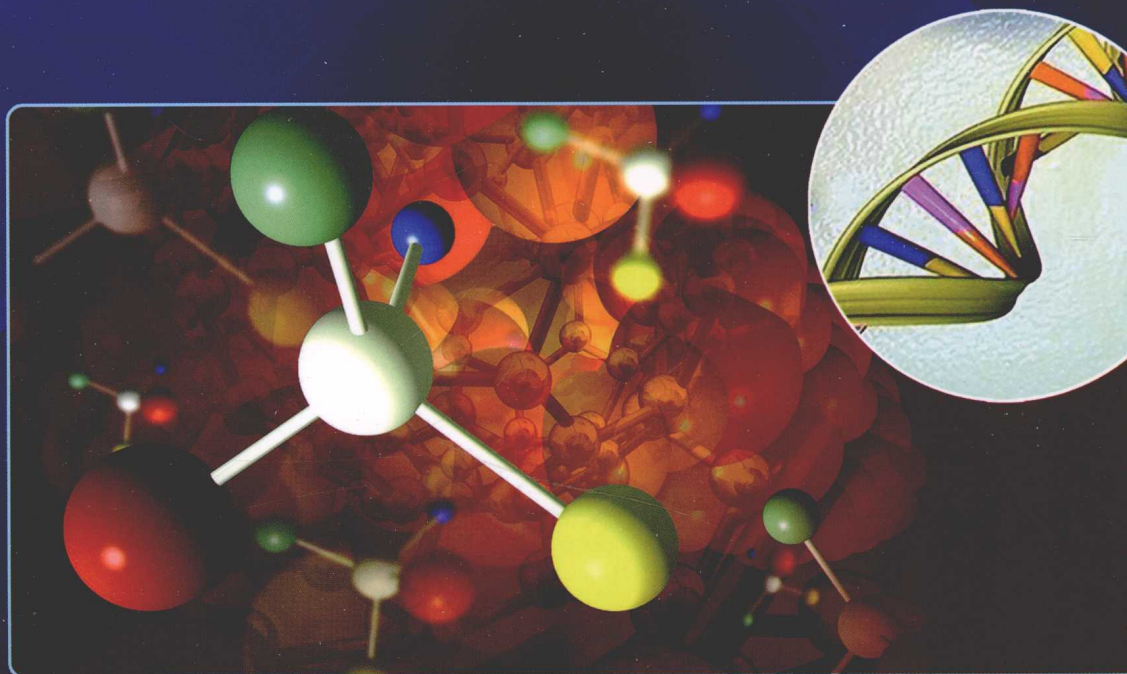


 21世纪高等院校教材

系统生物学

张自立 王振英 编著



1-43

0



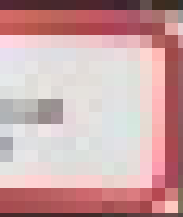
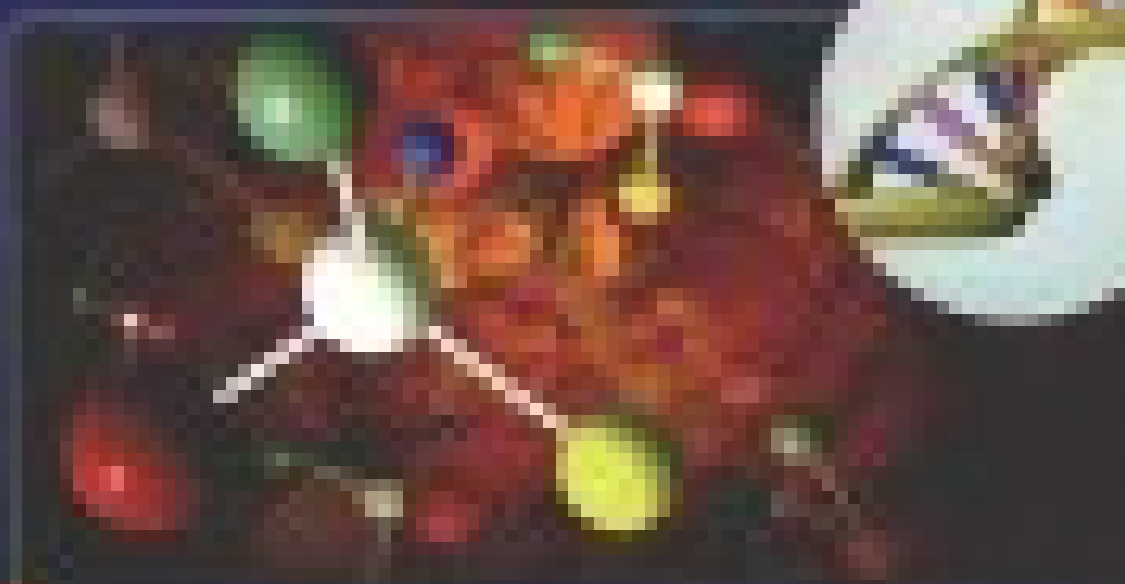
科学出版社
www.sciencep.com



