



漫谈生命与信息的关系

○ 文 / 李衍达

当人类基因组计划出来以后，就引起我对生命科学研究的关注，因为人类基因组计划提供了大量的数据——我对一切的数据都感兴趣，因为其间隐藏了很多我们需要分析的信息和信号——初步解读了人类的DNA序列，测出了人类的30亿个碱基对的次序。人类的DNA序列是一本天书，它隐藏了人类生命的奥秘，是解读生命本质的一本书。其实，生物学的研究在人类基因组计划出来以后，已经进入了一个新的水平，不仅进入到分子水平，也进入到研究生命本质的阶段。这个研究将会对人类社会产生很大的冲击，使21世纪人类社会的发展引发出无限的可能和想象。

我们知道DNA序列是由核苷酸所组成的，由四种核苷酸类型组成一个DNA序列，如果用信息学的概念来说，可以用四个字母来代表。它排列起来组成一个很长的序列，人类DNA序列大概有30亿个碱基对，所以非常长。这30亿个碱基对的基本功能大体分两部分：一部分是编码蛋白质——生物行使功能主要靠蛋白质。蛋白质又是由20个氨基酸所组成的序列，首先要编码不同的氨基酸。编码要由核苷酸来完成，所以这一部分编码就组成人体或者是生物体的各个部分的基本构件。这个

