

# 复杂的生命

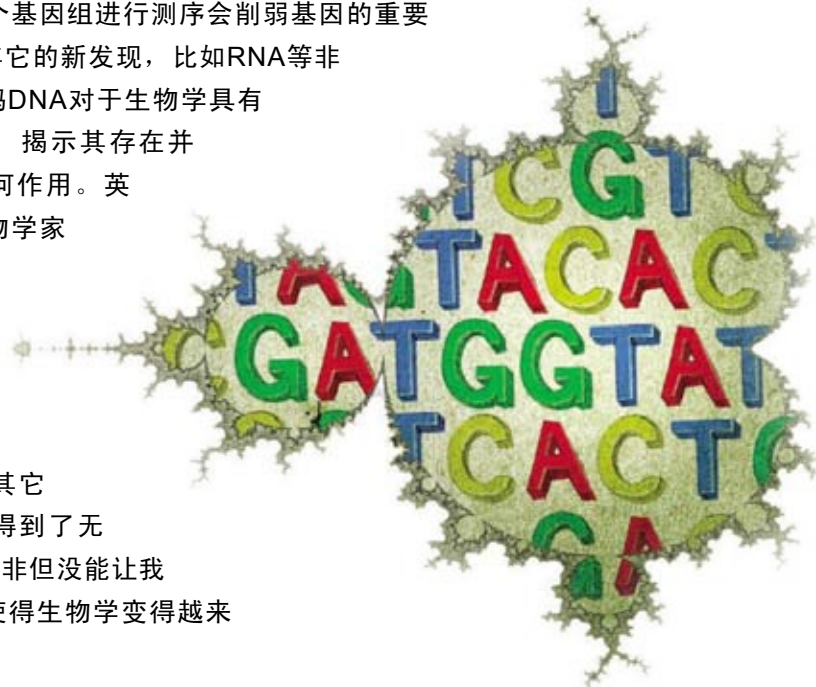
生物学家们研究得越深入，就越发发现生命的复杂。Erika Check Hayden将带领我们来看看有没有办法让生命看起来简单一点。

生物学一直以来都被很多人看作是一门简单的科学。人们认为生物学只是一门探索性的观察实验科学。20世纪初，就在Albert Einstein和Max Planck写出了数理方程，阐明了宇宙间的基本物理原理的同时，生物学家Ivan Petrovich Pavlov因为从狗流口水的现象中提炼出著名的条件反射和信号学说而获得了诺贝尔奖。

1953年，DNA的分子结构被发现，于是一切都改变了，生物学进入了分子时代，它变得量化性更强，也开始成为一门受人尊敬的科学。生物学家们开始着手探讨各种问题，比如物种进化和疾病起源等等。10年之后，开始逐渐兴起的人类基因组测序技术又将生物学往前推进了一大步。人们开始探索祖先的奥秘、破译进化的密码、发现疾病背后的分子机制，从而为找到更加精确的诊断方法和更加有效的个性化医疗措施打下了坚实的基础。人类基因组研究技术的出现，为成功描绘人类生物学蓝图带来了希望。

虽然这一天尚未到来，但是从某些方面来说，测序技术已经帮助我们阐明了某些问题。在人类基因组计划（Human Genome Project）开始以前，生物学家们估计人类基因组中至少包含有10万个编码基因。不过测序之后发现，人类基因组当中只有将近21,000个编码基因，尽管生物学家已了解其中部分基因的生物学功能，但对其它基因却毫无头绪。测序技术仅仅为我们打开了生物迷宫的大门，后面还有很多新的问题等待着我们去解决。

有人曾预测，对整个基因组进行测序会削弱基因的重要性，因为我们会有一些其它的新发现，比如RNA等非编码调控元件等。非编码DNA对于生物学具有非常重要的作用，然而，揭示其存在并未对阐明其功能起到任何作用。英国癌症研究所的细胞生物学家Mel Greaves指出，我们一厢情愿地认为基因组序列是一份清晰的蓝图，但实际情况并非如此。事实上，随着测序技术及其它技术的进步，我们确实得到了无数的数据，但是这些数据非但没能让我们更了解生物学，反而使得生物学变得越来



