

研究生创新教育系列丛书

医学分子遗传学

(第四版)

陈金中 汪旭 薛京伦 主编



科学出版社

医学分子遗传学

——基础与临床——

第三版

人民卫生出版社

研究生创新教育系列丛书

医学分子遗传学（第四版）

陈金中 汪 旭 薛京伦 主编

科学出版社

北 京

内 容 简 介

医学分子遗传学于1990年作为遗传学系列丛书的一部分出版。尽管系列丛书的多数未曾再版,但医学分子遗传学到现在已经是第四版了。它不仅仅是一部遗传学参考书,也被多个院校作为本科生与研究生教材,更新再版也成为一种义务。正是考虑到这一情况,本书依然把知识的完整性作为一个重要的考量;同时本书也力求反映遗传学从经典遗传学、分子遗传学到遗传转化医学的变化轨迹与趋势。本书也首次邀请海外学者参与编写,力求反映世界遗传学发展特点,希望可以使读者一定程度了解国外遗传学的发展。

我们一如既往地希望本书可以成为一本有关专业人员案头的参考书,也可以作为医学与生物学相关专业的教材和教学参考书。

图书在版编目(CIP)数据

医学分子遗传学 / 陈金中, 汪旭, 薛京伦主编. — 4 版. — 北京: 科学出版社, 2012
(研究生创新教育系列丛书)
ISBN 978-7-03-035830-1

I. ①医… II. ①陈…②汪…③薛… III. ①医学遗传学—分子遗传学—研究生—教材 IV. ①Q75

中国版本图书馆 CIP 数据核字(2012)第 249183 号

责任编辑: 李 悦 岳漫宇 刘 晶 / 责任校对: 郑金红
责任印制: 钱玉芬 / 封面设计: 陈 敬

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>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2013年1月第 四 版 开本: 787×1092 1/16

2013年1月第一次印刷 印张: 20 3/4 插页: 1

字数: 461 000

定价: 68.00 元

(如有印装质量问题, 我社负责调换)

前 言

无论最古老原始的微生物，还是万物之灵的人类，都拥有遗传繁衍这样的生命基本特征。核酸——DNA 和 RNA，是能够自我复制的遗传信息分子，是生命自我复制的真谛所在。研究这些遗传信息分子的结构、传递、功能及其调控规律，探索不同个体遗传差异与疾病易感性的关联，并将相关成果应用到医学研究与临床之中，是医学分子遗传学的重要内容，也是逾越遗传学基础研究和临床应用之间的屏障，将医学生物学基础研究成果有效地转化为临床药物、生物材料与方法的重要环节。

医学分子遗传学是遗传学-分子生物学-医学等多学科相互交融的交叉领域。本书在原有第一至第三版的基础上，进一步融入了我们和同仁的最新研究成果，引入了国内外研究的最新热点问题。本书首先介绍了人类基因组计划、基因表达调控、发育与疾病发生中的表观遗传学、肿瘤遗传学等基础内容和分子生物学的基础技术，并应近年兴起的转化医学所反映的生物医学科研与临床应用相结合的社会需求，力求体现医学遗传学科研服务于生物医学、医学诊断与治疗、营养与健康、司法鉴定和药物发现等主流方法。

根据潜在的读者范围与知识结构，本书前 4 章简要介绍了医学分子遗传学的基本理论与技术框架；基于人类基因组计划的发展沿革和后基因组时代人们对基因组功能及其调控本质探索的渴望，第五和第六章简要阐明人类基因组结构、人类基因表达与调控的一些基本原理；鉴于遗传物质的表观修饰在发育与疾病发生中的作用有可能推动医学的革命性发展，Beate Brand-Saberi 等 4 位德国教授在第七章严谨生动地介绍了表观遗传学在发育和疾病中的作用与原理。郭凌晨博士在第八章中大幅度地更新了遗传与肿瘤发生中的信号途径变化；基因诊断和基因治疗为本书前三版的重点内容，新版对原文作了少量更新并继续保留。陈金中副教授在本书的构架、组织中发挥了核心作用，完成多个章节的撰写与更新；李成涛研究员贡献了法医分子遗传学一章，对完善之前版本的知识体系大有裨益。汪旭教授和澳大利亚 Commonwealth Scientific and Industrial Research Organization (CSIRO) 的 Michael Fenech 教授全面更新了营养基因组学的内容，分别完成了基因组稳定性与环境、公共健康及个性化基因组学的撰写，他们从营养和基因科学到公共政策的角度阐明了基因组时代的健康概念与策略。感谢王明伟教授和周彩红博士为本书贡献了基因组学与药物创新一章，向读者展示了医学遗传学基础研究转化为医学的动力和端倪。

