	肿瘤基因组解剖计划	343
第十一章	遗传病的诊断和预防	348
第一节	遗传病的诊断技术	348
第二节	产前诊断技术	359
第三节	遗传筛查和遗传咨询	367
第四节	医学遗传学的伦理学思考	372

第十二章	基因治疗	376
第一节	概述	376
第二节	基因转移系统	382
第三节	单基因遗传病的基因治疗	393
第四节	多基因遗传病—糖尿病的基因治疗	400
第五节	体细胞遗传病—恶性肿瘤的基因治疗	403
第六节	获得性遗传病—HIV 感染的基因治疗	406
第十三章	生物信息学技术资源及其利用	415
第一节	计算机技术发展简况	415
第二节	Internet 简介	417
第三节	生物信息学概况	419
第四节	人类基因组研究的信息资源	420
第五节	网络期刊	426
第六节	人类基因组研究信息资源的利用	426
第七节	生物信息与蛋白质组学	427
第八节	生物信息与生物芯片	431
第九节	目前生物信息技术与资源的不足	432
中英名词对照		434

第一章 绪 论

基因及其载体染色体的研究是上世纪以来生命科学与技术发展的核心，自从 1865 年 Mendel 提出颗粒遗传学说以来，在这 137 年期间，人们一直在努力寻找基因、定位基因、分离基因、认识基因、操作基因和开发基因，由此形成了一系列的学科，在学科发展的基础上将形成 21 世纪最前沿而又最宏大的生物技术产业。

回顾这一百多年生命科学发展的历程，在基因的研究中最核心的研究对象是人，即查明人类染色体的数目，定位和克隆决定人体各种性状的基因、定位和克隆与人类疾病相关的致病基因以及能够用于疾病治疗与预防的功能基因，深入开展了基因功能组学的研究，形成了基因研究与医学相结合的一个崭新学科——医学遗传学。

在这一生命科学的核心领域中，由于研究成果突出，近 50 年来共获得 18 项诺贝尔生理学奖，11 项诺贝尔化学奖。并在此基础上带动了医学遗传学的发展，即基因诊断、基因药物及基因治疗，开启了一个崭新的基因医学时代。

第一节 学科发展基础

1865 年，Mendel 遗传物质的颗粒特性的发现，是遗传学发展史上的里程碑，被誉为遗传学研究的起点。但他的这一重大发现，也许超出了时代所能理解的极限，其重要意义并没有被人们所认识；同时，由于当时对染色体及减数分裂的现象未被发现，使 Mendel 本人也不能为其理论作出进一步的解释，以至“融合遗传”理论一直居统治地位。1882 年，Flemming 在人类的肿瘤细胞中发现了染色体的存在。随后，有丝分裂和减数分裂现象也相继被发现。这些重大突破，使 Weismann 等推想到染色体可能就是遗传物质的载体。1900 年，Mendel 定律被重新发现。Sutton 和 Boveri 等于 1903 年正式提出了遗传的染色体理论，认为遗传物质存在于染色体上。这一理论经 Morgan 等的连锁与交换分析，将某一特定的基因与某一特定的染色体直接联系起来而得到了证实，并由此创立了遗传的基因理论。Mendel 和 Morgan 等工作，使遗传学从过去的猜想、假设，发展成为一门严谨的实验科学。他们以新的实验结果为基础，创建新理论、建立新学说，而使新的理论和学说客观地反映事物内在的本质的联系，是科学发展史上成功的范例。遗传染色体理论确立了遗传物质在染色体上。染色体主要由 DNA 和蛋白质两大物质构成。确定 DNA 是遗传物质的实验是由 Griffith 和 Avery 等于 1928 年和 1944 年逐步完成的。他们通过光滑型和粗糙型肺炎双球菌的转化试验证明只有 DNA 能将粗糙型肺炎双球菌转化为光滑型；而蛋白质不具备这种转化功能，从而证明遗传物质是 DNA 而不是蛋白质。随后的一系列工作均证明在有 DNA 的生物中，DNA 是遗传物质；在没有 DNA 而只有 RNA 的生物中，如某些 RNA 病毒，RNA 是遗传物质。