中国科学院植物研究所

系统生物学

作者: 王保刚 编著



中国科学院植物研究所
植物研究所

21 世纪高等院校教材

系统生物学

张自立 王振英 编著

科学出版社

北 京

内 容 简 介

系统生物学是现代生物学新兴分支学科之一。它整合了各层面的生物信息数据,建立各种数学模型进行仿真实验,进而定量阐明和预测生物功能、表型及行为,它已成为当今生命科学的重大前沿领域之一。本书概述了系统生物学的基本概念和基本内容,介绍了基因组学、转录组学、蛋白质组学、糖组学、代谢物组学、相互作用组学、表型组学、数学建模与仿真、序列比对与数据库搜索、分子进化模型与系统树的构建等。

本书内容新颖简要,适于有生物学基础知识的本科生,以及研究生、科研人员学习,期望引导他们用新视觉认识生命科学。

图书在版编目(CIP)数据

系统生物学/张自立,王振英编著. —北京:科学出版社,2009

21 世纪高等院校教材

ISBN 978-7-03-024984-5

I. 系… II. ①张…②王… III. 生物学-系统科学-高等学校-教材
IV. Q111

中国版本图书馆 CIP 数据核字(2009)第 116809 号

责任编辑:王国栋 周 辉 席 慧/责任校对:张怡君

责任印制:张克忠/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>

北京市农林印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2009 年 7 月第 一 版 开本:B5 (720×1000)

2009 年 7 月第一次印刷 印张:10 3/4

印数:1—3 000 字数:202 000

定价:23.00 元

(如有印装质量问题,我社负责调换)

前 言

21 世纪的生物研究正从分子生物学走向系统生物学, 由精细的分解研究转向系统的整体研究。研究者整合各种生物信息的实验数据建立数学模型, 并通过实验验证完善模型, 使其能定量和预测生物系统的表型、功能和行为, 这样一门崭新的现代生物学的分支学科被称为系统生物学。系统生物学在形成过程中引起了数学、物理、化学、生物、医学及信息和计算机科学等学科领域众多科学家的广泛关注并参与其中, 现已成为一门多学科交叉的新型学科。

当前世界各地已纷纷成立了研究系统生物学的机构。例如, 2000 年 1 月美国科学家 Hood 率先创建了世界上第一个系统生物学研究所 (Institute for Systems Biology, ISB), 随后日本也建立了研究所。2003 年 9 月哈佛大学和麻省理工学院设立了系统生物学研究中心, 同年 12 月我国上海交通大学与中国科学院上海生命科学研究院合作成立了我国第一个系统生物学研究机构, 随后中国科学院生物物理研究所成立了系统生物学研究中心, 清华大学生物系设立了生物信息学与系统生物学研究所。自 2000 年以来, 系统生物学的国际会议十分活跃, 每年均组织交流活动。2003 年 3 月, *Science* 刊载了系统生物学专集, *Nature* 上也经常有相关论文发表。我国生物领域的学术刊物自 2002 年以来也陆续刊载介绍国外系统生物学发展的一些综述与专论。此外, 一些国家政府和大型制药公司已出资大力资助系统生物学的研究项目。例如, 2002 年美国能源部启动了 21 世纪系统生物学技术平台, 以推动环境生物技术和能源生物技术产业的发展。世界 6 大制药企业组成了以系统生物学技术为基础的新药研发系统的联合体。2003 年 9 月 15 日美国国立卫生研究院 (NIH) 的官方网站 (<http://www.nigms.nih.gov/news/releases/complex-centers-2003.html>) 发布了一条消息: NIH 将在此后 5 年中设立两个总量约 3 亿美元的特别基金, 用于资助哈佛大学和麻省理工学院各自建立一个新型的研究中心, 使得来自不同领域的科学家能协同工作、共同参与系统生物学研究以解决生物医学问题。

系统生物学是以整体性研究为特征的一门大科学, 其指导思想是强调“整体论”, 力图摆脱传统的“还原论”。这对半个世纪以来生物学的传统科研思维方法提出了极大的挑战, 不可避免地会出现多种争议。正确对待争议对发展学科至关重要。国际商业信息和技术咨询公司 (IDC) 的生命学科分析员 Zimmerman (2004) 对美国普林斯顿大学的 David Botstein 教授 (人类基因组计划奠基者之一、系统生物学研究倡导人之一) 专访时提出: “如何评价传统的生物学研究方