本书所涉及的领域发展迅速、内容宽泛，对文献的引用难免挂一漏万；由于水平有限，撰写错误在所难免，欢迎读者批评指正。

我们对联合基因科技（集团）有限公司（United Gene Holdings Co., Ltd.）为本书出版提供部分经费支持表示由衷的感谢！

薛京伦

于复旦大学 2012 年 4 月 8 日

第三版前言

医学分子遗传学是以遗传学理论为基础，医学遗传学为背景，借助分子生物学技术，从分子水平揭示疾病与遗传因素的关系，从而探索新的疾病诊疗技术和防治途径的学科。随着人类基因组结构与功能研究的深入和医学的飞速发展，人们不断发现，单基因遗传病、多基因遗传病、线粒体疾病、肿瘤、衰老、病毒性疾病、环境易感性疾病等都和遗传因素有千丝万缕的联系。因此，医学分子遗传学是一门涉及遗传学、医学、环境科学、生态学、分子生物学、生物信息学、人类学、伦理学和社会学等多学科的领域，医学分子遗传学的发展必将有力地推动人类健康和医学事业的进步。

这本“医学分子遗传学”是本课题组出版的同名专著的第二本，它凝聚了我们自 20 世纪 80 年代以来的研究和教学经历。从体细胞基因定位、人类遗传病基因治疗的基础和临床试验到目前向基因治疗载体安全性与有效性的挑战，不仅使我们整个课题组处于基因诊断和基因治疗研究的前沿，同时使本书独具特色，汇集了撰写者丰富的科研、教学实践与成就。

本书以遗传性疾病和疾病的遗传性因素为主线，从单基因疾病、多基因疾病、肿瘤、病毒性疾病及体细胞遗传病等多个角度，应用基因组结构和功能的知识，深入揭示基因突变和疾病发生的内在联系，同时阐述环境对遗传物质的作用，前 7 章着重介绍了人类基因组的结构特征、单基因和多基因疾病及肿瘤的分子遗传学，充分介绍了本学科的基础知识及新进展；第八到第十一章，介绍了医学分子遗传学中若干新的领域和内容，其中包括表观遗传学、线粒体医学、环境基因组研究与环境相关疾病遗传因素的探索、转基因动物作为遗传疾病的模型及在医学研究中的地位等；第十二和第十三章重点介绍了与本课题组 20 多年来的工作密切相关的基因诊断与基因治疗的基本理论、基本技术和基本策略，旨在为进一步实现疾病的预防和分子水平的基因诊断，最终实现疾病基因治疗的目的。

本书可用作生物和医学类本科生及研究生的教材，以及高校及科研院所青年科技工作者和管理者的参考书。

由于水平有限，错误在所难免，欢迎读者批评指正。

谨将此书献给复旦大学建校 100 周年！

薛京伦

于复旦大学 jlxue@fudan. ac. cn

2005 年 3 月 27 日

第二版序

1953年，J. D. Watson 和 F. H. C. Crick 提出了 DNA 双螺旋结构模型，这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。目前，遗传学科已经成为生命科学领域中最活跃和最引人注目的一个带头学科。

30 余年来，现代遗传学无论在理论研究还是在生产应用方面，都取得了一系列重大突破，尤其是 70 年代初重组 DNA 技术的建立和发展，为遗传学发展走上产业化道路奠定了基础。遗传工程的兴起，使人类有可能按照自己的意愿和需要，来直接操纵遗传物质，有目的地改造各种生物的遗传组成及至建立新的遗传特性。遗传工程已经并将继续对人们的日常生活产生巨大的影响，对人类社会发挥重要作用。今天，以遗传工程为主体的生物工程技术已同微电子技术、能源技术一起，成为关系到人类生存和社会进步的世界高技术领域的重要支柱。