1953年, Watson 和 Crick 提出的 DNA 双股螺旋结构的分子模型, 以及随后遗传密码的破译, 使人们认识到决定生物性状的遗传信息存在于 DNA 的碱基顺序之中; 基因即为一段具有遗传学功能的 DNA 序列。DNA 由 A/T 和 G/C 互补, 方向相反的两股多核苷酸链构成。

此后, 人类染色体技术、重组 DNA 技术与理论得以迅速的发展。在人类细胞遗传学研究方面, 1952 年美籍华人徐道觉发现染色体低渗制片技术, 1956 年美籍华人蒋有兴等查明人类染色体数目为 46 条即 46, XX 或 XY, 相继于 1960 年丹佛、1963 年伦敦、1966 年芝加哥召开了 3 次人类细胞遗传学国际会议, 制定并不断修改了国际命名体制, 1959 年 Lejennes 等证明了先天愚型为 21 号染色体三体型患者, 但由于方法学上的局限性, 在 1970 年以前的 10 年间人们仅记载了数种染色体病, 自 Caspersson 等于 1971 年发表第一张人类染色体显带照片, 以及 1971 年巴黎人类显带染色体国际命名会议以后, 至今已发现染色体结构畸变的类型近万种, 染色体综合征 100 余个。随着羊水穿刺等技术的建立, 染色体病的诊断与产前诊断得以迅速发展, 20 世纪 70 年代中形成了“临床细胞遗传学”这一崭新的学科。在分子遗传学及人类基因组学研究方面, 1958 年 Matthew Meselson 和 Franklin Stahl 证明 DNA 的复制涉及双螺旋互补链的分离; A.Kornberg 分离出 DNA 聚合酶 I, 这是第一个可在试管中合成 DNA 的酶。1960 年 S. Ochoa 发现 RNA 聚合酶, 该酶沿 DNA 单链的表面合成 RNA 链; S.Ochoa 发现信使 RNA, 并证实其带有编码蛋白质中氨基酸序列的信息。1961 年 Marshall W. Nirenberg 用人工合成的信使 RNA 分子 (多聚尿嘧啶核苷酸) 破译出第一个遗传密码。1965 年揭示在细菌中抗生素抗性基因常常携带于称为“质粒”的额外小染色体上。1966 年 Marshall W. Nirenberg 和 Har Gobind Khorana 共同确定了全部的遗传密码, 1968 年他们两人和转运 RNA 的发现者 Robert W. Holley 共享该年度的诺贝尔生理学或医学奖。1967 年分离出可将 DNA 链连接在一起的限制性内切酶。1970 年 Werner Arber 和 Hamilton O. Smith 分离出第一个可在特定位点切割 DNA 分子的限制性内切酶, 1971 年 D. Nathans 用这种酶来切割一种猴病毒的环状 DNA, 进行遗传学研究, 1978 年他们三人共享该年度诺贝尔生理学或医学奖。1972 年 Paul Berg 在美国斯坦福大学将 DNA 连接酶用于连接由限制性内切酶切开的 DNA 片段, 构建了第一个重组 DNA 分子。1973 年由于发现当外源 DNA 片段插入质粒 DNA 后, 能产生嵌合质粒, 当嵌合质粒重新导入大肠杆菌后, 仍具有功能, 即人们已能将某一基因克隆到细菌中并有效表达; 据此, 公众对重组 DNA 操作有可能产生具有潜在危险的新微生物产生了担心, 因而部分科学家于 1974 年发出了“世界各国暂停某些类型的重组 DNA 实验”的呼吁; 随后, 英政府发布公告, 要求在特定的实验室内谨慎进行重组 DNA 研究; 在美国加利福尼亚 Asilomar 地方召开的国际会议, 强烈要求制定关于控制重组 DNA 实验的操作守则, 同时呼吁发展那些不能逃逸出实验室的安全的细菌和质粒。1975 年美国国立卫生研究院 (NIH) 公布第一个重组 DNA 实验操作守则, 对许多类型的重组 DNA 实验作了限制; 随着公众对条例可能不起作用的忧虑加深, 《纽约时代杂志》上刊登文章呼吁禁止对重组 DNA 研究授予诺贝尔奖。1977 年美国成立了第一个专门利用重组 DNA 技术制造医学上重要药物的遗传工程公司 (Genentech), 同年, 科学家构建了第一个含有哺乳动物 DNA 的重组 DNA 分子, 发明了对大片段 DNA 进行快速序列分析的方法; Richard J. Roberts 和 Phillip A. Sharp 发现断裂 (split) 基因, 并于 1993 年获得诺贝尔生理学或医学奖。1978 年 Werner Arber、Hamilton O. Smith 和 D.Nathans 因在限制性内

