

分子遗传学

MOLECULAR GENETICS

□ 主编 路铁刚 丁 毅



高等教育出版社
Higher Education Press

分子遗传学

内容简介

MOLECULAR GENETICS

主编 路铁刚 丁毅

参加编著人员 (以姓氏汉语拼音为序)

丁毅 路铁刚 孙英莉

邢更妹 张春义 赵双宜



高等教育出版社

Higher Education Press

ISBN 7-04-01168-1
定价: 35.00元

内容简介

本书以基因和基因组的结构与功能为基础,从分子水平上阐述了基因变异、基因调控、遗传重组与转座、基因与发育、基因与免疫多样性等重大分子遗传学的核心命题,特别是结合近代基因概念的发展对表观遗传学和近年广泛流行的几种重要病毒病发生的分子机制进行了专题介绍和讨论,综述了基因组和后基因组研究的进展,最后对分子遗传学研究中常用技术的原理做了较为广泛而系统的介绍。

全书图文并茂,内容新颖,观点明确,介绍了分子遗传学的发展趋势,为读者提供了一个阅览和探究分子遗传学知识的新窗口和知识平台。本书是综合性大学、理工科大学、农林院校、师范院校和医学院校等生命科学类各专业本科生和研究生的专业课教材,也可供相应专业的科技工作者参考。

图书在版编目(CIP)数据

分子遗传学/路铁刚,丁毅主编. —北京:高等教育出版社,2008.10

ISBN 978-7-04-024415-1

I. 分… II. ①路…②丁… III. 分子遗传学-高等学校-教材 IV. Q75

中国版本图书馆 CIP 数据核字(2008)第 141851 号

策划编辑 王 莉 责任编辑 张晓晶 封面设计 张 楠 责任绘图 尹 莉
版式设计 王艳红 责任校对 王 超 责任印制 朱学忠

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100120
总 机 010-58581000

经 销 蓝色畅想图书发行有限公司
印 刷 河北省财政厅票证文印中心

开 本 850×1168 1/16
印 张 32.75
字 数 830 000

购书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landracom.com>
<http://www.landracom.com.cn>
畅想教育 <http://www.widedu.com>

版 次 2008 年 10 月第 1 版
印 次 2008 年 10 月第 1 次印刷
定 价 52.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 24415-00

前 言

以基因组学及后基因组学为主流的现代生命科学的迅速发展,为分子遗传学赋予了新的内涵并提供了良好的发展机遇。在新理论、新知识、新技术和新成果迅猛增长的形势下,为了适应教学改革的要求、满足广大本科生和研究生的求知欲望,迫切需要编写一本既传承分子遗传学的知识体系又反映当代生命科学发展前沿的教科书。为此,高等教育出版社于2006年组织部分高等院校和科研单位长期从事分子遗传学教学和科研的有关专家学者讨论了分子遗传学教材编写大纲。经过反复斟酌和修改,几易其稿构筑了本教材的框架基础。

根据上述原则,本教材从分子水平上阐述了基因组结构与功能、基因突变与DNA损伤修复、基因表达与调控、遗传重组与转座、基因与发育、基因与免疫多样性等分子遗传学的核心命题。特别是以相应的篇幅讨论了分子遗传学的几个重要发展分支,如表观遗传学、基因组与后基因组学的研究进展与动态,并专门介绍了分子遗传学研究中常用技术,以期拓宽学生知识面,在更深层次上理解和掌握基础理论和研究动向。

本教材共分为12章,根据各位作者研究方向和专长进行了分工:中国农业科学院生物技术研究所路铁刚,第5章和第12章;武汉大学生命科学学院丁毅,第2章和第11章;中国科学院高能物理研究所邢更妹,第1章和第6章;山东大学生命科学院赵双宜,第3章和第4章;中国农业科学院生物技术研究所张春义,第9章和第10章;兰州大学生命科学学院孙英莉,第7章和第8章。

王亚馥教授对全部书稿进行了认真的审阅,提出了许多宝贵建议,对保证本书的质量起了重要的作用。高等教育出版社王莉编辑和张晓晶编辑对本书的出版付出了大量的劳动,没有她们的支持和帮助,本书不可能这样顺利出版。张芊博士、彭昊博士、张伟博士、张治国博士、王琦琳、王元火同志为本书的稿件收集与整理做了大量工作。对为本书出版发行付出辛勤劳动的所有人员,我们在此一并致以衷心的感谢!

由于分子遗传学发展迅速,加之作为教科书的篇幅有限以及时间仓促,本书无法将分子遗传学的知识面面俱到,难免有不足甚至错误之处,恳请各位老师、同学和读者提出宝贵意见,以便再版时进行补充和更正。

编 者

2008年1月7日于北京

目 录

1	绪论	1
1.1	分子遗传学的涵义及其研究任务	2
1.1.1	分子遗传学的涵义	2
1.1.2	分子遗传学研究的任务	2
1.2	分子遗传学的建立	5
1.2.1	物理学的渗透	5
1.2.2	遗传物质是核酸	5
1.2.3	分子遗传学的诞生	8
1.3	中心法则及其发展	16
1.3.1	中心法则的涵义	16
1.3.2	RNA 编辑与中心法则	17
1.3.3	朊病毒与中心法则	18
1.4	基因概念及其发展	21
1.4.1	顺反子	21
1.4.2	操纵子与基因家族	22
1.4.3	外显子与内含子	24
1.4.4	重叠基因与转座因子	26
1.4.5	现代基因的概念与界定	27
1.5	基因组学与后基因组学	28
1.5.1	基因组与基因组学	28
1.5.2	后基因组学	30
1.6	生物信息学的兴起	31
1.6.1	生物信息学的涵义	31
1.6.2	生物信息学的研究内容	32
1.7	分子遗传学与社会	32
1.7.1	基因工程与现代生物技术	32
1.7.2	蛋白药物与疫苗研究	35
1.7.3	分子诊断与基因治疗	36
1.7.4	环境污染与生物净化	39
2	基因组的结构与功能	43
2.1	病毒基因组	44
2.1.1	病毒基因组的结构多样性	44
2.1.2	反转录病毒基因组的结构与功能	45