越复杂，而且其复杂程度正以几何倍数增长。随着研究的深入，生物学家们发现自己好像落入了曼德布洛特集（Mandelbrot set）一样。所谓曼德布洛特集就是由一个简单方程决定的一个空间，不过当你非常关注其中一个部分时你会发现它其实并非那么简单。

我们现在几乎能够获得任何信息并且对这些信息进行分析，因此生物学家们开始思考另一个重大问题——我们是否真的能够深入了解一个物种，或者一个细胞、一个细胞器或者一条信号通路，直到最终掌握各种细节呢？

如果我们非常熟悉各种输入信号、输出信号以及各种相互作用变量的信息，就能够凭借这些信息构建出一个预测模型。不过这个计划有多诱人还得看科学家们能做到什么程度，有人认为这已经可以足够了解生命的基本规律，但是另一些人却没有被越来越复杂的生命系统所困扰，他们转向去寻求另一个问题的答案。美国加利福尼亚大学伯克利分校（University of California, Berkeley）的生物化学家Jennifer Doudna这样说道：“这就好像在爬山，我们想不断往上攀登，但是爬得越高，就越会发现前面还有更多、更高的山峰等着我们。”

## 互联网一样的网络

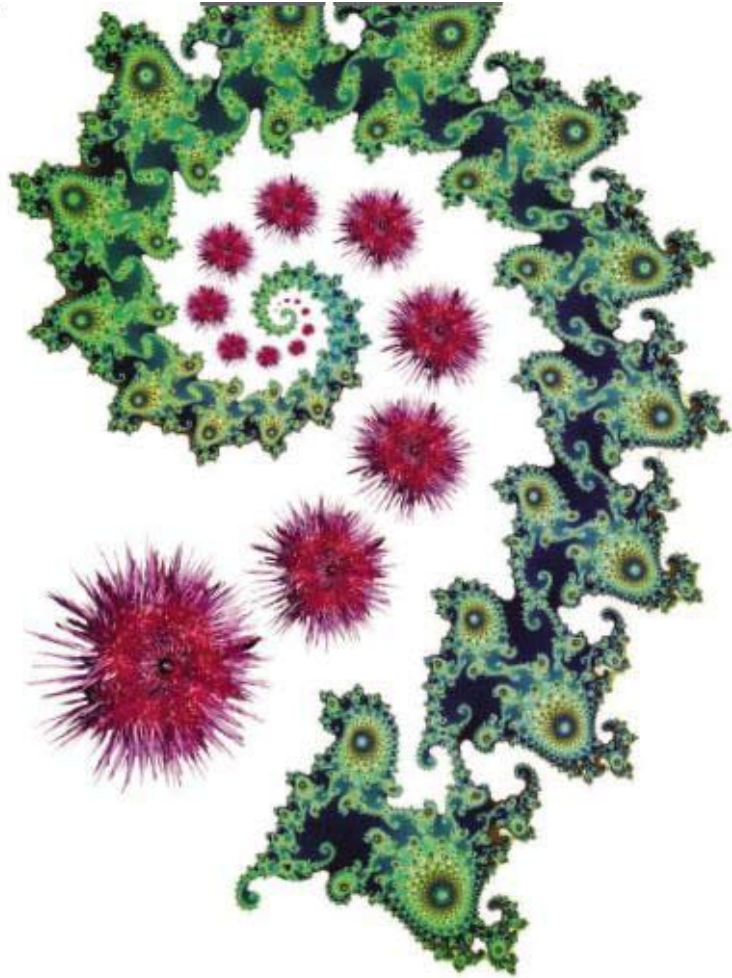
生物学家们曾经低估了生物学的复杂性。比如基因表达调控机制是在差不多50年前就已经或多或少地被阐明了。1961年，法国生物学家François Jacob和Jacques Monod提出了一种机制，他们认为调节蛋白通过与DNA相结合的方式调控基因的表达。5年之后，美国生物化学家Walter Gilbert发现了乳糖抑制子蛋白（lac repressor protein，该蛋白能够与DNA相结合来控制大肠杆菌的乳糖代谢），从而证明了上述假说。随后，科学家们不断发现了该机制的一些细节，大家相信已经掌握了事情的本质。在牛津大学出版社出版的1997年版的《基因6》（*Genes VI*）一书中这样写道：“基因表达调控这个难题的实质就是调控基因编码调控蛋白，而调控蛋白则通过与DNA上的特定位点相结合来控制基因的转录。”