法与整合型研究方法?” Botstein 教授的回答十分值得我们深思,他说:“不能说谁好谁坏。只能说这种整合型研究方法是不同的,是新的。它只是基因组序列和相关技术带来的一个机遇,该机遇使得我们能够同时考察许多问题,并在计算机的帮助下同时分析和理解所有的事情。所以它与传统的方法是完全互补的。但是,总会有一些摩擦存在,因为新的东西倾向于吸引一些人,而在新、旧之间总是存在人们能够理解的竞争。”

为了使广大的生物学本科生和青年科研工作者能及时了解这门新兴学科概况,我们编著了这本《系统生物学》,以期使仅有生物学基础知识的青年读者对这门交叉学科有初步的了解,也期望能拓宽读者的视野,引导他们学习用新视觉去认识生命科学,引发他们对新学科的兴趣,思考一些新的学术前沿问题。

本书在撰写过程中引用了有关此学科的大量文献,我们特向文献的原作者致以衷心的感谢。

本书得到天津市细胞生物学重点学科建设经费资助,谨此致谢。

限于我们的学识水平,书中存在的不妥与错误之处恳请读者批评指正。

张自立 王振英

2008年6月

目 录

前言

第 1 章 系统生物学概况	1
1.1 从分子生物学到系统生物学	1
1.1.1 分子生物学的诞生及发展	1
1.1.2 “基因决定论”和“还原论”的局限性	2
1.1.3 转向整体论新潮流	4
1.1.4 系统生物学的产生和发展	5
1.2 系统生物学的定义和研究内容	7
1.2.1 系统生物学的定义	7
1.2.2 系统生物学的研究内容	8
1.3 系统生物学的研究	9
1.3.1 系统生物学的基本工作流程	9
1.3.2 系统生物学的研究方法	10
1.4 系统生物学的应用前景	12
主要参考文献	13
第 2 章 基因组学	15
2.1 基因组学的提出及其任务	15
2.2 人类基因组计划	16
2.2.1 人类基因组计划的研究目标及技术路线	16
2.2.2 人类基因组计划的作图	17
2.2.3 人类基因组计划的测序	18
2.2.4 人类基因组计划的信息处理	21
2.2.5 人类基因组研究计划进展	24
2.3 基因组学及其分支学科	25
2.3.1 功能基因组学	25
2.3.2 比较基因组学	26
2.3.3 药物基因组学	28
主要参考文献	29
第 3 章 转录组学	31
3.1 转录组及转录组学	31

3.1.1 转录组及转录组学的定义	31
3.1.2 转录组学的研究内容	32
3.2 转录组学的研究方法	32
3.2.1 高通量 mRNA 表达分析技术	32
3.2.2 基因表达系列分析技术	33
3.2.3 转录物编目的研究方法	35
3.2.4 绘制动态转录物图的研究方法	36
3.2.5 转录物调节网络	37
3.3 对转录物研究的新突破	37
3.3.1 转录物的多样性	37
3.3.2 非编码 RNA 的类型和功能	38
主要参考文献	41
第 4 章 蛋白质组学	43
4.1 蛋白质组学的产生	43
4.2 蛋白质组及蛋白质组学的概念	43
4.3 双向凝胶电泳	44
4.3.1 双向凝胶电泳 (2-DE) 原理	45
4.3.2 图像分析与数据库构建	46
4.4 生物质谱技术	47
4.4.1 种类及其原理	47
4.4.2 肽质量指纹谱鉴定技术 (PMF)	48
4.4.3 肽序列标签串联质谱技术 (PST)	48
4.4.4 翻译后修饰蛋白质的鉴定	49
4.5 蛋白质组数据库	50
4.6 蛋白质芯片技术	50
4.6.1 蛋白质芯片的制备	50
4.6.2 靶蛋白与捕捉分子结合情况检测	51
4.7 分析蛋白质-蛋白质相互作用的酵母双杂交系统	53
4.7.1 酵母双杂交系统的基本原理	53
4.7.2 酵母双杂交系统的改进	53
4.8 蛋白质组研究进展	55
4.8.1 病毒蛋白质组研究	55
4.8.2 细菌蛋白质组研究	55
4.8.3 酿酒酵母蛋白质组研究	55
4.8.4 多细胞生物蛋白质组研究	58