2.1.3	病毒基因组中特征序列的结构与功能	46
2.2	原核生物基因组	48
2.2.1	大肠杆菌基因组	48
2.2.2	拟核结构	48
2.2.3	操纵子结构	49
2.3	真核生物基因组	49
2.3.1	真核生物基因组大小与 C 值悖理和 N 值悖理	49
2.3.2	真核生物 DNA 的复性动力学	51
2.3.3	基因家族	56
2.3.4	核小体结构与染色质	59
2.3.5	真核生物染色体的结构与功能	60
2.3.6	异染色质形成的分子机制	63
2.3.7	常染色质基因表达的分子基础	68
2.3.8	染色质的复制和转录	75
2.4	核外基因组	80
2.4.1	质粒基因组	80
2.4.2	线粒体基因组	83
2.4.3	叶绿体基因组	85
3	DNA 复制与基因表达	88
3.1	DNA 的复制	89
3.1.1	DNA 复制的一般特征	89
3.1.2	DNA 复制的酶学与机制	93
3.1.3	线状 DNA 的末端复制机制	101
3.2	转录	102
3.2.1	参与转录的组成成分与结构	103
3.2.2	转录的过程	108
3.3	RNA 的加工	113
3.3.1	rRNA 的加工	113

3.3.2	tRNA 的加工	114	5.1.3	显性突变和隐性突变	182
3.3.3	mRNA 前体的加工	115	5.1.4	同义突变、错义突变和无义突变	182
3.4	翻译	125	5.1.5	功能缺失型突变与功能获得型突变	183
3.4.1	参与翻译的组成成分与结构	125	5.1.6	非条件突变和条件突变	183
3.4.2	翻译的起始	128	5.1.7	回复突变与抑制突变	183
3.4.3	肽链的延伸和终止	133	5.1.8	显性负突变	184
4	基因表达的调控	140	5.2	基因突变的分子基础	185
4.1	原核生物基因表达的调控	141	5.2.1	碱基替换	185
4.1.1	乳糖操纵子——可诱导的负调控和正调控模型	141	5.2.2	插入或缺失突变	186
4.1.2	半乳糖操纵子	142	5.2.3	移码突变	186
4.1.3	阿拉伯糖操纵子	143	5.3	诱发基因突变的因素	187
4.1.4	色氨酸操纵子	144	5.3.1	基因突变的生物因素	187
4.1.5	转录的时序控制和翻译调节	147	5.3.2	基因突变的物理及化学因素	188
4.2	真核生物基因表达的调控	152	5.3.3	定点突变与突变热点	193
4.2.1	真核细胞转录调控机制	152	5.3.4	诱变剂的检测——Ames 测试	196
4.2.2	真核生物基因表达的组合控制	157	5.4	DNA 的修复	197
4.3	真核生物基因表达的多层次调控	159	5.4.1	光修复	197
4.3.1	细胞及染色体水平的调控	159	5.4.2	切除修复	198
4.3.2	DNA 甲基化与基因表达活性	160	5.4.3	错配修复	200
4.3.3	DNA 重排与基因表达调控	161	5.4.4	重组修复	200
4.3.4	转录后水平的调控	164	5.4.5	双链断裂修复系统	202
4.3.5	翻译和翻译后水平调控	170	5.4.6	SOS 修复	203
4.4	RNA 干涉与基因表达调控	173	5.5	突变体的创制与应用	204
4.4.1	RNA 干涉现象的发现	173	5.5.1	EMS 突变体	204
4.4.2	RNA 干涉与基因沉默	173	5.5.2	快中子突变体	205
4.5	RNA 编辑	175	5.5.3	T-DNA 标签突变体	206
4.5.1	位点特异性脱氨基作用	175	5.5.4	转座子标签突变体	208
4.5.2	gRNA 指导尿嘧啶插入或删除	176	5.5.5	突变体库的饱和度分析	209
5	基因突变与 DNA 损伤修复	180	5.6	突变体的筛选与检测	211
5.1	基因突变的类型	181	5.6.1	突变体的遗传筛选	211
5.1.1	自发突变和诱发突变	181	5.6.2	突变体位点的分子检测	213
5.1.2	体细胞突变与生殖细胞突变	181	6	遗传重组与转座	219
5.1.3	显性突变和隐性突变	182	6.1	遗传重组及其类型	220
5.1.4	同义突变、错义突变和无义突变	182	6.1.1	遗传重组的概念	220
5.1.5	功能缺失型突变与功能获得型突变	183	6.1.2	遗传重组的类型	222
5.1.6	非条件突变和条件突变	183	6.2	真核生物同源重组的分子机制	223
5.1.7	回复突变与抑制突变	183	6.2.1	同源重组的 Holliday 模型	223
5.1.8	显性负突变	184	6.2.2	双链断裂起始重组模型	226