切酶的发现与应用方面的贡献而授予诺贝尔生理学或医学奖；同年，生长激素抑制素 (somatostatin) 成为第一个用重组 DNA 技术产生的人类激素。1979 年 NIH 操作守则普遍放松。允许利用重组 DNA 技术研究病毒 DNA；同年，科学家用恶性生长细胞分离的 DNA 转染小鼠培养细胞株，以便分析恶性生长细胞中的癌基因。1980 年筹建了用重组 DNA 技术生产胰岛素的第一家工厂；同年，Paul Berg 因于 1972 年构建了第一个重组 DNA 分子，Walter Gilbert、Frederick Sanger 因发明了对 DNA 序列的快速而准确的测序方法共同获得了诺贝尔化学奖。1981 年第一个重组 DNA 公司 (Genentech) 出售通用股票，纽约华尔街金融中心估价，其总额在 2 亿美元以上；同年，NIH 取消了操作守则中对使用大肠杆菌和酵母实验菌株作为宿主繁殖重组 DNA 分子的基因克隆实验的限制；镰状细胞贫血症成为第一个可直接在基因水平上进行产前诊断的遗传疾病，其方法是利用限制性内切酶对 DNA 进行分析；第一只转基因小鼠和转基因果蝇诞生。1982 年 NIH 进一步全面放宽操作守则的限制；同年，科学家将克隆的大鼠生长激素基因连接到可诱导的小鼠基因控制信号上，注入小鼠的受精卵，然后植入受体母鼠体内。所产生的“超级小鼠”，体重为正常小鼠的两倍；用重组 DNA 方法生产的人胰岛素以“Humulin”的商品名进入市场；发现一种掺入烟草植物细胞的外源基因简单地以孟德尔规律通过花粉和卵细胞进行遗传；从人类膀胱癌细胞分离出癌基因，并在大肠杆菌上进行克隆，发现人膀胱癌基因与其正常的基因仅存在单一碱基的变化，从而导致蛋白质产物中一个氨基酸的变化。1983 年科学家发表了 λ 噬菌体的 DNA 全序列，共有 48502 碱基对；McClintock 由于在 20 世纪 50 年代提出并发现了可移动的遗传因子而获得诺贝尔生理学或医学奖；同年，科学家获得了第一例转基因植物；Kohler、Milstein 和 Jerne 由于发展了单克隆抗体技术，完善了极微量蛋白质的检测技术而分享了诺贝尔生理学或医学奖；斯坦福大学获得了关于重组 DNA 的专利。1985 年第一批转基因家畜出生。1988 年 J.D.Watson 出任“人类基因组计划”首席科学家。1989 年 Altman 和 Cech 由于发现某些 RNA 具有酶的功能（称为核酶）而共享诺贝尔化学奖；同时，DuPont 公司获得转肿瘤基因小鼠——“Oncomouse”。1990 年首例基因治疗临床实验在美国实施；同年，美国能源部 (DOE) 和美国国立卫生研究院 (NIH) 牵头启动人类基因组计划。1993 年 Mullis 由于发明 PCR 技术而与第一个设计基因定点突变的 Smith 共享诺贝尔化学奖。1994 年 Gilman 和 Rodbell 由于发现了 G 蛋白在细胞内信息传导中的作用而分享了诺贝尔生理学或医学奖；同年，第一批基因工程西红柿在美国上市；微生物基因组计划启动；绘制完详细的人类基因组遗传图谱，密度达到平均每 700kb 一个 Marker。1995 年 Nature 汇集发表了人类基因组全物理图，以及 3 号、16 号和 22 号人染色体的高密度物理图；同年，Lewis、Nüssle-Volhard 和 Wieschaus 由于在 20 世纪 40~70 年代先后独立鉴定了控制果蝇体节发育基因而分享诺贝尔生理医学奖。1996 年完成了酵母基因组 (1.25 (107bp) 全序列测定。1997 年 E.coli 基因组测序完成，共有 4600000 碱基对，大约 4000 个基因；同年，英国爱丁堡罗斯林研究所获得克隆羊——多利。1998 年 11 月 James Thompson 和 John Gearhart 领导的两个研究小组分别成功地培养出了人胚胎干细胞；同年，人类基因组计划宣布将加快研究进度，2001 年实现基因草图的绘制。2000 年果蝇基因组 (47350000 碱基对) 测序完成。2001 年 HGP 和 Celera 分别将他们的初步分析结果发表在 2 月 15 日的 Nature 和 2 月 16 日的 Science 上。