主要参考文献	59
第5章 糖组学	61
5.1 糖组与糖组学的研究内容	61
5.2 糖组学在生命科学中的意义	61
5.2.1 蛋白质组学必须面对糖蛋白	61
5.2.2 糖蛋白的定义	61
5.2.3 聚糖和糖蛋白的生物学作用	62
5.3 糖组学的研究方法	64
5.3.1 对 2-DE 分离糖蛋白结合质谱技术的改进	64
5.3.2 聚糖分子的微阵列技术	65
5.3.3 用敲除基因及转基因技术研究聚糖分子引起的表型变化	66
5.4 糖组学的国际合作和数据库	67
主要参考文献	69
第6章 代谢物组学	71
6.1 代谢物组学的定义和研究任务	71
6.1.1 代谢物组学的定义	71
6.1.2 代谢物组学的研究任务	71
6.2 研究代谢物组学的意义	72
6.2.1 代谢物组学是基因组学和蛋白质组学的补足	72
6.2.2 代谢物组学在医药界的应用	72
6.3 代谢物组学的研究方法	73
6.3.1 代谢物组的研究技术及其原理	73
6.3.2 用于代谢物组研究技术的比较	80
6.3.3 代谢物组分析的技术路线	81
6.4 代谢网络的研究	82
6.4.1 代谢网络的结构特征	82
6.4.2 用“整合”与“干扰”研究代谢网络	84
主要参考文献	88
第7章 相互作用组学	90
7.1 相互作用组学的研究方法	90
7.1.1 大规模蛋白质-蛋白质相互作用研究技术	90
7.1.2 大规模遗传学相互作用研究技术	93
7.2 蛋白质相互作用网络	94
7.2.1 丙型肝炎病毒 (HCV) 的蛋白质相互作用	94
7.2.2 病原菌幽门螺杆菌的蛋白质相互作用	96

7.2.3 酵母的蛋白质-蛋白质相互作用网络	96
7.2.4 果蝇的蛋白质-蛋白质相互作用网络	99
7.2.5 线虫的蛋白质-蛋白质相互作用网络	100
7.2.6 人类的蛋白质-蛋白质相互作用网络	101
主要参考文献	102
第8章 表型组学	104
8.1 什么是表型组学	104
8.1.1 基因型与环境的相关及互作	104
8.1.2 表型和表型组学	105
8.2 从基因组到表型组系统研究的方法	105
8.2.1 从大肠杆菌和酵母代谢缺失菌株预测生长表型	105
8.2.2 建立一种人类表型组——基因组的网络联系	107
8.2.3 微阵列技术在人类表型组、基因组和环境组系统研究中的应用	108
8.3 从基因组到表型组研究有关的数据库	110
主要参考文献	111
第9章 数学建模和仿真的基础知识	113
9.1 系统模型	113
9.1.1 什么是系统	113
9.1.2 什么是模型	115
9.1.3 生物系统中生化反应网络的数学描述	116
9.1.4 生物系统中的质量作用动力学模型	118
9.1.5 生物系统中有关细胞信号转导的建模	119
9.2 系统仿真	122
9.2.1 什么是系统仿真	122
9.2.2 系统仿真软件和相关数据库	123
9.2.3 系统生物学采用仿真技术的实用成果	124
9.3 实例：微生物细胞的建模与仿真	125
9.3.1 微生物数学模型的种类	126
9.3.2 微生物细胞的建模	127
9.3.3 用于微生物细胞模型的仿真平台	129
主要参考文献	131
第10章 序列比对和数据库搜索	134
10.1 数据库中序列表示的格式	134
10.1.1 FASTA (或 Pearson) 格式举例	134
10.1.2 GenBank flatfile 格式举例	135

10.2 序列比对·····	138
10.3 网络比对·····	140
10.3.1 成对网络比对研究 ·····	140
10.3.2 网络对位排列的算法 ·····	141
10.4 数据库中序列相似性检索·····	141
10.4.1 FASTA 程序 ·····	141
10.4.2 BLAST 程序 ·····	142
10.5 用隐马尔可夫模型预测新基因·····	143
10.5.1 隐马尔可夫模型 ·····	143
10.5.2 用隐马尔可夫模型预测新基因的举例 ·····	144
10.6 用人工神经网络预测蛋白质二级结构·····	145
10.6.1 简单神经网络模型 ·····	146
10.6.2 多层神经网络模型 ·····	146
主要参考文献·····	147
第 11 章 分子进化模型与系统树的构建 ·····	149
11.1 蛋白质编码序列进化·····	149
11.1.1 血红蛋白 α 链的进化距离和氨基酸替代率的估计 ·····	149
11.1.2 氨基酸的替代矩阵 ·····	151
11.2 DNA 序列的进化 ·····	152
11.2.1 核苷酸替代数的估计 ·····	152
11.2.2 Kimura 模型 ·····	153
11.3 系统树的构建·····	154
11.3.1 距离法·····	154
11.3.2 最大简约法 ·····	156
11.3.3 最大似然法 ·····	157
11.3.4 分子系统树的检验 ·····	157
11.3.5 对分子系统树的争议 ·····	158
11.4 分子系统发育软件·····	159
主要参考文献·····	159