6.2.3	重组与联合复合体	226
6.2.4	基因转换导致等位基因间的 重组	230
6.3	细菌同源重组的分子基础	232
6.3.1	RecBCD 识别 <i>chi</i> 序列引发 重组	232
6.3.2	RecA 催化单链同化	234
6.3.3	Ruv 系统解离 Holliday 连 接点	235
6.4	位点专一性重组的分子机制	237
6.4.1	λ 噬菌体的整合与切离	237
6.4.2	位点专一性重组的机制	239
6.4.3	λ 噬菌体重组发生在整合 体中	240
6.5	转座因子及其分类	241
6.5.1	转座因子的发现	241
6.5.2	DNA 转座	245
6.5.3	反转录转座子	246
6.6	原核生物中的转座因子	248
6.6.1	插入序列	248
6.6.2	转座子	250
6.6.3	转座噬菌体	252
6.7	真核生物中的转座子	253
6.7.1	酵母菌的转座子	253
6.7.2	果蝇的转座子	255
6.7.3	玉米的转座子	257
6.7.4	人类基因组中的转座子	258
6.8	转座作用的分子机制	260
6.8.1	DNA 转座机制	260
6.8.2	反转座子的转座机制	262
6.9	转座因子的遗传学效应与应用	269
6.9.1	改变染色体结构	269
6.9.2	诱发基因突变	270
6.9.3	调节基因表达	272
6.9.4	产生新的变异	273
6.9.5	转座子标记克隆目的基因	274
6.9.6	转座因子作为基因工程 载体	274
7	基因与发育	278
7.1	细胞分化和细胞决定	279
7.1.1	单细胞生物的细胞分化与 基因调控	279
7.1.2	多细胞生物的细胞分化与 细胞决定	279
7.1.3	线虫的细胞特化	281
7.1.4	程序性细胞死亡与凋亡	281
7.2	胚胎极性的决定	284
7.2.1	果蝇胚胎的极性	284
7.2.2	果蝇前-后轴形成	285
7.2.3	果蝇背-腹轴形成	289
7.2.4	分节基因与果蝇胚胎体节的 形成	291
7.2.5	同源异形基因	293
7.3	高等植物成花诱导及花器官发育的 基因调控	299
7.3.1	高等植物发育基本过程及其 特点	299
7.3.2	成花诱导中的基因调控	300
7.3.3	高等植物花器官发育的 ABC 模型	301
7.4	细胞周期的基因调控	303
7.4.1	细胞周期事件	303
7.4.2	细胞周期的调节因子	304
7.4.3	<i>cdc</i> 基因对酵母细胞周期的 调控	306
7.4.4	p53 蛋白对癌细胞生长的负 控制	308
7.4.5	细胞周期失控与肿瘤	308
8	基因与免疫多样性	311
8.1	免疫应答	312
8.1.1	体液免疫	312
8.1.2	细胞免疫	312
8.1.3	免疫系统的应答特点	313
8.2	抗体多样性产生的机制	314
8.2.1	抗体的结构与功能	314
8.2.2	抗体基因及其重组	315
8.2.3	抗体基因重组与抗体多 样性	317
8.2.4	免疫重组的方式	318
8.2.5	有效重排引发等位基因 排斥	321
8.2.6	体细胞突变增加免疫多 样性	321

8.2.7	假基因参与鸟类免疫球蛋白的装配	322
8.3	Ig 类型转换的机制	323
8.3.1	DNA 重组导致 Ig 类型的转换	323
8.3.2	RNA 加工改变早期重链基因表达	324
8.4	T 细胞受体基因的重排机制	326
8.4.1	T 细胞受体的类型以及对外源抗原的识别	326
8.4.2	TCR 的功能	326
8.4.3	T 细胞受体基因	327
8.4.4	TCR 基因的重排及其多样性产生的机制	328
8.5	主要组织相容性复合体	329
8.5.1	MHC 抗原的类型及其功能	329
8.5.2	MHC 抗原的结构	329
8.5.3	MHC 基因的定位	330
8.5.4	MHC 抗原基因的结构	331
8.5.5	I 型 MHC 基因的表达调控机制	332
8.5.6	II 型 MHC 基因的表达调控机制	333
9	表观遗传学	336
9.1	表观遗传变异的发现	337
9.2	真核生物染色质重塑	337
9.2.1	染色质重塑的概念	337
9.2.2	染色质重塑因子	337
9.2.3	染色质重塑与发育	340
9.2.4	染色质重塑与人类疾病	340
9.2.5	染色质重塑与基因剂量补偿	341
9.3	DNA 甲基化	342
9.3.1	DNA 甲基化酶	342
9.3.2	DNA 甲基化位点及甲基化类型	344
9.3.3	DNA 甲基化的转录抑制机制	345
9.3.4	DNA 甲基化与组蛋白修饰	346
9.4	基因组印记	347

9.4.1	基因组印记的发现	347
9.4.2	基因组印记与印记基因	349
9.4.3	印记基因的功能	350
9.4.4	DNA 甲基化与基因组印记	350
9.4.5	基因组印记模型	351
9.5	RNA 编辑	352
9.5.1	核基因组 RNA 编辑	353
9.5.2	线粒体基因组 RNA 编辑	355
9.5.3	叶绿体基因组 RNA 编辑	356
9.5.4	RNA 编辑的作用	357
9.6	DNA 甲基化的检测方法	357
9.6.1	全基因组水平甲基化分析	357
9.6.2	特异性位点的 DNA 甲基化的检测	358
9.6.3	甲基化新位点的寻找方法	360
9.7	表观遗传学研究的应用与展望	361
9.7.1	表观遗传学研究应用	361
9.7.2	表观遗传学研究展望	362
10	重要病毒疾病发生的分子机制	366
10.1	反转录病毒	367
10.1.1	反转录病毒颗粒结构特点及其功能	367
10.1.2	反转录病毒基因表达与加工	368
10.2	人类免疫缺陷病毒	369
10.2.1	HIV 粒子的形态结构与侵染	369
10.2.2	HIV 基因组复制与病毒增殖	373
10.2.3	HIV 的感染与致病机制	374
10.2.4	AIDS 的治疗和预防	375
10.3	严重急性呼吸综合征冠状病毒	375
10.3.1	SARS 冠状病毒的形态结构	375
10.3.2	SARS 冠状病毒的基因组结构及其编码蛋白的特点	376
10.3.3	SARS 病毒生活史与致病机制	378
10.3.4	SARS 的预防和治疗	379
10.4	禽流感病毒	379