第二节 展 望

20世纪50年代以来,随着染色体显带技术、DNA重组技术、PCR技术的发现和应用,细胞遗传学、分子遗传学进入了鼎盛发展时期,它与医学的结合相继形成了人类细胞遗传学(human cytogenetics)、人类生化遗传学(human biochemical genetics)、人类分子遗传学(human molecular genetics)、人类群体遗传学(human population genetics)、药物遗传学(pharmacogenetics)、免疫遗传学(immunogenetics)、辐射遗传学(radiation genetics)、发育遗传学(developmental genetics)、遗传毒理学(genetic toxicology)、体细胞遗传学(somatic cell genetics)、肿瘤遗传学(cancer genetics)、行为遗传学(behavior genetics)、临床遗传学(clinical genetics)等数十个医学遗传学学科。1990年以来,“国际人类基因组计划”的实施,极大地推动了人类与医学遗传学的发展。

尽管目前公布的基因组框架图还不十分完善,但还是为我们提供了一个当前最全面、最完整的基因组信息。利用日益丰富和不断完善的基因组数据,我们不仅可以确定遗传病、肿瘤、心血管病、糖尿病等危害人类生命健康最严重疾病的致病基因,寻找出个体化的防治药物和方法,对于癌症、药物依赖、基因表达、免疫、进化基因组学、膜交换、细胞骨架、细胞周期和生物钟研究以及进化生物学等众多学科都将产生难以预想的深远影响,将极大地推动人类与医学遗传学以及整个遗传学科的发展。

自1974年世界上克隆第一种Kappa轻链缺乏症的疾病基因以来,到1989年16年期间仅克隆了疾病基因106个,平均每年克隆6.6种;而从1990年人类基因组项目实施以来到2002年的12年间共克隆了1060个疾病基因,平均每年克隆88.3种。也就是说基因组计划实施以后每年克隆疾病基因的数目是以前的13倍。同时查明,人类基因组大小约为2.9~3.2Gb,不同个体之间99.99%的基因序列是相同的。整个基因组中GC碱基含量仅占38%,人的基因组中仅含3万~4万个基因,远少于原先估计的10万个基因,只是线虫或果蝇基因数量的两倍,人有而鼠没有的基因只有300个。还有大约200多个基因编码的蛋白质分子与细菌中的蛋白同源,但在酵母、线虫、果蝇及其他一些无脊椎动物中都未发现同源物。人类基因的大小差异较大,从65bp到2000kb不等,目前发现的最大编码dystrophin蛋白的基因有2400kb,虽然它含有79个外显子,但其中的编码序列只占整个基因序列的0.6%。在人类基因组中编码蛋白的序列不足5%,存在“热点”和大片“荒漠”,即在染色体上有基因成簇密集分布的区域,也有大片的区域只有Junk-DNA,分布着300多万个长片断重复序列,在非编码序列中重复序列则至少占基因组的50%。

“突变与自然选择”的“经济与适应”原则是生物生存与进化的最基本的法则。基于此,四种单核苷酸(A、T、G、C)的不同排列组合就构成了地球上数亿种生物,为什么人类具有30亿个核苷酸而编码蛋白的序列不足5%?如果这95%的核苷酸序列对生物体的生存无用,那为什么不在进化过程中淘汰?为什么人类只有3万~4万个基因却能编码出数十万种蛋白质?在进化中人是最高等生物,为什么人有、鼠没有的基因只有300个?上世纪末的研究证明,酶是蛋白质,但蛋白质以外的物质如某些RNA也具有酶的功能!DNA、RNA是遗传的物质基础,但其他物质如蛋白质是否在某种特殊情况也具有遗传的功能?受精卵已经带有致病基因,但为什么要到一定的年龄才发病?为什么同一个致病基因的同一突变在不同的

