

分子生物学之成功与局限*

Robin Holliday

前 言

科学的兴旺发达依靠理论和实验不断的相互作用。这种相互作用的本质是一场恒久的辩论，但它显然比仅仅提出并验证假说要复杂得多。理论和实验能相互作用就必须有一个统一化和概念化的理论框架，在此框架之内实验科学家能进行工作。实验的结果将支持、修改或者扩展已有的理论框架；或者在少数例子中能导致全盘修正过去的框架。这种最基本的理论和实验的相互作用在物理学和化学中可以很清楚地看到；在生物学科内，显而易见的也有遗传学和早年处于遗传密码被揭开时期的分子生物学。分子生物学的剖析能力正被运用于遗传学本身和其他众多的生物学领域，尤其是发育生物学和肿瘤学，还有神经生物学、寄生虫学和分类学等等。

在这篇评述中，我将集中论述分子生物学方法在发育生物学中的应用。控制基因代代相传的规律已完全清楚了，但是控制基因在发育过程中表达的规律在目前还是一片模糊。不过，在许多实验室都毫无疑问地假定，随着对不同胚胎发生时期或者各种特定细胞境况下的基因及其产物的更多信息的积累，自身就会导致新的原理的诞生。可是科学中已有许多先例表明，经验主义的归纳方法即仅仅大量积累资料，并不是获取成功的妙方。发育生物学的进展将依赖于概念化或理论性框架的提出，这个框架不仅能解释已有的资料，而且能对基因展现发育程序的战略提出有力的、可检验的预测。换言之，理论和实验的相互作用曾给分子生物学带来了早期成功，当分子生物学要应用到其他领域时，这一研究战略必须得到保持。

遗传学的起源和基础

在上世纪，许多生物学界的杰出人物没能正确评价遗传特性的遗传和生物个体发育的差异。著名的例外是孟德尔，他开始鉴定能遗传又能控制个体表型特性的因素。在他著名论文的开头，他描述了对于能用于这种研究的实验系统的要求，并清楚地解释了为什么豌豆能满足这些条件。如同众所周知的，孟德尔在

遗传研究中的革命性进展被他 的同人几乎完全忽略了 35 年。过后曾有许多解释说明为什么孟德尔的论文被人忽视，有些实在无关重要。在新近的一个严谨的学术分析中，Sandler & Sandler (1985) 认为以前的解释都不通。他们自己提出了令人信服的论据，认为最主要的原因是那个时代包括达尔文在内的许多生物学家都有一种信念，认为对遗传的了解和对发育的了解是不可分割的，不可能只研究其中一个而不研究另外一个。而孟德尔的天才正在于认识到遗传和发育是能够加以区别的，并且把实验和理论结合起来，揭示了遗传的基本定律。由于他的同时代人认为这是不可能做到的，因此这也就阻碍了他们理解孟德尔思想的重要意义。

当孟德尔论文的重要意义被 De Vries, Correns & von Tschermak 分别地再发现后，孟德尔的实验结果很快在其他许多实验系统得到证实。遗传学在这之后的进展很大程度上应归功于原来由孟德尔提供的理论框架，它使得检测动植物上许多不同的特征的可遗传性成为可能。在这方面很突出的是摩尔根和他的同事们用果蝇进行的工作。他们建立了遗传的染色体理论的各个主要方面，这个理论随后在许多动植物和稍晚在微生物中得到证实和完善。在所有这些研究中，都有理论和实验之间的持续的相互作用。染色体理论进一步发展并继续提出可检验的预测。实验的验证进一步加强了理论，并导致修正、补充或者完善，这里就不需要一一叙述了。

自然选择的进化论是所有生物学中的一个主要的统一理论。在本世纪 20 年代末 Fisher, Haldane & Wright 指出达尔文的进化论和孟德尔的遗传学是完全一致的，此时遗传学得到了有力的加强。不同领域间理论上的联系是科学发展的又一重要组成部分，因为每一学科都能加强另一学科。