10.4.1	AIV 粒子的结构	380	11.5.2	YAC 文库	417
10.4.2	AIV 的基因组及其编码		11.5.3	P1 噬菌体衍生人工染色体	
	蛋白	380		文库	418
10.4.3	AIV 的感染和复制	382	11.5.4	可转移的人工染色体 BiBAC/	
10.4.4	AIV 致病机制	385		TAC 文库	420
10.5	乙型肝炎病毒	386	11.6	后基因组学研究	421
10.5.1	HBV 的结构	386	11.6.1	比较基因组学	421
10.5.2	HBV 基因组结构	387	11.6.2	功能基因组学	424
10.5.3	HBV 致病机制	390	11.6.3	蛋白质组学	424
10.5.4	HBV 的预防和治疗	391	11.6.4	转录物组学	426
			11.6.5	代谢物组学	426
			11.6.6	化学基因组学	427
			11.6.7	生物信息学与后基因组学	428
			11.6.8	ENCODE 计划	428
11	基因组学与后基因组学	394	11.7	基因组的进化	429
11.1	人类基因组计划	395	11.7.1	基因组进化的分子基础	429
11.1.1	HGP 的提出和科学意义	395	11.7.2	基因组的起源	430
11.1.2	人类基因组的结构特点	396	11.7.3	基因组的进化	431
11.1.3	遗传图谱	398			
11.1.4	物理图谱	398			
11.1.5	人类基因组序列图	398			
11.2	模式生物基因组研究	401	12	分子遗传学研究技术	436
11.2.1	大肠杆菌基因组研究	401	12.1	基因表达与调控的研究技术	437
11.2.2	酵母基因组研究	403	12.1.1	差减杂交	437
11.2.3	果蝇基因组研究	403	12.1.2	mRNA 差异显示和限制性	
11.2.4	拟南芥基因组研究	404		差异显示	440
11.2.5	水稻基因组研究	404	12.1.3	基因表达连续分析	442
11.2.6	宏基因组学研究	405	12.1.4	微阵列技术	443
11.3	基因组测序与序列装配	406	12.1.5	基因表达快速分析技术	446
11.3.1	DNA 测序的策略	406	12.2	基因组学和蛋白质组学研究技术	448
11.3.2	DNA 序列的装配	407	12.2.1	DNA 大片段文库构建与	
11.3.3	人类基因组的测序与装配	408		基因组测序	448
11.4	基因组图谱构建与应用	409	12.2.2	酵母双杂交技术	449
11.4.1	人类基因组遗传图谱的		12.2.3	双分子荧光互补技术	450
	构建	409	12.2.4	蛋白质组学技术	452
11.4.2	植物基因组遗传图谱的		12.2.5	双向差异凝胶电泳与多维	
	构建	411		色谱和质谱技术	453
11.4.3	制作物理图谱的基本要素	412	12.3	基因遗传操作技术	455
11.4.4	细胞遗传学图谱	413	12.3.1	基因克隆技术	455
11.4.5	辐射杂种图谱	413	12.3.2	重组 DNA 分子构建技术	456
11.4.6	限制性酶切图谱	414	12.3.3	转基因技术	458
11.4.7	叠连群图谱	414	12.3.4	基因打靶技术	460
11.4.8	基因组图谱的应用	415	12.3.5	基因表达及功能鉴定技术	461
11.5	基因组 DNA 大片段文库的构建	416	12.3.6	体细胞克隆技术	464
11.5.1	BAC 文库	417			

12.4	遗传分子标记技术	466
12.4.1	分子标记的种类	466
12.4.2	RFLP 和 AFLP	468
12.4.3	RAPD 及相近的分子标记	470
12.4.4	CAPS 和微卫星标记	471
12.5	基因治疗技术	473
12.5.1	基因治疗策略	473
12.5.2	基因治疗中的转基因技术	473
12.5.3	基因治疗中的 RNA 修饰	475
12.5.4	胚胎干细胞与基因治疗	477
参考文献		
名词解释		
索引		

12.5.3	基因治疗中的 RNA 修饰	475
12.5.4	胚胎干细胞与基因治疗	477
参考文献		
名词解释		
索引		

1

绪论

自从 20 世纪 50 年代初,以遗传信息载体的 DNA 双螺旋结构的提出为标志宣告分子遗传学诞生以来,在这半个多世纪里,以研究基因为核心的分子遗传学的发展速度超过了生命科学发展史上的任何一个时期。人们真正在分子水平上揭示基因的结构与功能,阐明生物遗传与变异的本质,对生命世界的理性认识达到前所未有的深度与广度。特别是被誉为 20 世纪三大科学工程之一的人类基因组计划 (Human Genome Project, HGP) 的启动实施,到 2001 年人类基因组序列草图的完成,直到 2006 年人类最后一个染色体的成功破译,历时 16 年的人类基因组计划书写完了生命之书的最后一个章节,人类基因密码 99.99% 被解密,这是分子遗传学乃至整个生命科学发展史上的一个重要的里程碑。随着 HGP 的实施与完成,带动了各类生物的基因组研究,并迅速形成了一个新的分支学科,即基因组学 (genomics)。紧接着是“后基因组时代”(postgenome era) 新纪元的开始,其中功能基因组学 (functional genomics) 和蛋白质组学 (proteomics) 又成为研究的重心。这一系列新的分支学科的出现都与分子遗传学的发展密切相关,其研究内容是从基因组 (genome) 到转录物组 (transcriptome) 和蛋白质组 (proteome) 的结构与功能。实质就是从 DNA、RNA 和蛋白质等分子水平上研究遗传学的核心命题:基因型+环境=表型。这也正是分子遗传学研究的本质问题。总之,分子遗传学已成为当今最重要的前沿学科之一,并对生命科学及其相关学科的发展产生了重大而深远的影响。本章正是以这些研究进展与成果为基础,以论述分子遗传学的建立与发展,中心法则及其发展,基因概念及其发展等内容作为本书绪言。

1.1 分子遗传学的涵义及其研究任务

1.1.1 分子遗传学的涵义

遗传学是以基因作为研究的核心,它是研究基因的结构、功能、变异、传递和表达规律的学科。分子遗传学是遗传学的一个分支学科,是在分子水平上研究基因的结构与功能,以揭示生物遗传和变异以及表达的分子机制。它研究的范畴包含基因在生命系统中的储存、组织结构、基因的复制与传递的分子机制、基因表达与调控规律、基因表达产物的结构与功能、基因变异的分子机制、基因在控制细胞分裂、生长和分化以及形态发生与个体发育中的作用机制。总之,分子遗传学研究的内容是生物活细胞内与遗传变异相关的动态生命过程中所有的分子事件。

1.1.2 分子遗传学研究的任务

(1) 研究遗传物质的分子结构与传递机制 作为遗传物质的 DNA 或 RNA 具有什么样的结构?又是如何实现其传递遗传性状的功能的?所有已知的自然界有机体和许多病毒的遗传物质都是 DNA,有些病毒的遗传物质是 RNA。作为遗传物质必须具备的特性是:① 储存并表达遗传信息。② 能把遗传信息传递给子代。③ 物理和化学性质稳定。④ 含有遗传重组和变异的信息。1953 年 Watson 和 Crick 提出的 DNA 双螺旋结构模型,不仅表明了 DNA 序列由不同的核苷酸排列而成,遗传信息就储存在其中,而且精确阐明了遗传信息的复制与传递,特别是通过 DNA 双链分子的“半保留复制”(semiconservative replication),由亲链(parental chain)产生了两个与亲链完全相同的 DNA 双链分子,即两个子链 DNA 分子(progeny DNA molecular),保证了遗传物质从亲代向子代的传递,为分子遗传学的诞生奠定了理论基础[图 1-1(a)]。某些病毒,主要是反转录病毒,是以 RNA 作为遗传物质,尽管 RNA 的化学结构与 DNA 不同,但复制的基本原理是相似的,也是通过其互补链的合成来实现的,从而起着遗传信息的储存与传递的功能(详见第 3 章)。