第 1 章 系统生物学概况

生物体是一个多层次、多功能的复杂结构体系，研究其运动变化规律的科学叫做生物科学或生命科学。长久以来，科学家从群体、个体、细胞、分子等不同层次，以及形态解剖、生理生化、遗传发育、免疫、种群、进化等不同侧面探索着生命的运动规律，使生命科学迅猛发展，已从根本上改变了它在自然科学中的地位和作用。

当代生物技术也成为世界上最令人瞩目的高新技术之一，生物技术产业将成为 21 世纪的主导型产业之一，是许多国家产业结构调整的战略重点，人们把世界上日益严重的人口、环境、食物、资源等与人类生存和发展密切相关的诸多重大社会问题的解决，寄希望于生命科学与生物技术的进步。

回顾历史，20 世纪生物学经历了由表及里、由宏观到微观的发展历程，21 世纪远见卓识的科学家又试图把分子生物学的时代推向系统生物学时代。

1.1 从分子生物学到系统生物学

1.1.1 分子生物学的诞生及发展

1953 年 Watson 和 Crick 建立了 DNA 双螺旋模型，它标志着生物学迈进了分子生物学新时代。1958 年 Crick 对生物大分子 DNA 和蛋白质的相互关系提出了中心法则 (central dogma)，即 DNA 上的遗传信息可以通过复制传递给下一代的 DNA 分子，也能通过转录传递到 RNA，最后经翻译传递至蛋白质分子，简写成 DNA→RNA→蛋白质。该法奠定了分子生物学的理论基础。围绕中心法则的研究，并在它的影响下人们似乎在分子水平上证实了“一个基因决定一种酶”的传统观念。在此后 30 多年中，生命科学的研究热衷于把复杂的生物体层层分解，从个体到器官，从器官到组织，从组织到细胞，再从细胞到分子，最终力求在基因水平上寻找个体生命现象的根源。生命科学研究中几乎呈现出以基因为中心的势态，同时对各种基因的研究也侧重于对单个基因的寻找、分离和克隆。20 世纪 70 年代出现了基因工程技术，极大地加速了分子生物学的发展并带动了生物制药产业的崛起。最早的基因工程制药产品有人胰岛素 (Eli Lilly 公司, 1982)、人生长激素 (Genentech 公司, 1985) 和红细胞生成素 (Amgen 公司, 1989)，分别治疗糖尿病、侏儒症、恶性贫血等，在制药产业中产生了巨大影响。从 1982 年到 1989 年基因工程制药的平均开发期为 5 年，而同期的化学合

成新药期为 11 年。基因重组技术制药在第一个 10 年中有 13 种新药进入市场,第二个 10 年的中期达到了高峰期,仅 1996 年就有 11 种重组蛋白质新药被批准上市。从此分子生物学和生物技术跻身于发展最快、应用最广、潜力最大、竞争最激烈的科技领域之列。

1.1.2 “基因决定论”和“还原论”的局限性

随着分子生物学迅猛发展,人们对生物体的认识逐渐从表型深入到分子层次,毫无疑问这是人类认识上的重大进步。但是因此注重把复杂生物体层层分解到分子水平,侧重研究单个分子的功能就难免会陷进“还原论”(reductionism)和“基因决定论”(determinism)的误区,致使某些研究处于困境。例如,在高等动物、植物转基因研究中常出现外源目的基因表达沉默、表达水平太低、引起相邻其他基因突变、后代遗传特性不稳定等困扰,造成了转基因动植物产业化的瓶颈。现实情况显示基因工程制新药的速度从 20 世纪 90 年代高峰期后至今一直在减慢,大多数原因与上述瓶颈有关。Snoep 等(1995)报道研究者试图使微生物 *Zymomonas mobilis* 的糖酵解酶基因过表达提高丙酮酸产量,但结果却与预期目标相反,由于此基因过量表达引起了蛋白质合成的抑制,菌体生长减慢,导致丙酮酸产量下降。Leggewie 等(2003)报道:淀粉合成途径中编码淀粉合成酶的基因过表达,并没有使马铃薯块茎增大,试验结果反而出现块茎变小。作者认为这一结果暗示代谢过程是相互连接的网络(network)和多邻近途径(neighbourhood)相互作用。还原主义过于强调把一个复杂整体分解成单个的器官、组织、细胞,直至单个基因或蛋白质分子入手逐个研究的方法,虽然有利于获得对单个分子的详细信息,但事实表明单个分子脱离了整个系统的协调作用就难以发挥应有功效。众所周知人类某些基因对于某些疾病发生起重要作用,但越来越多的研究结果表明,疾病并非由单个基因决定。例如,妇女细胞中 *BRCA1* 基因发生突变,该妇女到 60 岁后约有 70% 的概率会患乳腺癌。那么还有 30% 的女性为什么未患癌症? 研究结果表明其他基因的表达产物对此疾病的影响不能忽视,人体内还存在多种抑癌基因,它能抑制癌变。此外患者的生活环境也会起一定的影响作用。美国科学家 Dulbecco(1975 年因发现肿瘤病毒而获得诺贝尔奖)多年致力于癌基因研究,由科研实践和经验得出:要彻底阐明癌症的发生、演变侵袭和转移的机制,必须对人体细胞的基因组(即全部基因)在整体上进行全序列分析。他在 *Science*(1986)发表的《癌症研究的转折点:测序人类基因组》一文中呼吁科学家必须转向人类细胞中全部基因的研究。