由于基因影响或控制成年个体的表型，遗传学家们清楚地意识到发育的正常过程必然依赖基因。他们也知道基因突变能彻底打乱正常发育而产生异常或古怪的表型。但是他们当中的绝大多数并没有关心发育

* 本文译自 J. Theor. Biol., 1988, 132:253-262

过程,也就是联结从生殖细胞基因的信息传递到最后发育完善的成体的特殊表型特征之间的漫长的一连串事件。大多数遗传学家甚至没有意识到需要对于基因表达有生物化学的了解。象 Garrod 的“代谢的先天错误”(“Inborn Errors of Metabolism”(1909))一书被遗传学家们忽视了近30年(Haldane是一个重要的例外),一直到微生物遗传学的兴起。

分子生物学的出现

一般认为,有两大研究领域一起奠定了分子生物学学科。其一是微生物遗传学,简单地说,它证明了单纯DNA就能够传递遗传特性,并且提供了一个基因控制一个酶或蛋白质的证据。其二是核酸和蛋白质的化学和三维分子结构的研究。Watson和Crick在1953年阐明的DNA双螺旋模型象是催化剂,一举使人们有可能在分子水平了解遗传信息如何被传递、可遗传的突变如何产生以及原则上四个碱基为何可能是无穷尽地编码特异性的源泉。

分子生物学的早期是充满着严谨、智慧的学科,其中理论性概念对于它的发展是至关重要的。DNA碱基顺序和氨基酸顺序的线性对应的基本原理在得到坚实的实验证据的前几年就已建立了。关于遗传密码的理论直接引导了T₄噬菌体系列的精美实验,从而肯定了确定一个氨基酸需要三对DNA碱基。在此之前几年,20个氨基酸的重要性亦已认识到了。尽管实际阐明每一个氨基酸由三联体碱基编码是需要精妙的实验,然而理论上的预测实际上是这项工作不可或缺的先决条件。蛋白质合成机制的揭开,是理论和实验协同相互作用的又一个极好的例子。tRNA和mRNA的存在都是在实验证明之前已经预测到了。基因调节的一般机理,起动机和蛋白质阻遏分子的概念大多来自遗传学和生理学研究的推理,而不是其后严谨的实验验证的直接结果。

我在此仅提到了显示分子生物学早期特征的许多主要发现中的一部分。它们说明了根本的方法学:不同学科之间的相互作用,尤其是遗传学、X光晶体衍射和生物化学,为以后的发展提供了一个统一的理论框架。当然有一些后来被证明是不正确的理论上的建议。实验工作是非常重要的,因为在一个理论框架内,只有它能最后分辨孰真孰伪。

但令人惊讶的是在分子生物学这段辉煌的阶段之后紧接着的是一段间隔期,在此期间人们普遍地认为所有重大问题几乎都已被解决,而且某些工作领域例

如微生物遗传学,似乎已成为多余的(参看stent, 1968)。其后的结果是,许多知名的分子生物学家进入了所谓的“动物园时代”(“zoo period”)。他们放弃了对噬菌体或细菌的研究,而开始从复杂的真核细胞系统为实验对象(如果蝇、鱼类、线虫、轮虫和水蛭等),而且,许多研究的特定目标都是探求这些生物个体中的中枢神经系统是如何工作的。这类工作中的某一些取得了很大成功,并延续至今。然而,紧接着令人意想不到的,“动物园时代”很快被分子生物学的全新的发展所取代。

现代分子生物学

宿主诱致的DNA修饰和限制已为少数微生物遗传学家研究了许多年,但是一般仍认为是一个相当模糊的领域。然而预测的威力又提供了突破的机会。遗传学研究已清楚地表明能够识别特定的短的碱基序列的酶一定是存在的,而且已为这些酶的发现铺平了道路。其结果是导致了新技术发展的蓬勃兴起:对不同来源的DNA进行剪切和重新连接、克隆基因;发展了转移DNA进入大肠杆菌*E. coli*和各种真核细胞的手段;DNA序列分析;Southern、Northern和Western印迹法;定点诱导突变和原位杂交等等。这一系列的新技术导致了目前分子生物学向越来越多的生物学领域的扩展。