在基因组生物学诞生仅仅十年之后，我们就发现了这一机制。现在生物学家们的关注点都已经转向了非编码DNA。这些过去曾被称作“垃圾DNA（junk DNA）”的“废物”，现在开始散发出迷人的光芒。来自国际合作项目DNA百科全书（Encyclopedia of DNA Elements）的科研人员发现，在基因组的某个部位中只含有部分的蛋白编码序列，大约有74%~93%的DNA能够被转录成RNA。大部分非编码DNA都能起到调控作用。他们还发现各种小RNA似乎也能在DNA水平和RNA转录水平上调控基因的表达，不过具体的作用机制还不清楚，他们正在开始进行这方面的研究。美国宾夕法尼亚大学（University of Pennsylvania）的计算生物学家Joshua Plotkin说：“这些调控因子的存在表明我们对于基本生物机制，例如细胞是如何工作的等机制的了解还是非常肤浅的。”

即使是一个分子，其复杂程度也非常大。比如最早于1979年发现的p53蛋白，最初曾经被误认为是促癌因子（cancer promoter），不过不久人们就替它“平反”了。经过多年研究，现在p53蛋白已经是众人皆知的抑癌因子了。鉴于其具有促使损伤基因组的细胞死亡，从而抑制肿瘤生长的作用，p53还获得了“基因组卫士”的美誉。至今，几乎没有其它任何一种蛋白被关注（研究）的程度能够超过p53蛋白，人们甚至还召开过专门的p53蛋白国际会议以更好地交流大家对p53的研究进展。p53蛋白的故事说明，生物（哪怕是分子）的复杂性是远远超出我们想象的。

1990年，有好几个实验室都发现p53蛋白可以通过直接与DNA相结合来控制基因转录，这也为Jacob-Monod提出的基因表达调控模型提供了证据支持。不过，随着对基因表达调控机制的研究不断深入，我们又发现了p53蛋白的更多的作用。就在去年，日本科学家还报道他们发现p53蛋白能够为多种小RNA的加工处理过程提供帮助，而这些小RNA都具有调控细胞周期检查点的作用，这进一步拓展了p53蛋白的抑癌机制。

其实很久以前，我们就已经知道p53蛋白借助化学和遗传相互作用在细胞内的蛋白质网络中起到了中枢调控的作用。我们现在知道p53蛋白能够与基因组中的数千个位点相结合，这些位点中有很多都是远离蛋白编码基因的，距离可以达到数千个碱基。p53蛋白会影响细胞生长、死亡和DNA修复，还会影响细胞的结构。此外，p53蛋白可与各种蛋白质结合，通过这种方式来改变自身活性。这种蛋白间的相互作用还可以被各种化学修饰物，例如磷酸盐基团和甲基基团调控。p53蛋白经过可变剪接可以形成9种同源异构体，每一种异构体都具有自身独特的活性和相应的化学修饰物。生物学家们现在意识到，p53蛋白除了能够控制肿瘤生长之外还参与了其它众多的生理过程，比如生育过程、胚胎早期发育过程等。实际上，我们试图完全了解p53蛋白的努力看起来似乎是那么的苍白无力。现在，生物学家们已经开始把研究方向转到了研究p53蛋白网络方面。p53蛋白网络就好像插图表现的那样，简直就是一座蛋白间相互作用的迷宫（想了解更多关于p53的内容，请登陆[www.lifeomics.com](http://www.lifeomics.com)查看《生命奥秘》2010年07月刊专题——p53基因研究30年记）。