众所周知，遗传学研究在我国曾几度波折，历经沧桑，有过一段坎坷曲折的历程。随着我国“四化”建设的步伐，遗传学研究逐渐走上健康发展的道路，改革开放政策的实施，更推动着我国的遗传学研究走向世界。目前，遗传学研究在我国已得到广泛开展，某些领域还达到了国际先进水平。然而，面对世界新技术革命的潮流，我们必须清醒地认识到我国科学技术发展中还有薄弱环节；而要改变这一切，跻身于世界科技强国之列，首要的战略措施就是必须抓紧人才的培养，这在任何国家都是一样的。造就和培养一大批具有高水平的科学家是我国科学技术实现现代化的体现和保证。因此，作为生命科学带头学科的遗传学，就更要求造就一大批一流的遗传学家，为现代遗传学的发展、为遗传工程和生物技术在我国各个领域开展服务，这是一项有战略意义的重要措施。

就是在这种形势下，复旦大学出版社为了加速我国遗传学人才的培养，适应国内的遗传学教学的需要，特约请了复旦大学遗传学研究所和国内的遗传学专家撰写《遗传学丛书》各分册。在这些论著中，作者不仅详细介绍了遗传学各个分支领域的基本理论和基础知识，还充分反映了最新的研究成果和进展，力求内容新颖、资料丰富、文笔流畅。这套丛书对遗传学专业的大学生、研究生和正在从事遗传学及其相关领域的教学、科研工作者无疑是一套水平较高的专业参考书。我相信，这套丛书的出版必将对我国蓬勃发展的遗传学事业起到积极的推动和促进作用，为我国科学技术现代化的早日实现做出应有的贡献。

谈家桢

1988 年 1 月于上海复旦大学遗传学研究所

第二版前言

医学分子遗传学是近年来在医学遗传学基础上发展起来的一门现代新兴学科，它运用分子生物学技术，从 DNA 水平、RNA 水平及蛋白质水平对遗传性疾病或疾病的遗传因素进行研究，揭示基因突变与疾病发生的关系，建立在分子水平上对遗传性疾病等的诊断方法，进一步实现对遗传性疾病等的基因治疗，达到从根本上治愈遗传病的目的。

数十年来，医学的发展和进步令人眼花缭乱、目不暇接，生物科学对医学的影响最为直接和深刻。而遗传学是生物科学的基石，本世纪分子遗传学领域里的重要发现对医学的发展起了决定性的作用，医学分子遗传学已经成为遗传学和医学领域里最为活跃的学科之一。

我们从 1984 年起为遗传学专业的本科生开设了医学分子遗传学课程，并于 1988 年编写了《医学分子遗传学》教材，它在历年的教学过程中，受到了老师和同学的好评。鉴于学科发展的迅速和知识更新的加快，原有教材难以反映最新的医学分子遗传学研究进展，我们为了及时将这个研究领域的成果系统地介绍给大家，在原有教材的基础上，并结合科研实践，编写了这本书，供大家参考和使用。

全书共分 18 章，包括绪论、医学分子遗传学基础、基因的表达调控、单基因病、多基因病、肿瘤、病毒性疾病、免疫系统疾病、线粒体疾病、细胞遗传学与分子遗传学、转基因动物、环境诱变剂、医学分子遗传学研究热点、基因定位、基因克隆、基因诊断和基因治疗。为了便于理解和复习，每章后均有小结和思考题，全书的末尾列出了主要的参考文献，可供进一步查阅。教材的讲授时间为 60 学时，一学期讲完，因材施教，讲授内容可以有所取舍。

本书各章分别由卢大儒（第一、第四、第十一、第十三、第十五、第十六、第十七和第十八章）、施前（第二、第三章）、邱晓赟（第四章）、高啸波（第五、第六章）、王琪（第七章）、包赟（第八章）、张克忠（第九章）、郑冰（第十章）、戴旭民（第十一章）、胡以平（第十二章）、刑永娜（第十三、第十四章）、谈珉（第十四章）、周其南（第十四章）撰写。由于书籍出版的周期较长，而这一领域的进展又如此迅速，本书难免会遗漏一些重要内容，同时由于我们的知识水平有限，尽管尽了最大的努力，肯定还会存在不少问题和错误，欢迎批评指正。

1997 年 8 月
于复旦大学遗传学研究所

第一版前言

医学分子遗传学是近年来出现的一门新兴的边缘学科，它是遗传学的一个重要分支，是医学遗传学与现代生物学技术结合的产物。这门学科从诞生至今不到 10 年，但在这一阶段中所取得的进展，使人类和医学遗传的研究完全进入了一个崭新的阶段，对整个生命科学的研究产生了巨大的影响。这方面的资料大多零散地刊登在各种杂志上，国内外还没有一本系统的教科书。为了将这一领域的最新研究成果及时系统地介绍给大家，我们在已经开设了 4 年的体细胞遗传学和医学分子遗传学课程的基础上，编写了这本书，供大家参考和使用。