科学家们面对现代分子技术的众多及其威力颇感到困惑。Brenner,早期分子生物学的先锋人物之一,在一次电台采访中评论说,现在许多科学家已变成“呆子”,他们不再需要去思考重要的生物学问题,相反地却只是钻进实验室,然后用所能利用的技术收集大量资料。所显示的实验技术是令人羡慕的,所得的实验结果也是“辉煌”得可予发表的。但是研究的战略在根本上是有缺陷的。本质上,这是回到了培根的科学哲学。它坚持对归纳方法的信念,这就是收集事实和信息本身必然导致一般原理的形成,由此导致不断前进。

内含子的问题

对真核生物基因结构的研究的阐明,正体现了当前分子生物学所存问题的本质。基因的克隆和从mRNA反转录得到cDNA导致了在这些基因中外显子(编码序列)和内含子(非编码序列)的发现。这个发现是完全出乎意料并和预测的相悖。DNA碱基序列和氨基酸碱基序列的线性对应已在原核生物中充分地确立了,理论上没有充分的理由说明为何在更复杂的生

物体系中情况应该有所不同。内含子的发现是现代技术值得炫耀的一个例子，但是至今基因的外显子—内含子结构并没有提供人们深入了解染色体的结构或者基因表达的控制。而且，内含子还得为它的存在寻找解释。也许它们只是些偶然地将自己插入到染色体基因中的自私 DNA (selfish DNA)。也许它们在基因进化上是重要的，可以使不同的蛋白质区 (protein domains) 组合进一个多肽。也许它们为未来蛋白质的进化提供机制 (顺便说一下，某一性状的达尔文式选择只能在许多代之后才能表现出来)。再或者，也许内含子具有目前还全然不晓的功能。内含子提出了什么挑战吗？它们提供了彻底了解基因间相互作用的钥匙吗？或者它们根本不相干，只是堆积于高等生物基因组中的无用 DNA (junk DNA) 的一部分？

尽管这些可能性都被提出来了，但是认真讨论内含子的重要性的程度比起内含子在许多基因中存在的广泛而详尽的记载来只是次要的。因此，强有力技术的使用在这里只导致了对基因结构的描述，而对基因表达、分化或者发育等基本生物学问题至今几乎没有形成什么见解。

发育的分子生物学

在高等有机体中基因有两项主要功能：首先，担负代与代之间的信息传递；其次，展现个体发育的遗传程序 (genetic programme)。如同我已阐述过的，遗传学家们在揭示基因的第一项功能方面获得了很大成功：支配基因和染色体传递的原理已搞得很清楚了；对于第二项功能所知有限，更谈不上规律。发育生物学的进展在很大程度上与基因的研究无关。这与其说是发育生物学苦于缺乏理论，倒不如说是在发育过程中基因作用的理论的匮乏。举例来说，一些理论性概念，如位置信息 (positional information) 它可能支配在三维空间中细胞如何构成组织 (Wolpert, 1969)，这可以用实验检验并且已经得到正的结果。但是位置信息和贮存于基因中的发育遗传程序还很难联系起来。尽管有关发育的某些重要理论已在基因水平上阐述发育控制的程序，但它们至今还没有能直接与目前的分子生物学研究挂上钩 (Kauffman et al., 1978; Meinhardt, 1986)。

现在许多发育生物学家的主要目标是用分子生物学的新技术将基因表达的调控和产物同整个生物个体的发育联结起来。不幸的是，这种类型的研究在很大程度上也是经验主义的。其战略包括下列布署：探测

在特定发育阶段表达的基因；查看转录产物在胚胎或组织中的分布；分离和序列测定这些基因；将基因产物的氨基酸顺序同其他已知的蛋白质比较；找到基因的起动力区域并且寻找可能和这区域相互作用的反式作用蛋白质因子；将克隆到的有兴趣的基因注入转基因动物 (transgenic animal)，然后检测它们的表达以及对表型的影响；鉴定能影响细胞行为的扩散性形态建成因子 (morphogens) 或生长因子及它们的受体。人们可以看到在将来这种或那种分子的信息堆积如山，但是有多少问题会得到解释呢？