(2) 研究遗传信息表达的分子机制 作为遗传物质不仅要具有精确复制和传递的机制,同时要具有控制性状表达实现功能的机制。我们知道,生命活动和功能的表达是通过蛋白质来实现的。蛋白质分为酶、运输蛋白、贮藏蛋白、毒蛋白、抗体、激素、细胞因子、基因表达调控因子和结构蛋白等等。蛋白质不仅构成了生命体的基本结构,而且生物的生长、发育、繁殖、代谢物的合成和分解、能量的产生和利用、对外界刺激的反应及运动等都是蛋白质功能的具体体现。同时,不同生物有稳定遗传的不同的结构和性状,显然这些表现都是由基因决定的。也就是说,基因决定生物的遗传性状是通过蛋白质来实现的。基因将所携带的遗传信息传递给蛋白质,不同基因决定不同的蛋白质。那么基因又是如何将遗传信息传递给蛋白质的?早在 20 世纪 50 年代,关于遗传信息的流向与表达,Crick 提出了中心法则(central dogma),它指出了遗传信息的流动方向,即遗传信息的复制、转录、翻译形成蛋白质的全过程,也就是 DNA、RNA 到蛋白质之间的线性关系[图 1-1(b)]。中心法则对分子遗传学的建立与发展具有重要意义,也是分子遗传学的重要基础。经过近半个多世纪的研究发展,中心法则不断地得到修正和补充,使其逐步得到完善。

(3) 研究基因表达调控的分子机制 基因表达(gene expression)是指基因通过转录和翻译最终产生功能产物的过程。这些功能产物为蛋白质或 RNA(如 tRNA、rRNA 等)。我们都知道,

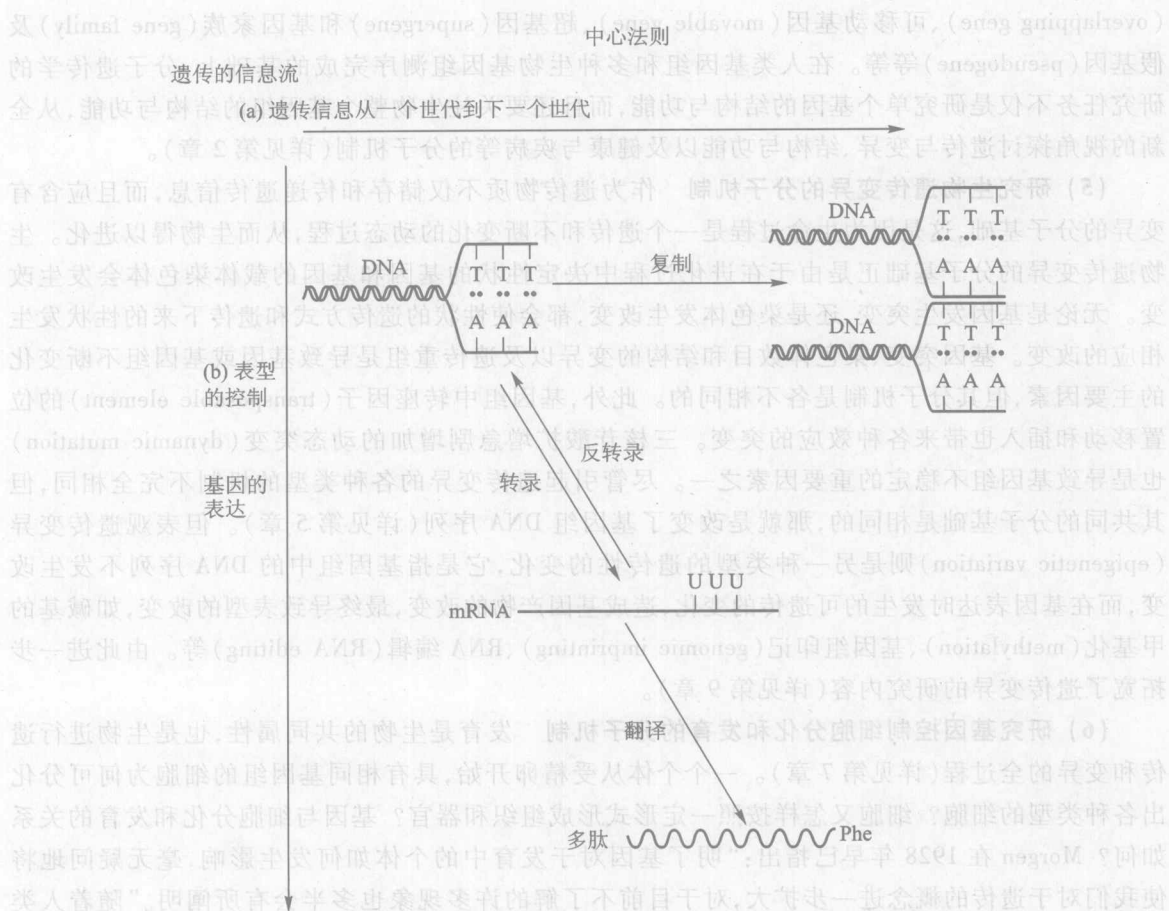


图 1-1 依照中心法则的遗传信息流向(引自 Sunstad et al, 2003)

生物在不同内外条件下,或在不同发育时期,以至不同组织或器官中,尽管具有相同的基因组,但是其表达产物是不同的。显然基因表达过程受到一系列精确的调控,调控可以发生在基因表达的任何阶段。早在 20 世纪 60 年代, Jacob 和 Monod 提出的大肠杆菌乳糖操纵子(*lac operon*)模型已表明,遗传信息的表达与调控是统一的,基因不仅是遗传信息的载体,同时又具有调控基因表达活性的功能。这一套相互制约的基因使生物在不同环境条件下表现出不同的遗传特性。该模型为研究基因表达调控的分子机制奠定了基础。只是原核生物基因与真核生物基因的结构不同,如原核生物基因一般成簇排列,构成一个操纵子,一个操纵子中的几个基因都受同一个启动子的调控;而真核基因则多半是独立存在,各有自己的启动子。因此,真核基因与原核基因表达调控机制也不相同。随着人类和相关生物基因组计划的完成而进入功能基因组时代,研究真核基因表达调控的分子机制将是分子遗传学所面临的一项十分迫切而艰巨的任务(详见第 4 章)。