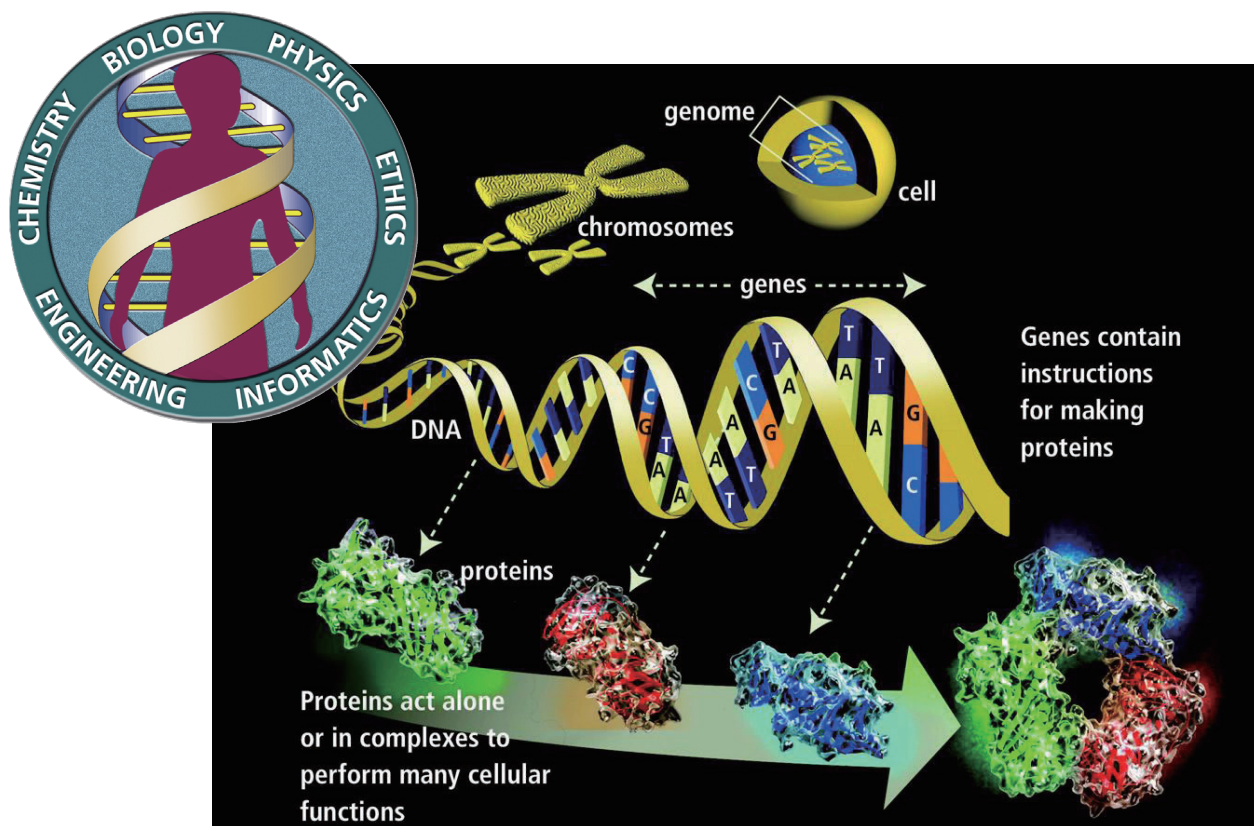
部分大概占DNA序列的2%到3%。

其余大部分序列除了一些重复序列和备用序列外，现在我们猜测可能是用于解决蛋白质和基因、蛋白质之间、或者是基因之间的相互调控。也就是说，这个序列从开始发育成一个不同的细胞、组织、器官到整体的发育过程，需要一个程序来调控：什么基因开始要打开？要表达？这些基因怎么相互作用？最后怎么在空间上和时间内配合起来组成一个整体？等等。所以这个调控程序是非常复杂的。打个比方，一台计算机，我要设计一个程序，从DNA序列开始，将它组合起来，组成生物的不同器官和组织，然后还要它合起来。这个程序非常严格，不能出错，错了鼻子可能长在脑袋顶上，空间就不对了。时间不对也不成。这是非常有意思的一个程序。我设想过这样的程序，但觉得非常困难，我甚至不知道应该怎么编，可各个生物体竟然都编了一个这样的程序。我猜想，生物不是像我们的计算机软件一样有一个集中控制的程序，它很可能采用分布自适应控制的方式，即每个细胞只与它周围的细胞和生物分子发生作用，而每个细胞的状态及其发展也只由它周围的细胞和生物分子的状态所决定。这样，不仅可以并行作用，而且具



李衍达

1936年生，广东南海人，祖籍佛山，信号处理与智能控制专家，1959年清华大学自动控制系毕业，中国科学院院士，清华大学信息科学技术学院院长，国务院学位委员会委员。主要从事信号处理理论和地震勘探数据处理方法的研究。曾获国家自然科学基金，国家教委科技进步奖等多项。



有鲁棒性（编者注：鲁棒性即英文robustness一词的音译，也可意译为稳健性），它不会因为某个环节出错而导致整个系统崩溃。

基因是由核苷酸组成的，而核苷酸本身也是一种物质，但是没有生命。由没有生命的物质，组合成了有生命的生物，可以说生命的出现是从无生命物质中产生的，可以简单的叫作“无中生有”。过去很多人包括很多科学家都不理解这种现象，所以对生命的认识就产生很多奇怪的想法。有人认为生命是有灵魂的，人死了，灵魂跑掉了，就从生又变成了死。有人为了寻找这个灵魂，曾经做过这样的实验：人死之前称一下体重，死了以后再称一称，看有没有什么变化，试图以此来反映这个灵魂的存在。当然这也是一个可以再探索的问题。

科学发展到今天，特别是近代研究的复杂系统，其实对这个问题提出了一个可能的答案。所谓复杂系统，它有一个非常有意思的特性，就是系统的整体特性可以是其各个分部所没有的。各个具有特性的分部组成一个整体系统，它可以出现每一个分部所没有的、新的特性。打个比方，现在国内很多地方的出租车司机最头疼的就是堵车，堵车其实是一种整体现象

的涌现，是多辆车实现相互联系的时候出现的，因为单个出租车是不会出现这种现象的。这个现象对所有的车都产生影响了，这就是所谓系统出现的一种新的特性。这个新的特性是系统之间相互作用而产生的，很多出租车司机每天晚上都会预测明天哪个地方堵车，结果大多数司机都发现，恰恰昨天不堵车的地方今天就堵起来了，因为大家都想着那个地方不堵往那里走。这也说明这种涌现是一个难以预测的现象，并符合复杂系统理论。所以复杂系统理论涌现特性的出现，概括来说，就是无中生有。了解了复杂系统的现象，再把它应用到生命现象，就很容易理解生命的出现。可以说生命的研究离不开复杂系统的研究，而复杂系统的研究是离不开信息科学的。

有人认为21世纪是生命科学的世纪，我认为是很合理的。因为生命奥秘的解开，对于我们的健康，对整个地球的生态环境的理解都极其重要。现在有人在研究人造生命，因为解开了生命的奥秘以后，就可以复制它，就可以制造人造的具有生命的生物。还有对人类寿命的研究，英国的科学家预测人的寿命大概是150多岁，一般认为正常的存活年龄大概在120岁左右。人体个体之间的寿命有差异，大概四分之一是由

基因决定的，四分之三是由生活方式决定的。因为人还要不断地跟环境进行交流，如果抽烟、喝酒，生活方式非常不好，那么寿命当然就会减下去。所以说尽管DNA对极限是有影响，但在极限以前的生死不仅仅是由DNA决定的，更重要的是由生活方式决定的。

那么，信息科学与生命科学，它们之间到底存在着什么样的关系呢？信息科学是研究信息获取、传输、处理和应用的科学。即信息是怎么得到的，如何获得信息，怎么传输信息，然后信息相互之间的作用又影响到系统的调控，也包括系统内部的调控关系。人们往往想到的是通信、手机、电视机跟信息科学有关，网络跟信息科学有关，调控跟信息科学有关，传感器跟信息科学有关，但生命又跟信息科学何关呢？