从分子水平考虑发育，果蝇有其明显的优势，因为不仅有巧妙的传统的遗传工程技术，还有现代化的新技术能使用。在果蝇已得到了大量的影响发育的有趣的突变体，其中同源异形突变体和母性影响突变体尤其有意义。果蝇早期发育的研究进展神速，但是即使在此领域，研究的方式大部分也是依靠信息的收集。当这样一些结果发表在最有声望的杂志上的时候，在这些文章的讨论部分很少能见到任何可以帮助阐明发育过程的令人困惑的复杂性的一般性原理。另一种对于将来进行发育的分子生物学研究有巨大潜力的有机体是线虫 *C. elegans*。从卵到成体的完整的细胞谱系已经搞清楚，许多影响发育或特异细胞类型的突变也已找到。然而，分子水平的研究并未能揭开在发育过程中导致特定的细胞谱系行为的机制。

这一问题的根本可以用果蝇的一个例子来具体说明。果蝇的幼虫含有一系列的成虫盘 (imaginal discs)，它们在化蛹时将产生成体果蝇。每个成虫盘都具有特定的程序，或者都已被决定，将产生出成虫的某一特定部分。这种决定的特异性是可遗传的，因为成虫盘组织能在一个适宜的微环境中 (如成蝇的腹腔) 传代扩增，然后当用蜕皮激素刺激时，仍能发育为特定的成虫结构。这种体细胞的可遗传性机制尚不清楚，但是，可以假定它是依赖于 DNA 和蛋白质的相互作用。因为 DNA 在所有的成虫盘中必然是相同的，成虫盘之间的差异必然在于特异的蛋白质，它们以某种方式控制可遗传性。如果是这样，那就提出了一个全新的遗传学概念，也就是说，存在某种蛋白质或染色质的遗传。更进一步，转决定 (transdetermination) 的过程已被充分证实了，在此过程中，特定的成虫盘组织有时能自发地转变成另一种特异性的组织而且后者也是可遗传的。那么，是不是一种蛋白质和 DNA 相互作用的模式 (pattern) 能够“突变”成某种新的可遗传的模式呢？

遗传学和分子生物学所有以前的概念都是以 DNA 结构为基础来说明信息的转移途径和遗传机制。是否需要一个完全不同的机制来解释体细胞的遗传呢? 特异蛋白质遗传的倡议者们应该站出来并且把他们的建议的变革性本质讲清楚。或许这会导致一场对一些基本问题更有收获的争论。

生物医学研究中的一些问题

分子生物学的各种技术目前正被用于研究人类重要的遗传性疾病。有关的基因可被定位, 然后克隆和序列分析; 接着推导出基因产物的氨基酸顺序, 以及这种产物是否显示出同其他具有已知功能的蛋白质有同源性。一个负责重要遗传疾病基因的发现往往会伴随着大量的宣扬; 其正常基因的功能不久将搞清楚, 其缺陷基因或者它的效应将有可能纠正。但是, 事实是对基因和它的产物的定位和鉴定本身并不能解释为什么有缺陷的表型。实际上目前在重要的生物学领域中, 包括发育, 有许多新的蛋白质被发现例子, 它们真正的细胞功能或者异常功能还都全然不清楚。数量巨大的卷入这种类型研究的工作往往将陷入僵局, 这实在是很不幸的; 部分地是因为在复杂有机体中的许多基本生物学过程并没有被了解。

问题的另一个例子来自癌基因与肿瘤的研究。有关癌基因的遗传性问题常在分子水平上加以解释。原癌基因(proto-oncogene)被激活而变为癌基因可以是突变所致, 也可以是转位到一个更活跃的启动子后过量表达而造成。尽管已费了很大的努力进行研究, 但是原癌基因同正常生长的关系以及被激活的癌基因同恶性生长的关系都仍不清楚。遗传物质的传递、DNA 的合成和有丝分裂的基本原理都是早期研究的成果, 而对癌基因的研究并没有使我们增加多少对细胞周期的认识。或许更好的战略是多费点人力和物力, 从理论上也从实验上, 研究细胞分裂本身。