由于对遗传病发病机制的研究已进入基因的结构、表达和调控阶段，所以在前面几章中简要地介绍了有关人体基因的结构和功能方面的基础内容。在各章末尾都列出了主要的参考文献，可供进一步查阅。全书共分 15 章，包括绪论、医学遗传学基础、人体基因组的结构解剖、人体基因的表达与调控、流式细胞分类学、单基因病分子遗传学、染色体异常的细胞和分子遗传学、多基因病分子遗传学、肿瘤分子遗传学、免疫系统疾病分子遗传学、限制性片段长度多态性、基因定位、基因诊断、基因药物学和基因治疗。

在历年来的教学过程中，不少老师和学生提出了许多宝贵的意见，才使这本书以今天这样的面目出版，在此一并致以深切的谢意，并恳切希望能继续得到各位读者和同行的批评指正。

由本书出版的周期较长，而这一领域的进展又是如此快，所以我们正在把所有最新的资料都输入软盘，希望以后能以微机（IBM-PC）软盘的形式为大家提供及时而又价廉的第二版。

谨以此书献给我们敬爱的导师刘祖洞教授。

薛京伦

1988 年 8 月 5 日

于复旦大学遗传学研究所

目 录

前言

第三版前言

第二版序

第二版前言

第一版前言

第一章 生物大分子与中心法则	1
第一节 生物大分子的基本构件	1
第二节 DNA 结构与复制	2
第三节 RNA 分类、转录与转录后过程	4
第四节 蛋白质的生物合成	10
第五节 生物学中心法则	16
参考文献	17
第二章 染色体——细胞分裂中的遗传物质	19
第一节 染色体的结构与组装	19
第二节 细胞分裂与染色体运动	24
第三节 染色体异常与疾病	26
参考文献	27
第三章 世代传递中的遗传物质	28
第一节 遗传基因世代传递的物质基础	28
第二节 遗传方式	29
第三节 基因的进化	36
参考文献	39
第四章 分子遗传技术方法	40
第一节 PCR 技术	40
第二节 分子杂交	42
第三节 基因工程	44
参考文献	47
第五章 人类基因组	48
第一节 基因组计划	48
第二节 模式生物基因组	50
第三节 人类核基因组概论	51
第四节 线粒体基因组	53

参考文献	57
第六章 人类基因的表达调控	59
第一节 DNA 水平的调控	60
第二节 顺反调节	61
第三节 异构体	63
第四节 miRNA 调节基因表达	63
参考文献	65
第七章 发育与疾病中的表观遗传学：原理与技术	67
第一节 DNA 甲基化	67
第二节 核小体和组蛋白修饰	69
第三节 microRNA	75
第四节 检测表观遗传改变的技术方法	77
第五节 基因组表观遗传图谱是肿瘤发生、增殖及预后的标记	78
第八章 肿瘤分子遗传学	110
第一节 癌基因与抑癌基因	110
第二节 细胞信号传递与肿瘤的生长控制	115
参考文献	131
第九章 基因组稳定性与环境	133
第一节 环境基因组计划与环境基因组学	133
第二节 营养基因组学	139
参考文献	155
第十章 遗传工程小鼠与医学遗传学研究	159
第一节 概述	159
第二节 基因工程小鼠的遗传修饰策略	161
第三节 基因工程小鼠制备的途径及其发展	168
第四节 基因工程小鼠模型在遗传病研究中的应用	172
第五节 基因工程小鼠模型在肿瘤研究中的应用	175
参考文献	186
第十一章 疾病基因克隆与基因诊断	187
第一节 遗传作图与遗传标记	187
第二节 两点和多点作图	187
第三节 鉴定疾病基因	188
第四节 复杂疾病易患性基因的鉴定	189
第五节 分子病理学与基因检测	191
第六节 基因诊断	193
参考文献	204
第十二章 法医分子遗传学	206
第一节 法医遗传学的发展	206