## 信息大爆炸

上面介绍的p53蛋白的故事只是一个例子，还有很多例子都能表明生物学家们对生物机制的了解是多么的肤浅，这些例子也让我们知道生物学家们的观念是如何发生转变的。不过这还得感谢基因组技术的大发展，让我们迎来了基因组时代。知道了p53蛋白的序列之后，计算生物学家们就能够在整个基因组当中寻找可能与p53蛋白相结合的部位了，也能预测出其它蛋白或化学修饰物与p53蛋白发生结合的位点。这进一步扩展了我们对蛋白间相互作用机制的了解，同时也得以让我们修正对信号通路所持有的旧观念。按照以往的观点，p53蛋白应该能够引发一系列的下游作用结果。

“我们起初认为信号通路是简单的、直线式的。不过现在我们发现，细胞内的信号通路实际上是复杂得多的网络式的结构。”加拿大多伦多大学（University of Toronto）的细胞生物学家Tony Pawson这样说道。

自从人类基因组计划开展以来，我们就获得了海量的基因数据。仅仅对一个基因进行克隆并开展研究就能在具有很高影响力的杂志上发表论文的日子已经一去不复返了。现在的研究方式都是对整个人类基因组进行测序，甚至对好几个基因组进行测序，然后进行比较。不过不幸的是，这种海量的信息并不总会给我们带来有意义的结果。

“在很多情况下，你都会进行高通量的研究，但其实有许多生物学问题是不需要这样做的。我们现在存在一个误区，那就是将收集数据简单地等同于开展研究了，认为这样做就会对生物机制了解得更清楚，这显然是错误的。”美国波士顿大学（Boston University）的生物工程学家James Collins说道。

另一种最新的理论——系统生物学（systems biology）可以帮助科研人员更好地了解生物的复杂性。我们希望将p53蛋白网络、整个细胞甚至一群细胞之间所有的相互作用进行一次分类，然后将数据输入计算机模型，这样可以从一个更为系统的角度来观察生物是如何运作的。

在后基因组时代，系统生物学家们开展了一系列的研究工作，希望能够构建酵母细胞、大肠杆菌、肝脏乃至虚拟人的系统模型。到目前为止，所有这些尝试都遇上了同样的问题——那就是完全没有办法将与每一个相互作用都相关的所有数据都纳入到一个系统模型当中。



## 系统的错误

在多数情况下，这些系统模型很快就会变得非常复杂，人们根本无法从中获取任何有价值的信息，只能再次落入相互作用的迷宫。

美国普林斯顿大学（Princeton University）的遗传学家Leonid Kruglyak认为，事后看来，希望通过从系统整体的角度来研究生物相互作用并以此来发现系统特征，似乎是一种不太现实的做法，因此，目前关于胞内和细胞之间的相互作用机制我们仍然一无所知。为此，Kruglyak将生物学同物理学进行了一番对比。如果我们在对量子力学、量子色动力学或者相对论都一无所知的情况下，能够建立起一套类似大型强子对撞机（Large Hadron Collider）之类的粒子加速器

(particle accelerator) 吗? Kruglyak表示: “你可以对所有的材料都检测一遍, 但是你根本不知道该如何分析结果, 因为你对其中的生理过程完全不了解。我们对这些生理过程——无论从生物学角度, 还是依据目前的理论基础进行预测, 都抱有过于乐观甚至天真的想法。事实上, 我们只能获得大量实验数据, 进行数据挖掘工作, 然后以非常普通的方式去研究这些数据的意义。”

但这并不意味着生物学家们在对曼德布洛特集没有丝毫了解的情况下就会不顾一切地深入研究它。有些生物学家们说, 更“聪明”一些的系统方法对他们的研究领域帮助很大, 他们能够借此发现一些生物学规律。“生物学已经进入了一个新时代, 科学在这里是建立在各种解释和假说(预测)的基础之上的, 而不再是根据各种现象进行因果推理。”美国加州理工学院(California Institute of Technology)的发育生物学家Eric Davidson这样说道。