对癌基因和生长因子的研究也引出了一些基本的生物学问题。比如, 生长因子先同细胞膜相互作用, 通过一连串细胞内事件(这也正是许多研究的主题)能触发 DNA 合成。细胞的表型明显地因一种外界物质——生长因子, 而发生改变。但这改变的表型在何种意义上是可遗传的? 细胞之外的信息能够从什么途径影响遗传系统? 很明显许多科学家都相信, 是细胞之间的通过可扩散性物质的通讯为稳定的细胞分化, 或者为转化细胞的正常表型的丢失提供了基础。如果外界信号能改变基因表达的模式, 并且转变细胞的表型,

那么这显然暗示出对分子生物学的“中心法则”的否认(“中心法则”认为信息从核酸流向蛋白质, 而不是相反)。这不是在细胞水平上回归到了拉马克主义上吗? 这些基本的问题很少被讨论过。

外遗传学(Epigenetics)理论的亟待建立

Waddington 既是一位遗传学家, 又是一位胚胎学家, 他认识到在生物个体发育过程中需要严格的规律来掌管基因活性的模式。他创造了外遗传学(Epigenetics)这个术语(Waddington, 1953), 它是研究在发育过程中怎样通过基因开关的机制, 或者产生基因活性的瞬息变化, 或者产生永久的基因活性模式。由于外遗传学的规律或法则都尚未了解, 因此这个术语常常带点贬意地被使用: 如果一个基因的激活或失活不能用传统的机制比如突变或阻遏物的控制来解释, 那么就祈求于外遗传学性的, 也就是未知的机制。事实上, 有些外遗传学基本原理应当说是清楚的。外遗传学涉及基因活性的时间控制, 基因活性在体细胞分裂过程中的分离, 以及活跃与不活跃状态的可遗传性。它还涉及到个体生命周期的特定时间这些活跃与不活跃状态的逆转。用真兽亚纲的雌性哺乳动物的 X 染色体可以说明这些基本原理。在雌性胚胎的细胞中, 两条 X 染色体中的任意一条随机地失活, 而且这种失活状态在以后的细胞分裂中非常稳定地遗传下去。这两条 X 染色体基本上具有相同的 DNA 而且它们处在共同的核质和细胞质之中, 因此, 其中一条 X 染色体的失活只是一种稳定的外遗传性的转变, 在基因或染色体水平上是可遗传的。但这种转变在减数分裂之时或前不久却因 X 染色体的重新激活而逆转。这种染色体活性的可遗传性和可逆性也能通过染色体印记(chromosome imprinting)得到证明。在这种情况下, 两条同源亲本染色体或染色体的部分在活性上显然不同; 也的确, 在哺乳动物能够进行正常发育之前它们需要彼此互补。在自交系的小鼠中来自父方和母方的常染色体在遗传上必应是相同的。按孟德尔基因的正常遗传, 这种活性印记将被重叠, 因而在通过生殖细胞传递时将被逆转(至少两条同源染色体中的一条是这样)。

至少已提出过三个不同的机制来解释发育过程中基因活性的外遗传控制。可移动因子(transposable elements)的发现和免疫细胞内免疫球蛋白基因的结构重组现象, 使得相信 DNA 顺序的重新排列对于高等有机体发育有普遍重要性的观点流行起来。从上面曾提到的例子看, 这样的一些变化可能是可逆转的。体细