个体间发病的轻重不同？为什么同一种药物对不同的人敏感度不一？等等。为了回答这些问题，科学家正在延伸人类基因组计划，提出了所谓模式生物的基因组学、功能基因组学、结构基因组学、疾病基因组学、肿瘤基因组学、药物基因组学、人类基因组多样性计划和人类基因组信息学，与时俱进的科学研究与发现有可能进一步揭示生命的奥秘，人类研究自身的这一认识过程将不断深化，但永无止境。

参 考 文 献

1. 夏家辉等.染色体病.北京: 科学出版社, 1989.
2. 夏家辉主编.医学遗传学讲座.长沙: 湖南科学技术出版社, 1998.
3. J.D.Watson and F.H.C. crick. Molecular Structure of Nucleic acids. Nature. 1953, 171:737-738
4. Adams MD, et al. The sequence of the human genome. Science, 2001, 291:1304-1351.
5. International Human Genome Sequencing Consortium. Initial Sequencing and analysis of the human genome. Nature, 2001, 409:860-921.
6. Xia JH, et al. Mutations in the gene encoding gap junction protein beta-3 associated with autosomal dominant hearing impairment. Nature Genetics, 1998, 18:370-373.
7. <http://www.ncbi.nlm.nih.gov/omim/>

(夏家辉)

第二章 染色体

第一节 染色体数目与结构

一、人类的正常核型

染色体 (chromosome) 位于细胞核中呈线状结构, 是遗传信息的载体。染色体一词来自希腊文 chroma (颜色) 和 soma (体), 研究染色体和细胞分裂是细胞遗传学的主要内容。染色质 (chromatin) 和染色体是同一种物质的不同存在形式, 染色质是指间期细胞核内遗传物质存在的形式; 染色体是细胞分裂时遗传物质存在的特定形式, 是间期细胞染色质多级螺旋折叠的结果。在 1950 年以前, 人们认为人类染色体数目为 48 条, 人类的性别和果蝇一样, 是由常染色体和性染色体数目的比率决定的, 这两个结论错误的主要原因是染色体技术的限制, 当时只能观察分裂旺盛的细胞染色体, 如肿瘤细胞、男性性腺细胞、培养生长细胞和骨髓标本, 而且染色体标本的质量比较差。1956 年, Tjio 和 Leven 证实了培养的人胚胎成纤维细胞有 46 条染色体, 而不是 48 条; 由于确定了人类染色体数目, 随后的三年中, 诊断出染色体数目异常的三个主要疾病, 即 Down 综合征, Turner 综合征和 Klinefelter 综合征; 世界各地不断开展了人类染色体的研究, 发现的染色体变异日益增多, 为了使人类染色体的识别和分析标准化以及国际间的交流方便, 在 1960 年 (丹佛)、1963 年 (伦敦)、1966 年 (芝加哥) 相继召开了三次人类细胞遗传学国际会议, 讨论并制定国际命名体制。

人类体细胞为二倍体 (diploid), 即 $2n$, 含 46 条 (23 对) 染色体, 其中 1~22 对染色体为男女共有, 称常染色体 (autosomal chromosome), 另一对的组成与性别有关, 称性染色体 (sex chromosome), 女性为 XX, 男性为 XY。配子为单倍体 (haploid), 即 n , 含 23 条染色体。一个体细胞中特定数目和大小的染色体称核型 (karyotype)。人类的正常核型是 46, XX (XY)。

二、染色体的结构特征

(一) 核小体和染色质纤维

人类基因组 DNA 凝缩在细胞核中构成了染色质, 染色质除了 DNA 外还有大量的组蛋白, 可分为五种 H1、H2A、H2B、H3 和 H4; 其中的四种组蛋白形成八聚体 (H2A、H2B、H3、H4 各两个分子), 大约 146bp 的 DNA 盘绕在八聚体构成的核心结构外面 1.75 圈, 形成了一个核小体 (nucleosome), 这些盘绕的 DNA 不易被核酸酶消化; 核小体是染色质的基本