第二节 短串联重复序列	207
第三节 单核苷酸多态性	218
第四节 RNA 分子	221
第五节 非人类 DNA	228
参考文献	231
第十三章 基因治疗概论	235
第一节 基因治疗的一般概念	235
第二节 基因治疗的历史事件和基本原则	236
第三节 基因治疗的理论与技术基础	238
第四节 基因治疗的现状与范例	247
第五节 基因治疗展望	255
参考文献	255
第十四章 公共健康及个体基因组学	257
第一节 暴露组、营养组和基因组的多样性	257
第二节 公共健康基因组学、营养遗传学及营养基因组学	258
第三节 DNA 损伤防范所需营养参考值	260
第四节 整合理化因素、心理社会环境和营养对基因组影响的综合分析方法	264
第十五章 基因组学与药物创新	283
第一节 新药研究的发展趋势	283
第二节 基于靶标的药物发现	291
第三节 药物基因组学	303
参考文献	310

彩版

第一章

生物大分子与中心法则

第一节 生物大分子的基本构件

生命的基本特点是由生物大分子来实现的，组成生命体的生物大分子包括核酸、蛋白质和多糖。多糖是由至少 10 个单糖组成的聚合碳水化合物。相同的单糖组成的多糖有淀粉、纤维素和糖原；以不同单糖聚合形成杂多糖，如阿拉伯胶是由戊糖和半乳糖等组成。少于 10 个残基的短糖链称为寡糖。结合了蛋白质和脂类的寡糖称为复合多糖，如糖蛋白、糖脂和蛋白聚糖。多糖可以作为生物体的骨架，而其作为动植物储藏的养分也可能具备特别的生物活性，如糖原和淀粉储藏能量、肝素抗凝血。

蛋白质是生命的物质基础，可以说没有蛋白质就没有生命。蛋白质旧称朊，故疯牛病的蛋白质病毒被命名为朊病毒。氨基酸是组成蛋白质的基本单位，氨基酸通过脱水缩合成肽链。通常蛋白质是由一条或多条多肽链组成的生物大分子。组成蛋白质的基本分子构件是 20 种氨基酸（表 1-1），在蛋白质中，某些氨基酸残基还可以因翻译后修饰而发生变化，从而对蛋白质进行激活或调控。多肽的氨基酸序列界定了蛋白质的基本潜能。多个蛋白质可以通过结合在一起形成的蛋白质复合物和折叠成一定的空间结构进一步丰富蛋白质的功能。

表 1-1 氨基酸的基本分子结构和性质

缩写	全名	中文名	支链	相对分子质量	等电点	解离常数 (羧基)	解离常数 (氨基)	R 基
G. Gly	glycine	甘氨酸	亲水性	75.07	6.06	2.35	9.78	—H
A. Ala	alanine	丙氨酸	疏水性	89.09	6.11	2.35	9.87	—CH ₃
V. Val	valine	缬氨酸	疏水性	117.15	6.00	2.39	9.74	—CH—(CH ₃) ₂
L. Leu	leucine	亮氨酸	疏水性	131.17	6.01	2.33	9.74	—CH ₂ —CH (CH ₃) ₂
I. Ile	isoleucine	异亮氨酸	疏水性	131.17	6.05	2.32	9.76	—CH (CH ₃)—CH ₂ —CH ₃
F. Phe	phenylalanine	苯丙氨酸	疏水性	165.19	5.49	2.20	9.31	—CH ₂ —C ₆ H ₅
W. Trp	tryptophan	色氨酸	疏水性	204.23	5.89	2.46	9.41	—CH ₂ —C ₈ NH ₆
Y. Tyr	tyrosine	酪氨酸	疏水性	181.19	5.64	2.20	9.21	—CH ₂ —C ₆ H ₄ —OH
D. Asp	aspartic acid	天冬氨酸	酸性	133.10	2.85	1.99	9.90	—CH ₂ —COOH
N. Asn	asparagine	天冬酰胺	亲水性	132.12	5.41	2.14	8.72	—CH ₂ —CONH ₂
E. Glu	glutamic acid	谷氨酸	酸性	147.13	3.15	2.10	9.47	—(CH ₂) ₂ —COOH

续表

缩写	全名	中文名	支链	相对分子质量	等电点	解离常数 (羧基)	解离常数 (氨基)	R 基
K. Lys	lysine	赖氨酸	碱性	146.19	9.60	2.16	9.06	$-(CH_2)_4-NH_2$
Q. Gln	glutamine	谷氨酰胺	亲水性	146.15	5.65	2.17	9.13	$-(CH_2)_2-CONH_2$
M. Met	methionine	甲硫氨酸	疏水性	149.21	5.74	2.13	9.28	$-(CH_2)-S-CH_3$
S. Ser	serine	丝氨酸	亲水性	105.09	5.68	2.19	9.21	$-CH_2-OH$
T. Thr	threonine	苏氨酸	亲水性	119.12	5.60	2.09	9.10	$-CH(CH_3)-OH$
C. Cys	cysteine	半胱氨酸	亲水性	121.16	5.05	1.92	10.70	$-CH_2-SH$
P. Pro	proline	脯氨酸	疏水性	115.13	6.30	1.95	10.64	$-C_3H_5$
H. His	histidine	组氨酸	碱性	155.16	7.60	1.80	9.33	$-CH_2-C_3N_2H_3$
R. Arg	arginine	精氨酸	碱性	174.20	10.76	1.82	8.99	$-(CH_2)_3-NH-C(=NH)-NH_2$