细胞核是全能的,这已为众所周知。一些特定类型的碱基顺序变化也是可逆的,比如在单股DNA多核苷酸的茎和环结构之内的互换,但是这些目前还没有在高等有机体中证实。

第二种分子机制是前面提到过的染色质或蛋白质遗传。这不只限于DNA和蛋白质的相互作用,因为这种机制提出,某些有特异顺序的蛋白质与DNA协同地结合,而结合后的复合物能经过细胞分裂稳定地传递下去(Weintraub, 1979, 1985)。蛋白质遗传在分子生物学中引出一个全新的原理,但是它确实提出了可检验的预测。目前需要的是一个模式型实验系统,这系统既可以证明在一定的细胞类型中与一定的基因活动(或失活)相关联的DNA-蛋白质复合体的可遗传性和异常的稳定性,也可证明在具有不同的基因活动的细胞中这种复合体的缺乏或者变异。

第三种可能机制是从细菌中所知的DNA修饰酶的性质而引伸出的。这种机制提出在特定的顺序中胞嘧啶要么被甲基化,要么保持未甲基化,而且甲基化的模式可遗传下去。在过去几年中积累起来的许多证据表明,在DNA甲基化和基因表达之间存在某种关系,包括X染色体的失活和染色体印记(参考Razin et al. 1984, Holliday, 1987, Monk, 1986, 1987等人的综述)。但是目前还不知道, DNA甲基化对于主要的发育途径的控制是否重要。

以本文之篇幅来详尽地讨论赞成或反对以上这些分子水平上的外遗传性机制的论据,显然是不可能的。这些机制的意义和重要性在于它们作出预测,让人们能够用现有的分子技术去支持或反驳这些预测。然而,关于基因活性和发育仍然非常需要一个更全面的综合的分子理论。

结 语

科学研究的进展不是来自单纯的资料和信息收集,而是依赖于理论框架的存在,实验科学家能够围绕它进行工作。在生物学科中,遗传学尤其依赖于推理过程和实验验证。分子生物学的创建同样依靠严谨的学术气氛,在这气氛中许多重大的预测被实验所证实。理论和实验的相互作用曾持续了许多年,但是现在已为其他不同的方法途径所取代。分子生物学新

技术的威力使得资料积累凌驾于其他所有活动之上。这也就是说,很大程度上现代分子生物学倒退到了经验主义,希望会从分子水平资料的堆积中产生出一般原理。然而,目前大量的实验资料在诸如发育生物学和肿瘤学领域中仍是无法解释的。这些资料最终将派上用处,但是只是在有了一个合适的理论分析的时候,它能根据不同实验系统的研究整合成完整的统一的整体。

对于支配基因代代相传的法则已很清楚了,但对外遗传学,也就是基因在展现发育的程序中所采取的战略,几乎一无所知。分子生物学为这领域的进展提供了手段,但是至今仍未取得重大突破,这很大程度上是由于错误的信念,认为采用这些方法步骤的本身就能提供对基本生物学问题的答案!

(范国平译 曾弥白校)

参 考 文 献

- [1] Brenner, S., 1988, In: A Passion for Science. (Wolpert, L. & Richards, A., eds.) Oxford University Press, Oxford.
- [2] Garrod, A. E., 1909, Inborn Errors of Metabolism. Oxford University Press, Oxford.
- [3] Holliday, R., 1987, *Science* 238, 163.
- [4] Kauffman, S. A. et al., 1978, *Science* 199, 259.
- [5] Meinhardt, H., 1986, *J. Cell Sci., Suppl.* 4, 357.
- [6] Monk, M., 1986, *Bioessays* 4, 204.
- [7] Monk, M., 1987, *Nature* 328, 203.
- [8] Razin, A. et al., 1984, DNA Methylation: Biochemistry and Biological Significance. Springer-Verlag, New York.
- [9] Sandler, I & Sandler, L., 1985, *Hist. Phil. Life Sci.* 7, 3.
- [10] Stent, G. S., 1968, *Science* 160, 390.
- [11] Waddington, C.H., 1953, Evolution. In: *Symp. Soc. Exp. Biol.* 7, pp. 186, Cambridge University Press, Cambridge.
- [12] Weintraub, H., 1979, *Nucl. Acids. Res.* 7, 781.
- [13] Weintraub, H., 1985, *Cell* 42, 705.
- [14] Wolpert, L., 1969, *J. theor. Biol.* 25, 1.