结构单位，其直径约 10nm；每个核小体串珠似地连接在一起，核小体之间的 DNA 约 15~55bp 不等，因种属而异。研究发现 H1 组蛋白为一组密切相关的蛋白质，其数量相当于核心组蛋白的一半，易从染色质中抽提出来，所有的 H1 组蛋白被除去后，也不会影响到核小体的结构，表明了 H1 组蛋白是位于核心组蛋白之外。每 6 个核小体以 H1 组蛋白为中心形成一个螺旋，时而紧时而松，处于动态，其外径约 30nm，这种螺旋管结构形成染色质纤维 (chromatin fiber)，进一步螺旋化和折叠，最终形成了有丝分裂中期的染色体 (图 2-1)。人类每条染色体都是由一条 DNA 分子盘绕组蛋白高度螺旋化构成的；每条 DNA 分子平均长度为 10cm，由 $(2\sim3) \times 10^8\text{bp}$ 组成，新复制的 DNA 很快和组蛋白组装成核小体。组蛋白是一类碱性蛋白质，核小体的每个组蛋白都有 20~40 个氨基酸残基从核心伸出，使其 N 端带正电荷的赖氨酸基团可以和与其自身盘绕的 DNA 磷酸相互作用，也可以和其附近 DNA 以及核小体之间的 DNA 磷酸相互作用；如果赖氨酸基团乙酰化，就会终止与 DNA 的这种作用。组蛋白乙酰化越多，染色质纤维丝的致密程度和高度螺旋化就越低，DNA 转录活性越强。在染色质中还有非组蛋白成分，为染色体提供支架作用，也和 DNA 转录有关。

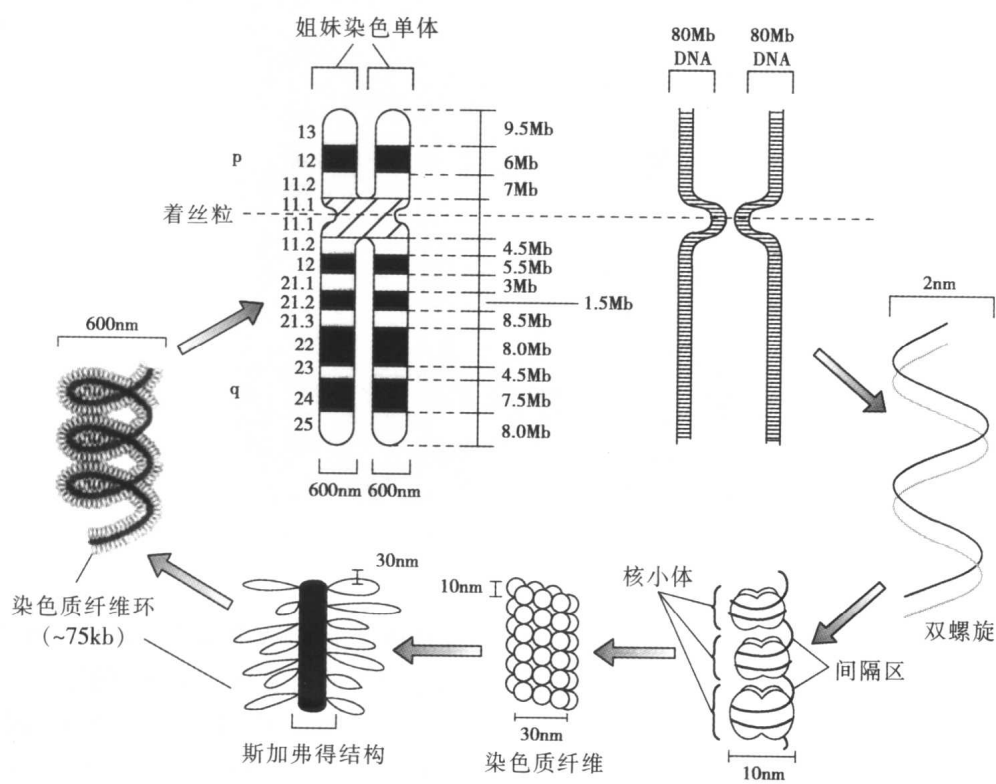


图 2-1 从 DNA 双螺旋到中期染色体的结构

(二) 着丝粒

在显微镜下可见正常染色体有一狭窄部分即着丝粒 (centromere)，在染色体的形态上表现为一个缢痕 (constriction) (图 2-2)。着丝粒连着两个姐妹染色单体，保证了细胞分裂时姐妹染色单体的分离；因此，着丝粒是细胞分裂时染色体分离的一种装置。如果没有着丝粒的染色体片段，不能被纺锤体附着，也就不能进入子细胞核。在细胞分裂前期，着丝粒上有一对动粒 (kinetochore) 分别附着在姐妹染色单体上，可以和纺锤体的微管相连，在分裂后