这种进展并不是源自自上而下的分析, 即将各种数据全都一股脑导入模型, 希望从众多错综复杂的头绪中理清背后的机制。相反, 实际情况是科学家们在对实验室中比较容易进行控制的, 比如模式生物的进程的各个参与元件进行系统分析之后才能得出结论。他们使用的仍然是系统的方法, 但是采用的则是传统的自下而上的研究策略。

Davidson发现了一个在机体发育过程中基因调控机制是如何发挥作用的实例。Davidson研究小组花费了差不多十年的时间来研究海胆(sea-urchin)的发育问题。他们采用的方法是系统地敲除能够发育成骨架的细胞当中每一个转录因子编码基因。该小组通过观察细胞丢失这些蛋白之后对海胆发育造成的影响以及分析每一次敲除之后对其它转录因子造成的影响, 最终绘制了一幅能够阐明这些转录因子如何共同协作形成海胆骨架的机制图。这幅图就是根据Jacob-Monod的调控理论构建的, 只不过它非常详细, 将所有的调控因子与DNA之间的相互作用全都包括在内, 试图从中总结出一条能够适用于其它物种发育过程的普遍规律。比如海胆胚胎基因组中的转录因子编码基因被母体的蛋白首次激活后, 随即激活下游的转录因子, 下游因子形成了一个正反馈环路, 从而被永久活化。从果蝇到人类都是按照海胆的这种方式来进行调控的, 它们利用各种“模块化”的基因来调控机体发育进程, 各个模块之间都是相互独立的, 这样在进化过程中只需要改动一个模块就能起到作用, 因而不会破坏进程的完整性。换句话说, 其实各个物种的发育机制都是相通的。

“这种基因组调控系统控制了整个机体发育进程, 基因组调控机制的变动会影响机体进化的想法是我们以前从来没想到过的一种生物学理论。”Davidson介绍说。他说这是自1963年他们成立第一个实验室以来的最大进展, 其余大部分的进化理论其实都是没有价值的。

Davidson将他们的工作称作“通用法则”, 如果你掌握了它, 就能了解所有你想了解的东西。Davidson认为, 人类基因组计划能够帮助生物学家们更好地了解生物系统, 而不是局限于某一个基因、蛋白或者其它什么调控元件等非常细微的东西。

### Davidson认为人类基因组计划的好处有两点

首先, 基因组计划能够对模式生物, 例如海胆的基因组进行测序, 发现模式生物发育过程中所有活化的转录因子;

其次, 基因组计划还催生出了一大批新兴的生物学家, 比如计算生物学家等。

## 见仁见智

那么为什么在Davidson看来如此简单、有序的事情在其他生物学家眼里却是那么漫无秩序呢？这个问题就见仁见智了。比如对模式系统进行研究的科学家能够用研究人类生物学的科学家们没有用过的方法来操纵这些系统，从而获得更加确定的答案。而且每位科学家思考生物学问题时所采用的基本哲学方式都各不相同。“是人自己把问题复杂化了。我见过很多生物学家，他们中有些人是简化问题的高手，而另一些人则善于分解问题。前者能够从宏观角度把握问题，而后者更专注细节问题。”美国加利福尼亚大学伯克利分校（University of California, Berkeley）的细胞分子生物学家Randy Schekman这样说道。

Mark Johnston是美国丹佛科罗拉多大学医学院（University of Colorado School of Medicine）的酵母遗传学家，他被公认为一个归纳高手。他曾经就酿酒酵母（*Saccharomyces cerevisiae*）做过一个非常“乐观”的预测。Johnston认为到2007年，每一个酿酒酵母的基因及其相互作用都被研究清楚之后，所有有关酿酒酵母的问题就都迎刃而解了。他甚至还非常严肃地指出这必将在几十年之内实现。Johnston和Davidson一样，他们均认为有关酵母生活史的很多细节，比如DNA合成和修复的基本机制都已经被彻底研究清楚了。Johnston介绍说科学家们已经清楚地知道酵母5,800多个基因当中2/3基因的作用，剩下的也都很快会被研究清楚。Johnston主要对葡萄糖感知通路（glucose-sensing pathway）开展研究，他说如果能对这条通路里的相互作用进行定量研究就会心满意足了，这虽然是一件很难的工作，但并非不可能。