核酸是由多个核苷酸聚合成的生物大分子。不同的核酸，其化学组成、核苷酸排列顺序等均不相同。核酸可分为 RNA 和 DNA。DNA 是储存、复制和传递遗传信息的物质基础，RNA 在蛋白质合成过程中起着重要作用，也可以作为遗传信息的储存载体和发挥类似蛋白质的酶活性作用而广泛参与生命过程。DNA 由 6 种小分子组成：脱氧核糖、磷酸和 4 种碱基（A、G、T、C），由这些小分子组成了 4 种核苷酸，这 4 种核苷酸组成了 DNA。RNA 同样也有 6 种小分子，核糖、磷酸和 4 种碱基（A、G、U、C）。从简单的分子组成比例上就可以看出，核酸分子的骨架应该由糖和磷酸组成，而信息密码隐藏在碱基的排列组合变化中。

第二节 DNA 结构与复制

一、DNA 的结构

Maurice Wilkins 和 Rosalind Franklin 依据对 X 射线衍射照片分析，提出 DNA 是由两条长链组成的双螺旋。Erwin Chargaff 测定了 DNA 的分子组成，发现 DNA 中的 4 种碱基的含量并不是等量的，但是 A 和 T 的含量总是相等，G 和 C 的含量也相等。James Dewey Watson 和 Francis Harry Compton Crick 首先意识到该比值的重要性，并请剑桥的 John Griffith 计算出 A 吸引 T、G 吸引 C、A+T 的宽度与 G+C 的宽度相等。随后，他们结合 X 射线衍射照片构建出了 DNA 分子双螺旋结构模型。

揭示 DNA 双螺旋结构为现代分子生物学的标志性成就，它不仅说明了 DNA 为什么是遗传信息的携带者，而且说明了基因的复制和突变等机理。1954 年，James Dewey Watson 和 Francis Harry Compton Crick 发表了有关 DNA 双螺旋结构的论文，虽然只有一千余字，但其奠定了现代分子生物学的基础。论文包括如下主要内容。

(1) DNA 由脱氧核糖和磷酸基通过酯键交替连接而成。其主链有两条，它们绕一共同轴心以右手方向盘旋，相互平行而走向相反，形成双螺旋构型。主链处于螺旋的外

侧,由糖和磷酸构成的主链具备亲水性。

(2) 碱基位于螺旋的内侧,它们以垂直于螺旋轴的取向通过糖苷键与主链糖基相连。同一平面的碱基在两条主链间形成碱基对。配对碱基总是 A 与 T 和 G 与 C。碱基对以氢键维系, A 与 T 间形成两个氢键, G 与 C 间形成三个氢键。两种碱基对的几何大小又十分相近,具备了形成氢键的适宜键长和键角条件。每对碱基处于各自的平面上,但螺旋周期内的各碱基对平面的取向均不同。双螺旋结构在满足两条链碱基互补的前提下, DNA 的一级结构不受限制。

二、DNA 复制

DNA 复制是指 DNA 双链在细胞分裂以前进行的复制过程,复制的结果是一条双链变成两条一样的双链,每条双链都含有原来双链的一条单链。这个过程通过半保留复制(semiconservative replication)机制得以完成。DNA 复制的特点可以归纳为以下几点。

(1) 半保留复制:1958 年 Matthew Meselson 和 Franklin Stahl 的实验证明 DNA 在复制时,以亲代 DNA 链作模板,合成完全相同的两个双链子代 DNA,每个子代 DNA 中都含有一股亲代 DNA 链,这种现象称为 DNA 的半保留复制。

(2) 有复制起始点:DNA 的复制需在特定位置开始,这些具有特定核苷酸序列的片段,即复制起始点。在原核生物中,因为较小的基因组规模,复制起始点通常为一个,而真核生物庞大的基因组复制一般有多个复制起点,从而保证复制在一定时间内完成。