期微管拉着两条染色单体向细胞两极运动，因此动粒起着核心作用，控制着微管的装配和染色体的移动。

着丝粒的结构和功能是由特异性 DNA 序列决定的。在简单的真核细胞中，这个序列非常短，例如酵母的着丝粒元件大约 110bp，是由一个富含 AT 的 80~90bp 中心序列和两个高度保守的 8bp 和 11bp 侧翼序列组成；着丝粒可以相互交换，即某酵母细胞的着丝粒元件可以被另一酵母细胞所代替，但不影响酵母细胞的分裂。在哺乳动物中，着丝粒是由几十万个碱基对组成，包括 DNA 重复序列、染色体特异序列和非特异序列。提取的基因组 DNA 经密度梯度离心后，着丝粒 DNA 在基因组 DNA 主峰旁形成“卫星”状，又称卫星 DNA，可进一步分为 α 、 β 、 γ 等类型。人类染色体每条都有 α 卫星，由 171bp 的重复序列组成，占每条染色体 DNA 的 3%~5%，构成着丝粒异染色质的大部分。 α 卫星的 DNA 重复单位有一个特异着丝粒蛋白结合位点，可能对着丝粒功能起重要作用。 β 、 γ 等其他类型卫星 DNA 的分布因不同染色体而异。

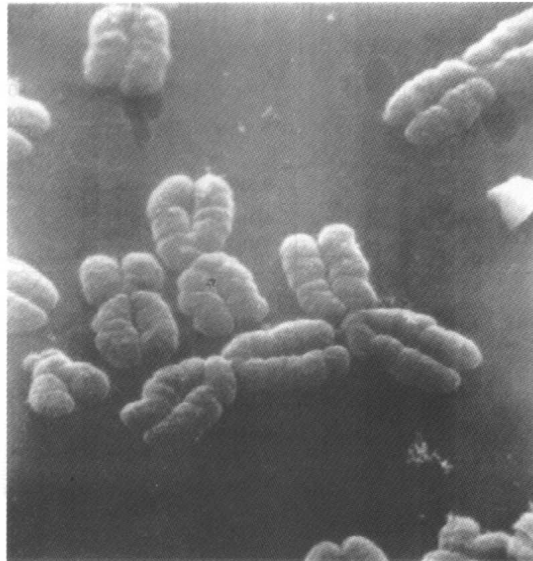


图 2-2 电子显微镜下的染色体及其着丝粒

(三) 端粒

端粒 (telomere) 是一个位于染色体末端的特殊结构，由 DNA 和蛋白质组成。在 20 世纪 30 年代，Müller H T 和 McClintock 发现真核细胞染色体的两端有一种特殊的组成，就像一顶“帽子”套在染色体的端部，使染色体保持其个体性 (individuality)，染色体之间不会出现端-端融合，也不会端部松散而导致解体。Müller 用希腊文 telos (末端) 和 meros (部分) 组成了 telomere (端粒) 命名了这种染色体端部的特殊组成。端粒 DNA 为一串联重复序列，双链中的一条 3' 端为富含 TG 的序列，互补链为富含 CA 的序列；这个重复序列单位在进化中高度保守，如草履虫的 TTGGGG、锥虫的 TAGGG、拟南芥的 TTTAGGG、人类的 TTAGGG。

染色体端粒的复制是通过端粒酶 (telomerase) 从延伸后随链亲链开始的 (图 2-3)。端粒酶由 RNA 和蛋白质组成，端粒酶 RNA 成分中的一段短序列作为模板，不断延长后随链亲链 3' 端 DNA，而延长的后随链亲链又为后随链的合成提供模板，最终使端粒 DNA 序列 3' 端