受中心法则的影响,似乎遗传信息的流向是严格按单程路线传递,即 DNA→RNA→蛋白质。基因决定论也强调 DNA 序列通过转录、翻译决定了蛋白质的氨基酸序列,为此生物体内只有在 DNA 序列发生了变化以后才能出现可

遗传的变异。但是大量分子生物学试验结果表明：基因表达与蛋白质水平之间并不完全一致，借助 mRNA 交替剪接、编辑或翻译起始与终止位点变化或翻译后蛋白质前体的自我剪接和加工修饰等多种途径，致使 1 个基因能编码多种不同的蛋白质。近年来发现的表观遗传（epigenetic inheritance）现象说明了即使基因的 DNA 序列并未改变，而表型也可能出现可遗传的变化。这类研究已有大量报道。Vegelman 和 Freinberg（1983）发现癌症患者细胞内表观遗传现象导致癌症的遗传。此后 10 年中表观遗传现象的研究在癌变病理学、诊断和预防中发挥了一定作用。Greener（2005）在《癌症表观遗传学进入主流地位》一文中引述了英国剑桥 Babraham 研究所 Reik 的见解，Reik 认为：“大部分癌症的发生是由遗传因素和表观遗传因素相互作用引起的”。例如，细胞中抑癌基因的 DNA 甲基化（hypermethylation）程度过高使 *p16INK4a* 基因沉默导致多种癌症发生。Esteller 也发现 *MGMT* 基因甲基化过量，表达被抑制，细胞中 E-cadherin 蛋白减少，会促进癌症发病及转移；反之，如果原癌基因启动区的 DNA 甲基化不足会使得卵巢癌和乳腺癌发病率增高。目前已有足够资料表明表观遗传并非稀有现象。

中心法则强调蛋白质中的氨基酸序列由 DNA 模板上的核苷酸顺序决定。但是 Lipmann 等（1971, 1976, 1984）证实了存在以蛋白质模板合成肽链的事实，如抗生素多肽（antibiotic polypeptide）、谷胱甘肽、胞壁质交联肽等的合成是不以 DNA 为模板，而用多酶体系为模板，按一定顺序吸附特定的一种氨基酸并将它们逐一连接形成肽链。合成过程中研究者加入 DNase、RNase 处理，并不影响多肽的产生。20 世纪 90 年代 Mizukami 等（1992）、Jensen 等（1992）、Far-chaus 等（1998）都进一步证明微生物能以非核糖体途径合成包括脂肽在内的多种天然多肽，其装配由非核糖体肽合成酶（nonribosomal peptide synthetase）的多模块酶以硫模板机制完成。研究者用调控多模块顺序，有计划有目的地变换多肽中氨基酸顺序，达到提高产品表面活性剂的活性。此技术已用于生产实践。图 1.1 是脂肽 Surfactin 合成酶及合成过程示意图。

中心法则将 RNA（mRNA）作为传递信息和蛋白质合成（rRNA、tRNA）的中间因素，对 RNA 生物功能的多样性未能充分重现。2001 年中国、美国、日本、德国、法国、英国和美国 Celera 公司联合公布了人类基因组图谱及初步分析结果，指出：与蛋白质有关的基因仅占 2%，而编码非蛋白质的各种 RNA 基因的比例要大得多。随后不少科学家研究证明非编码蛋白 RNA（ncRNA）可以通过多种遗传机制调控 DNA 的结构、RNA 的表达及蛋白质的翻译和蛋白质的功能，在生物体正常生长发育中起重要作用，这些作用完全超越了中心法则的认识。