(3) 需要 RNA 引物:DNA 聚合酶必须以一段具有 3' 端自由羟基(3'-OH)的 RNA 作为引物,才能开始合成子代 DNA 链。RNA 引物的大小,在原核生物中通常为 50~100 个核苷酸,而在真核生物中约为 10 个核苷酸。

(4) 双向复制:DNA 复制时,以复制起始点为中心,向两个方向进行复制。但在低等生物中,也可进行单向复制。

(5) 半不连续复制(semidiscontinuous replication):DNA 聚合酶只能以 5'→3' 方向聚合子代 DNA 链,两条亲代 DNA 链作为模板合成子代 DNA 链时的方式是不同的。以 3'→5' 方向的亲代 DNA 链作模板的子链在聚合时基本上是连续进行的,这一条链被称为前导链。而以 5'→3' 方向的亲代 DNA 链为模板的子链在聚合时则是不连续的,这条链被称为随从链。DNA 在复制时,由随从链所形成的多个子代 DNA 短链称为冈崎片段。

一般认为, DNA 复制一旦开始,就会将该 DNA 分子全部复制完毕。其实在 DNA 上也存在着复制终止位点, DNA 复制将在复制终止位点处终止。在 DNA 复制终止阶段,令人困惑的一个问题是线性 DNA 分子两端是如何完成其复制的? 已知 DNA 复制都要有 RNA 引物参与。当 RNA 引物被切除后,中间所遗留的间隙由 DNA 聚合酶 I 所催化填充。在线性分子的两端以 5'→3' 为模板的随从链的合成,其末端的 RNA 引物被切除后是无法被 DNA 聚合酶催化填充的。1941 年 Barbara McClintock 提出了端粒(telomere)的假说,认为染色体末端必然存在一种特殊结构——端粒。其作用包括保持

染色体末端稳定和参与染色体核纤层相连定位。

1978年,四膜虫的端粒结构首先被测定。1990年Calvin Harley就把端粒与衰老挂上了钩,提出“细胞越老,其端粒长度越短”的观点,细胞分裂一次,其端粒的DNA丢失30~200 bp。端粒的复制不能由DNA聚合酶催化,而是由一种特殊的反转录酶——端粒酶完成。真核生物染色体末端DNA复制是由端粒酶将一个末端DNA序列加在刚刚完成复制的DNA末端。例如,在四膜虫细胞中的线性DNA分子末端有30~70拷贝的5'-TTGGGG-3'序列,端粒酶可以将TTGGGG序列加在事先已存在的单链DNA末端的TTGGGG序列上。这样有较长的末端单链DNA,可以被引发酶重新引发或其他的酶蛋白引发而合成RNA引物,并由DNA聚合酶将其变成双链DNA。这样就可以避免其DNA随着复制的不断进行而逐渐变短。正常人体细胞中检测不到端粒酶。在一些良性病变细胞、体外培养的成纤维细胞中也检测不到端粒酶活性。但在生殖细胞及胎儿细胞中此酶为阳性,恶性肿瘤细胞具有高活性的端粒酶。人类肿瘤中广泛地存在着较高水平的端粒酶,用其作为肿瘤治疗的靶点是较受关注的热点。

第三节 RNA 分类、转录与转录后过程

一、RNA 分类与结构

在DNA指导的RNA聚合酶催化下,生物体以DNA的一条链为模板,按照碱基配对原则,合成一条与DNA链的一定区段互补的RNA链,这个过程称为转录。经转录生成的多种RNA,主要包括rRNA、tRNA、mRNA、snRNA和hnRNA等(图1-1)。