并不是每个人都认同Johnston的观点。美国布兰德斯大学（Brandeis University）的分子生物学家James Haber认为难以确定是否能在二、三十年内解决生物基本机制的问题。他说：“科研进展是否能够促使问题最终得到解决仅仅是一个说法问题。不过我们已经对有一些问题，比如染色体在核内是如何被安排的问题开展研究了。”Johnston认为将他所期望的对葡萄糖感知通路进行的定量研究推广到酵母的其它研究领域是不可能也是没有必要的。“你必须决定你了解到何种程度，有人会说他们永远都不会满足，因此我们必须不断前进。”Johnston说道。人们对这两种研究策略（简化问题和分解问题）之间的差异感兴趣并不只是出于旁观者的好奇心这么简单，这种差异每天都在日常研究工作中发挥作用，而且各大学术期刊的编辑们也都在以此判断究竟哪一种方法更值得支持，哪一种文章更值得发表。这也与人们究竟能不能彻底了解“生命的本质”这一生物学的终极命题密切相关。



## 复杂的程度

有研究人员，比如日本东京系统生物研究所（Systems Biology Institute）的系统生物学家Hiroaki Kitano就认为，生物系统看起来好像变得越来越复杂恰恰就是因为我们对它的了解越来越深入。Hiroaki说：“生物学就是一个系统，迟早我们会彻底把它给弄清楚。”

不过另外一些人则没有这么乐观，他们认为生物学家们永远都不可能彻底弄清楚生命到底是怎么一回事，不过这也并非什么大不了的事。美国约翰霍普金斯大学（Johns Hopkins University）的肿瘤基因组研究学者Bert Vogelstein认为基因组时代的来临能让我们第一次有望解决生命系统的复杂性问题。他认为对健康人群和患病人群的基因组进行测序之后获得的这些序列能帮助我们发现导致疾病发生的遗传病因，从而开发出新的治疗方案。肿瘤与其它疾病一样，要比人们预计的复杂得多。比如现在通过对肿瘤个体患者细胞的基因组进行测序就能发现每名患者体内大约有50多个遗传突变位点，不过每名患者突变的情况都各不相同。因此，我们以往那种希望找到某个单个基因靶标并将其作为治疗靶点的药物设计方案，正逐渐转变为针对大多数肿瘤细胞中都普遍存在的信号网络来设计药物的策略。

Vogelstein认为，即使我们永远都无法彻底了解生物学，也不用担心，因为只要我们掌握的知识足以帮助我们攻克疾病就够了，而且这并非不可能完成的任务。他还补充道：“人类的确非常善于充分利用仅有的知识，最大化地发挥其功用。有一点很重要，不要等到把一切都弄清楚之后再开展下一步的工作，因为要弄清楚全部情况并非一朝一夕之功。”实际上，针对错综复杂的信号通路的药物可能是目前最有希望获得成功的抗癌药。用于治疗病毒感染性疾病、肿瘤疾病和黄斑变性（macular degeneration，该疾病是西方国家主要的致盲性疾病）等疾病的针对小RNA的药物也已经进入了临床试验阶段。所以，越来越高的复杂性并不会阻碍我们的进步，而且这对于希望迎接挑战而不是害怕面对挑战的研究人员来说其实是一件值得庆祝的事情。

美国劳伦斯·伯克利国家实验室（Lawrence Berkeley National Laboratory）从事肿瘤研究的Mina Bissell说，她在人类基因组计划开展期间几乎被那种乐观的预计（即人类基因组计划结束之后，所有问题都会迎刃而解）给弄得快绝望了。她说：“到时候，著名人物会站起来宣布说：‘我们已经解决了所有问题。’生命的确很复杂，但这也正是它的迷人所在。”幸好Mina不用再担心了，至少目前看来，像曼德布洛特集一样复杂的生命系统还没有任何迹象能够表明可以被破解。

原文检索：

Erika Check Hayden. (2010) Life is complicated. *Nature*, 464:664-667.

筱玥/编译