(4) 研究基因和基因组的结构与功能 任何一种生物的基因组都具有单倍体细胞内所含的整套染色体(chromosome),它蕴藏着一套完整的遗传信息,决定了该生物体的遗传特性。基因组中每条染色体由一个 DNA 分子组成,而每个 DNA 分子则包含很多基因。迄今已有几百种生物基因组完成了全序列分析,表明基因组大小和所含的基因数目与生物进化程度呈现不完全对应的关系。随着研究的进展,基因结构和类型也愈加丰富,如断裂基因(split gene)、重叠基因

(overlapping gene)、可移动基因(movable gene)、超基因(supergene)和基因家族(gene family)及假基因(pseudogene)等等。在人类基因组和多种生物基因组测序完成的基础上,分子遗传学的研究任务不仅是研究单个基因的结构与功能,而且还要关注生物整个基因组的结构与功能,从全新的视角探讨遗传与变异、结构与功能以及健康与疾病等的分子机制(详见第2章)。

(5) 研究生物遗传变异的分子机制 作为遗传物质不仅储存和传递遗传信息,而且应含有变异的分子基础,这是因为生命过程是一个遗传和不断变化的动态过程,从而生物得以进化。生物遗传变异的分子基础正是由于在进化过程中决定性状的基因和基因的载体染色体会发生改变。无论是基因发生突变,还是染色体发生改变,都会使性状的遗传方式和遗传下来的性状发生相应的改变。基因突变、染色体数目和结构的变异以及遗传重组是导致基因或基因组不断变化的主要因素,但其分子机制是各不相同的。此外,基因组中转座因子(transposable element)的位置移动和插入也带来各种效应的突变。三核苷酸扩增急剧增加的动态突变(dynamic mutation)也是导致基因组不稳定的重要因素之一。尽管引起遗传变异的各种类型的机制不完全相同,但其共同的分子基础是相同的,那就是改变了基因组DNA序列(详见第5章)。但表观遗传变异(epigenetic variation)则是另一种类型的遗传性的变化,它是指基因组中的DNA序列不发生改变,而在基因表达时发生的可遗传的变化,造成基因产物的改变,最终导致表型的改变,如碱基的甲基化(methylation)、基因组印记(genomic imprinting)、RNA编辑(RNA editing)等。由此进一步拓宽了遗传变异的研究内容(详见第9章)。

(6) 研究基因控制细胞分化和发育的分子机制 发育是生物的共同属性,也是生物进行遗传和变异的全过程(详见第7章)。一个个体从受精卵开始,具有相同基因组的细胞为何可分化出各种类型的细胞?细胞又怎样按照一定形式形成组织和器官?基因与细胞分化和发育的关系如何?Morgen在1928年早已指出:“明了基因对于发育中的个体如何发生影响,毫无疑问地将使我们对于遗传的概念进一步扩大,对于目前不了解的许多现象也多半会有所阐明。”随着人类基因组计划和一些重要动植物基因组计划的完成,功能基因组的研究已成为重要课题,而发育生物学正是研究基因组功能的主要领域(详见第11章)。当今发育生物学的研究真正在分子水平以基因和基因组为对象,研究各类生物基因组在发育不同的时空中基因差异表达和调控的网络,以揭示细胞分化、发育、生殖、遗传、变异以及癌变和衰老等发育生物学全过程的分子机制,进一步认识许多重大生物学问题的分子机制。研究表明基因控制着发育的图式形成(pattern formation),发育是在特定的时间和空间基因差异表达(temporal-spatial differential gene expression)的结果,是生物体基因型与内外环境因子相互作用并逐步转化为表型的过程(图1-2)。

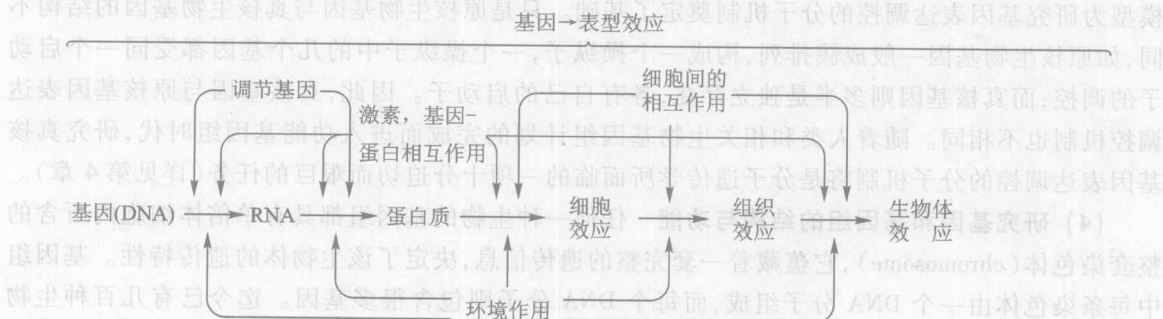


图1-2 基因影响生物表型效应的复杂途径(引自 Sunstad et al, 2003)

1.2 分子遗传学的建立

1.2.1 物理学的渗透

众所周知,生命活动是一种物质运动形式,而遗传与变异是重要的生命活动之一,所以遗传与变异也应该是一种物质的运动形式,同样遵循着支配自然界物质运动的基本规律。如生命活动同样服从于物理学的热力学定律,为此吸引了一批物理学家关心生物学问题,并跨入生物学研究领域。其中最为突出的是:① Bohr 于 1931 年发表了“光和生命”的演讲,指出:“我想大家都同意牛顿的观点,即科学的真正基础是相信在同样条件下,自然界总是显示出同样的规律性。因此要是我们能像分析原子现象那样,深入分析生命有机体的机制,我们就很难在有机体和无机物的性质上找到差别。”他还明确指出:“当生物学应用新的概念和研究方法时,就能上升到新的认识水平。”(《自然》杂志 131 卷)。② 作为洛克菲勒基金会资助的物理研究员,Delbruck 于 1931 年来到丹麦哥本哈根,在 Bohr 的影响下决定投入生物学研究。1937 年他到美国加州理工学院开始进行噬菌体研究。同时,他与一些物理学家和化学家组建成有名的“噬菌体研究组”。他们主要研究生物信息的物理学基础,所以被称为“信息学派”。③ Schrodinger 是量子力学的创始人之一,同样对生物的复制和繁殖问题感兴趣。1943 年 2 月,他在爱尔兰都柏林三一学院做了著名的讲演《生命是什么?》(What is life?),并于 1944 年公开出版。该书可以说是分子生物学的雏形,他也是第一个用热力学和量子力学理论来阐明生命本质的人。