可见，“基因决定论”和“还原论”存在局限性。

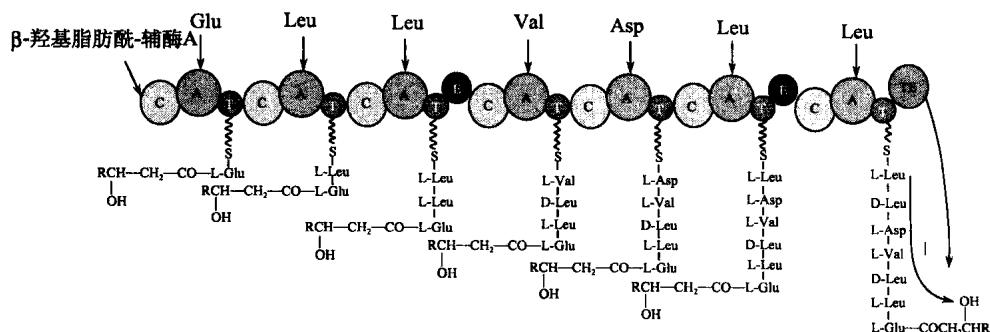


图 1.1 脂肽 Surfactin 合成酶及合成过程示意图 (Walsh et al., 1997)

1.1.3 转向整体论新潮流

近年来学术界对“基因决定论”和“还原论”提出质疑, Strohman (2001) 指出: 决定论忽视众多基因的相互作用, 基因和遗传背景的相互作用, 基因和环境的相互作用, 许多表观遗传的事实远远超越了决定论的视野。他认为基因的调节循环应该是一种“动态表观式 (dynamic epigenetics) 的图案”。Silverman (2004) 在《科学家》杂志发表了《重新考虑基因决定论》的文章, 他认为基因决定论在分子遗传学上的阐述离开了正确性和可靠性, 文章强调基因调节中有多种不确定因素 (uncertainty)。Mattick (2004) 认为细胞中很多种 RNA 和蛋白质可能同时相互调节遗传信息。Lolle 等 (2005) 在拟南芥细胞中发现大量超越基因组外的 RNA 负载非孟德尔式的遗传信息, 起着遗传反馈作用以补充孟德尔式遗传。Goodman 等 (2005) 认为复杂多细胞生物的基因调控受多个层次的相互制约: 有机体个体调节、细胞调节、转录调节、翻译调节和后加工调节等相互影响, 且各层次均离不开它们所处环境的作用。同时, 他们还提出了在整体水平上基因调控的模型 (图 1.2)。

Dulbecco (1986) 提出了进行人类基因组计划研究的建议, 客观上就是对“基因决定论”和“还原论”的传统思维的挑战, 当然会有不少反对意见, 经历一年的争论, 学术界才取得了共识。1987 年美国国会通过了人类基因组计划, 1990 年正式启动拨款。基因组学问世不久, 科学家开始了蛋白质组研究。2000 年报道了第一个完整的蛋白质组——支原体蛋白质组。2001 年国际人类蛋白质组组织 (HUPO) 宣告成立, 第二年召开了首届国际人类蛋白质组大会。几乎与此同时, 转录组学、糖组学、代谢组学和代谢物流率组学 (fluxomics) 也展开了大量的研究工作, 据统计, 近 10 年中在生命科学研究领域竟提出了近百种