其实信息科学本身是一种横向科学，涵盖了所有的事物。因为信息的获取并不仅限于设备，也不仅限于网络，所有的事物都有信息。抽象地说，信息是客观事物的运动和特征的一种表述。有云有雨，这就是气象特征的一种表述，这就是它提供给我们的信息。气候有信息，海洋有信息，海浪的有无、海水的温度等也都是信息，所以万事万物都有信息。正因为有信息，我们才能对它进行研究、分析，所以信息科学并不像我们所想象的那么狭窄。

同样，对生物信息重要性的认识也是一个逐步加深的过程。以前生物学家把生物信息只看作是管理生物学数据，后来研究发现生物信息在某种意义上反映了生物深层次、本质的问题。举例子来说，遗传是生物的一种最重要的本质性的特性，如果没有遗传，就没有生命。那么遗传是遗传什么呢？其实遗传是信息的遗传。把所有生命的某种生物特征用信息编码下来，然后遗传下来。有人说，也可能遗传是一种物质的遗传，因为刚生下的婴儿是由父母的物质组成的，某种意义上有物质的遗传。其实婴儿七、八斤重的物质在以后的新陈

代谢当中早就代谢掉了，长大后身上保留的父母的物质基本上没有了。但是他会保存父母的特性，他的脾气、习惯、行为都可能跟父母很像，那是因为父母的信息遗传给了他。所以在本质上来说，遗传本身是一个信息的问题。

我们研究生命，核心的问题不在它的组成部分，而是生物内部的生物分子和蛋白质相互作用所形成的一个活的整体。所以生命的奥秘很大部分在于生物分子之间的相互作用、相互调控，组成一个活的整体。这就牵扯到一个调控的机理，靠什么来调控呢？靠信息。生物内部的信号，通过转导来调控生物体，调控哪个基因表达、哪个基因关闭，调控哪些蛋白质应该到哪儿去。所以调控关系是决定生物本性的一个很重要的部分。而研究调控问题，又是跟信息密切相关的。因为生物体内部的调控，现在很可能跟一般的设备调控机理是一样的。生物的本质问题跟调控有关，所以调控专家可以介入，而调控则是信息科学的一部分。更重要的是，生命体内部具有复杂系统的性质。如果没有复杂系统这个概念，生命就根本不可能出现。而复杂系统也是信息科学研究很重要的内容。所以从遗传一直到调控，一直到复杂系统的性质都是信息科学研究的重要内容。因此可以说，信息科学家的介入其实对解开生命这个奥秘，起着核心的、不可替代的作用。

一般人认为信息科学在生物生命科学里所起的作用是工具作用，表现在对生物信息数据的管理、对生物信息的分类、对序列用计算机进行比对、对蛋白质的折叠进行预测等等。但我认为生物体内部的信息系统、信息问题是生命产生的一个核心问题，遗传信息就是生命的一个重要特性，其实它就是生物特性的信息编码。而现在进一步发现，生命的产生跟调控分不开，跟生物分子之间的调控分不开，从而表明信息科学在生命科学里起着核心的作用，生物信息学将来很可能是生命科学的本质性的、主

流的、关键的学科之一。因为只有从信息的模型的角度，才能够真正从本质上来理解生命科学；也只有从复杂系统和信息模型的角度，才能够真正的理解生命。当然我们也必须承认，化学、物理学、数学的应用等在生命科学里也都起着很重要的作用。

现在世界上有很多发达国家都建立了生命交叉科学的研究中心，美国很多著名的大学包括斯坦福、MIT都建立了生命研究中心，物理、数学、计算机、信息等学科都介入了。为什么这样呢？首先，人类基因组计划提供了那么大的数据，生物学家觉得难以管理了，所以需要计算机学家帮他建立数据库。这么多的序列需要分析，所以又需要计算机专家帮他进行分析。为了分析这个数据，他又需要建模，所以数学家也需要了。当然，化学家、数学家都在建模里面发挥作用。进一步分析，又发现光建模还不够，还得研究模型内部这些生物分子之间的关系，涉及到调控的问题，所以现在我们发现很多调控专家也介入了。这其实是生命科学研究进入到一个更深层次的表现。它说明生物学已经脱离了描述性的，实验性的科学，而进入到一个定量分析、可以预测、甚至可以寻找生物内部规律的一个新的阶段。在这个新的阶段，光靠生物学家大概是不够的。所以物理学家、数学家、计算机专家、调控专家介入，我认为是必然的。这也说明原来科学本来是一个整体，只是人为地把这些东西分开了。实际上当要研究一个综合性的、实质性的问题时，所有的科学手段都要利用。只有这样，才能够真正对事物形成理解。所以，不仅是生物学，其实所有的科学都进入了一个交叉融合的阶段，交叉学科变成了一个前沿学科。从科学本意来说，这是必然的。我们不应该人为地把学科硬性隔离开来，而应该从问题研究的实质出发。生物科学或者说生命科学的交叉学科中心的成立就反映了这样一种趋势，不过它反映得更加突出、更加明显而已。❏