在《生命是什么?》一书中作者第一个用物理和化学的语言对生命进行了系统的描述,把信息理论、量子力学的概念引入生命过程之中,为现代分子生物学提供了完整的理论框架,在分子水平上沟通了物理学、化学和生物学之间的关系,产生了深远的影响,并吸引了大量的学者,包括 DNA 分子双螺旋结构模型创立者 Crick 和 Watson 都是在该书的影响下进入了遗传学研究领域,并为分子生物学和分子遗传学的诞生奠定了基础。

1.2.2 遗传物质是核酸

早在 1865 年, Mendel 就提出生物性状的遗传是受细胞内颗粒性遗传因子(inheritance factor)所控制。40 余年后, Johannsen 于 1909 年提出用“基因”(gene)一词代替 Mendel 提出的“遗传因子”。但当时对遗传因子或基因的认识只是一种逻辑推理产物。基因究竟是什么物质分子还不得而知。在 Mendel 遗传规律被重新发现两年后的 1902 年, Sutton 和 Boveri 在观察蝗虫减数分裂过程中的染色体时发现,遗传因子的行为与配子形成和受精过程中的染色体行为是完全平行的,并由此认为控制性状的遗传因子位于细胞核内染色体上,从而奠定了遗传的染色体学说(chromosome theory of inheritance)的基础。由于他们的工作把细胞学和遗传学结合起来,极大地推动了染色体理论向基因论的过渡进程。

Morgan 进一步证明了基因是以线性形式排列在染色体上,并在染色体上占有一定位置,基因的传递同基因所在的染色体的传递是连锁的,由此提出连锁定律。连锁定律与孟德尔的分离定律和独立分配定律合称为遗传学三大定律。连锁定律不仅将一个特定基因与特定染色体的特定位置联系起来,还指出基因如同果蝇唾腺染色体上横纹那样直线地排列在染色体上。这样,基因就不再是代表性状的抽象逻辑推理的产物,而是在染色体上占有一定空间的实体,并且是物质的。在此基

础上, Morgan 于 1926 年发表了著名的《基因论》, 论述了基因在上下代之间的传递规律, 认为基因控制相应的性状, 基因可以发生突变、交换和重组。由此提出, 基因既是一个功能单位、一个突变单位, 也是一个交换单位的“三位一体”的概念; 并预言基因是一个有机的化学实体。显然, 《基因论》对遗传学的研究产生了巨大的影响, 也为分子遗传学的建立奠定了理论基础。

正是在这种背景条件下, 一些物理学家和化学家关注《生命是什么?》的同时, 微生物学家和生物化学家也向遗传学靠拢。早在 1923 年, 英国医生、生物化学家 Garrod 根据对人体的一种先天性代谢疾病尿黑酸症的研究, 发现该病是一种隐性遗传病, 当这种纯合隐性基因存在时就不能产生尿黑酸氧化酶 (homogentisic acid oxidase), 使蛋白质代谢产物不能最终分解为二氧化碳和水而积累于血液中。这样, 一部分尿黑酸多聚物沉积于软骨及其他结缔组织中, 使患者年老时发生褐黄病; 还有一部分尿黑酸则随尿液排出, 暴露于空气中氧化成黑色素, 使尿迅速转为黑色, 故称为尿黑酸症 (alkaptonuria)。他认为这种疾病是由于单个基因发生突变后产生一种不具功能的产物, 从而导致功能障碍 (图 1-3)。Garrod 的这种一个突变基因决定一种代谢障碍 (one mutant gene-one metabolic block) 的观点在当时并未受到关注。直到 1941 年, Beadle 和 Tatum 对粗糙脉孢菌 (*Neurospora crassa*) 的生化突变型进行研究时才发现了 Garrod 的工作, 明确提出了“一个基因一个酶” (one gene-one enzyme) 的理论。后来将“一个基因一个酶”改为“一个基因一种多肽”