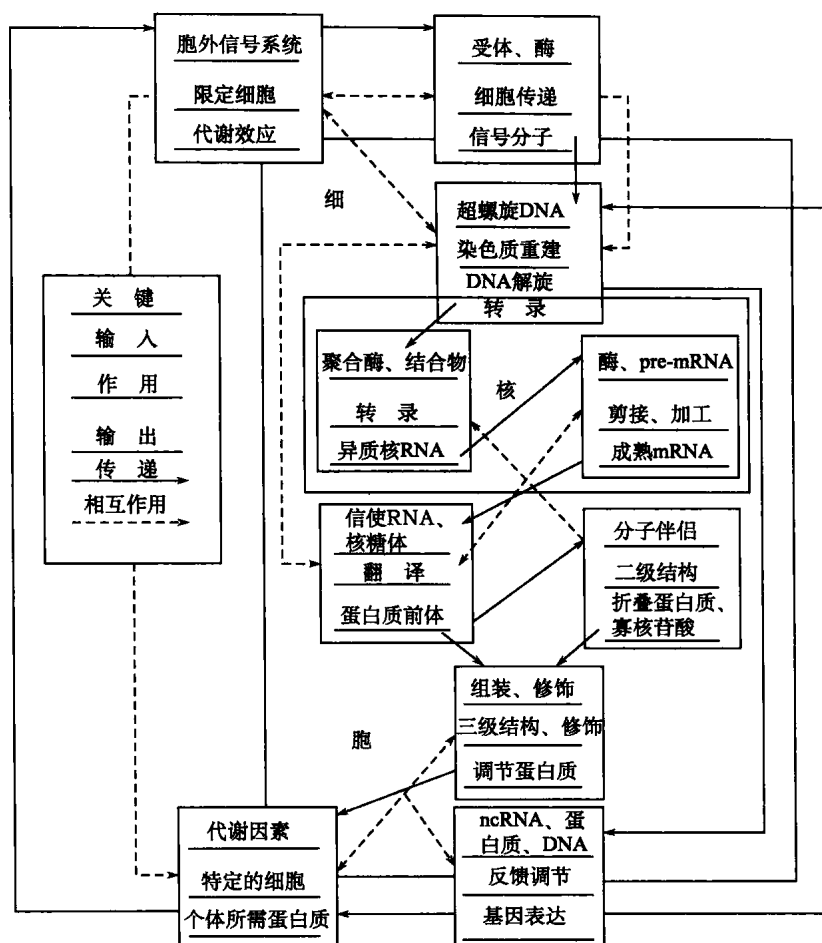


图 1.2 整体水平的基因调控模型 (Goodman, 2005)

“组学”研究，它意味着转向整体论的新潮流已经到来。

1.1.4 系统生物学的产生和发展

从人类基因组计划实施的第一天起，研究者就面临对每天 DNA 测序获得的海量数据进行收集、储存和分析，这就必须与数学、物理、化学、信息科学和计算机技术相结合、渗透、交叉和融合，否则就无法深入研究。如果追溯历史，把生物看做一个系统与数学和计算机技术结合共同研究生命规律，则早在 20 世纪 50 年代就已经开始。

Wiener (1948) 早已明确指出：生物系统和控制系统能用同样的科学方法进行研究。20 世纪 50 年代初，Hodgkin 和 Huxley 建立了神经细胞膜产生动作

电位膜电导变化的模型，揭示了神经电生理的内在机制，并可通过仿真预测神经电的一切基本现象。

1987年《数学生物科学》杂志还发表过《生物化学系统理论和代谢控制理论》一文，但是由于当时试验测量技术的通量有限，模型的数据不够充分，该文章未能受到生物学家应有的重视。随着人类基因组计划的实施，各种高通量测量技术的快速发展，收集到关于生物系统性能方面的大量数据，同时也随着计算机的发展，求解复杂系统数学模型的功能也越来越强，采用计算机技术对复杂生物系统进行数字仿真才日益为人们所重视和应用。

跨入21世纪的前夕，美国科学家Hood等(1999)创建了世界上第一个系统生物研究所(Institute for Systems Biology)，他指出过去30年中生物学家们已习惯于对个别基因或个别蛋白质进行研究，而系统生物学则用综合方法研究系统的所有元素并追踪这些元素在系统发挥功能时的行为。例如，要研究汽车怎样工作，传统生物学家会去孤立地研究各个部件包括变速器、点火装置、刹车等。而系统方法则首先规定系统的所有元素；然后研究在系统工作时各个元素如何同系统中其他元素相互作用；最后，系统方法要求建立一个数学模型，这个模型将定量描述系统的本质和属性，并能预测生物功能、表型和行为。

继Hood等创建了系统生物研究所后，系统生物学开始引起了广泛的关注。美国、日本、德国、韩国和中国等都相继建立了相关的研究机构，提出了研究计划。自2000年起国际上每年都有一次系统生物学会会议，分别在美国、日本和欧洲等地举办，下面列举了国际上系统生物研究机构 and 公司的简介(表1.1)。

表 1.1 系统生物研究机构和公司简介(资治科和孙之荣, 2005)

研究机构/公司	主页地址	目标和使命
Institute for Systems Biology	http://www.systemsbiology.org/	理解人类生理和疾病
Molecular Science Institute	http://www.molsci.org	预测细胞和器官响应基因缺失和环境变化时的行为
Watson School of Biological Sciences	http://www.cshl.org/gradschool/index.html	为生物新时代培养生物学家
ERATO-Kitano Systems Biology Project	http://www.symbio.jst.go.jp/ http://www.systems-biology.org/	生物系统建模与仿真模拟
Alliance for Cellular Signaling	http://www.cellularsignaling.org/	建立细胞中所有蛋白质的相互作用图谱
Computational Systems Biology Lab	http://csbl.bmb.uga.edu/resources/researchgroup.php	发展解决生物学问题的计算工具