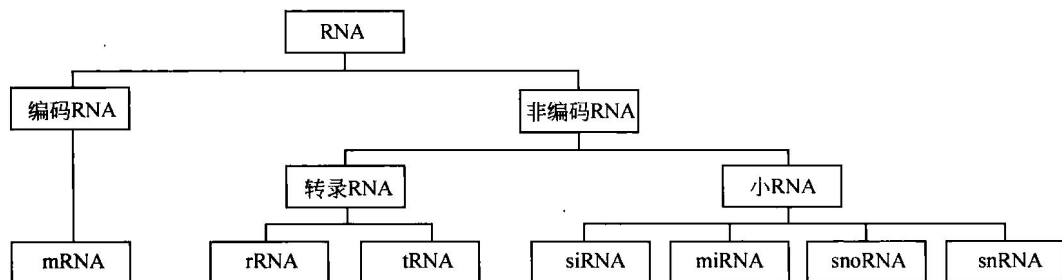


图 1-1 RNA 的分类

绝大部分RNA分子都是线性单链,但是RNA分子的某些区域可自身回折进行碱基互补配对,形成局部双螺旋。在RNA局部双螺旋中A与U配对、G与C配对。除此以外,还存在非标准配对,如G与U配对。RNA分子中的双螺旋与A型DNA双螺旋相似,而非互补区则膨胀形成凸出或者环,这种短的双螺旋区域和环称为发夹结构。发夹结构是RNA中最常见的二级结构形式,二级结构进一步折叠形成三级结构,RNA只有在具有三级结构时才能成为有活性的分子。RNA也能与蛋白质形成核蛋白复合物——RNA的四级结构。

(一) tRNA

tRNA 约占总 RNA 的 15%，tRNA 主要的生理功能是在蛋白质生物合成中转运氨基酸和识别密码子，细胞内每种氨基酸都有其相应的一种或几种 tRNA，因此 tRNA 的种类很多，在细菌中有 30~40 种 tRNA，在动物和植物中有 50~100 种 tRNA。从数量上直观分析就可以推断，部分真核生物蛋白在细菌中可能无法顺利表达。

tRNA 一级结构：tRNA 是单链分子，含 73~93 个核苷酸，有 10% 的稀有碱基，如 DHU、rT 和 ψ 及不少碱基被甲基化。其 3' 端为 CCA-OH，5' 端多为 pG，分子中大约 30% 的碱基是保守的。

tRNA 二级结构：tRNA 二级结构为三叶草形。配对碱基形成局部双螺旋而构成臂，不配对的单链部分则形成环。三叶草形结构由 4 臂 4 环组成。氨基酸臂由 7 对碱基组成，双螺旋区的 3' 端为一个 4 个碱基的单链区-NCCA-OH 3'，腺苷酸残基的羟基可与氨基酸 α 羧基结合而携带氨基酸。DHU 环以含有 2 个稀有 DHU 而得名，由 8~14 个碱基组成；DHU 臂由 3~4 对碱基组成。反密码子环由 7 个碱基组成，其中 3 个核苷酸组成反密码子 (anticodon)，在蛋白质生物合成时，可与 mRNA 上相应的密码子配对。反密码子臂由 5 对碱基组成。不同生物携带的一定反密码子 tRNA 的丰度不同，这种差异会造成转基因过程中出现所谓稀有密码影响转基因表达的问题，需要依据受体细胞对转基因进行修改。额外环在不同 tRNA 分子中变化较大，可在 4~21 个碱基之间变动，又称为可变环，是 tRNA 分类的重要指标。T ψ C 环含有 7 个碱基，所有的 tRNA 在此环中都含 T ψ C 序列，T ψ C 臂由 5 对碱基组成。

tRNA 的三级结构：20 世纪 70 年代初科学家用 X 射线衍射分析发现 tRNA 的三级结构为倒“L”形。tRNA 三级结构的特点是氨基酸臂与 T ψ C 臂构成“L”的横，—CCAOH-3' 端就在这一横的端点上，是结合氨基酸的部位；而 DHU 臂与反密码子臂及反密码子环共同构成“L”的竖，反密码子环在一竖的端点上，能与 mRNA 上对应的密码子识别；DHU 环与 T ψ C 环在“L”的拐角上。三级结构氢键的形成与 tRNA 中不变的核苷酸密切相关，各种 tRNA 三级结构都呈倒“L”形。

(二) mRNA

原核生物中 mRNA 转录后直接进行蛋白质翻译。转录和翻译不仅发生在同一空间，而且两个过程几乎是同时进行的。原核生物的 mRNA 结构简单，往往含有几个功能上相关的蛋白质编码序列，可翻译出几种蛋白质，因此被称为多顺反子。在原核生物 mRNA 中编码序列之间有间隔序列，可能与核糖体的识别和结合有关。在 5' 端与 3' 端有与翻译起始和终止有关的非编码序列，原核生物 mRNA 中没有修饰碱基，5' 端无帽结构，3' 端无聚腺苷酸 (polyA) 的尾。原核生物转录后约 1min mRNA 就开始降解。所以原核生物一般不使用基于 mRNA 的文库。

真核细胞成熟 mRNA 是由其前体——核内不均一 RNA (heterogeneous nuclear