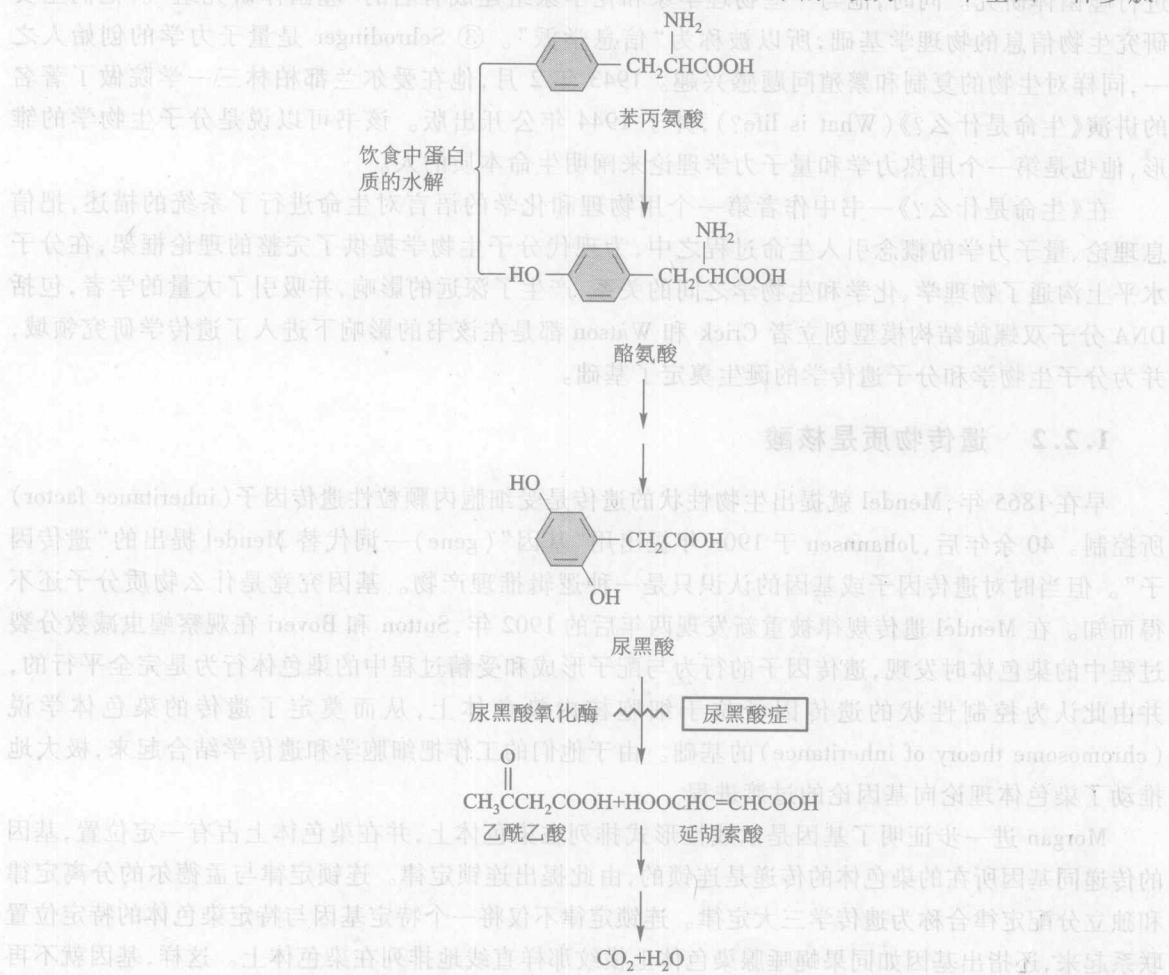


图 1-3 基因突变所引起的人类尿黑酸症 (引自 Sunstad et al, 2003)

(one gene-one polypeptide)。这表明基因是通过控制多肽的合成而影响生物遗传性状的发育和表达(图1-4)。

野生型的孢子被照射后, 菌株与野生型的杂交

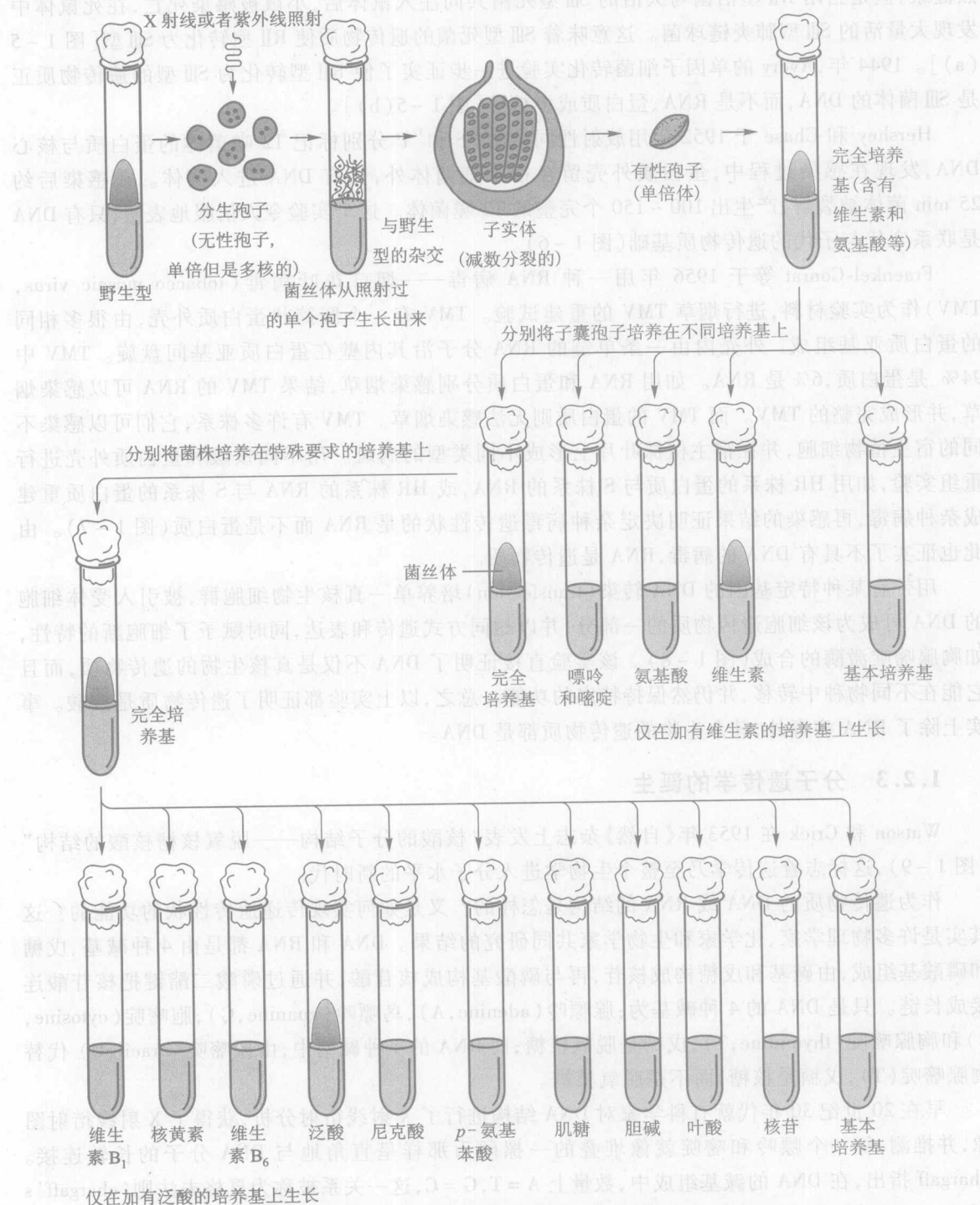


图1-4 Beadle和Tatum的实验图解(引自Sunstad et al, 2003)