

基因沉思录

关于DNA中的美学、信息学、数学、化学、物理学及社会学

江泓 著

中国科学技术出版社



基因沉思录

关于DNA中的美学、信息学、数学、化学、物理学及社会学

江泓著

中国科学技术出版社

图书在版编目(CIP)数据

基因沉思录/江泓著. —北京: 中国科学技术出版社, 2004. 3

ISBN 7-5046-3745-9

I. 基... II. 江... III. 基因—基本知识 IV. Q343.1

中国版本图书馆CIP数据核字(2004)第016460号

中国科学技术出版社出版

北京市海淀区中关村南大街16号 邮政编码: 100081

电话: 010-62103210 传真: 010-62183872

科学普及出版社发行部发行

<http://www.kjphooks.com.cn>

北京长宁印刷有限公司印刷

开本: 787毫米×1092毫米 1/32 印张: 3.5 字数: 80千字

2004年4月第1版 2004年4月第1次印刷

印数: 1-1000册 定价: 12.80元

序

不论我们自觉自愿与否，我们已进入了后基因组时代。这是一个令人担忧的时代，因为对基因的合成和改造已经开始，我们不知道下一步会有什么从未见过的有害生物被合成出来，我们也不知道将有什么怪物会被制造出来。同时，这又是一个振奋人心的时代，因为我们对于生物和人类本身的认识正在深入，更高产、更美味的食品正在被产出；更多的疾病的诊断和治疗方法正在被推出；多个物种的基因组图谱正在被绘出，多年辛勤工作所获得的零散资料，片段的 DNA 正在汇集成完整的基因组，一幅完整的图画，而只有完整的图画方可给出深层的含义。我们有幸能成为看到这些完整蓝图的第一代，也肩负了破译这些蓝图含义的历史使命。目前我们已经得到了某些生物基因组的 DNA 图谱，但这并不等于完全理解了它们的含意，因为我们尚不知这些基因是如何相互作用的，不知道这些基因产物之间的相互关联，总之尚有大量工作等待我们去完成。这工作是艰辛的，也是有趣的。

科学的魅力在于它对未知的解答和它内在逻辑的美丽。因看到科学的美丽而产生对它的热爱是许多人踏上科学之路的原因。展现现代生物科学与其他基础学科的内在此联系也就部分地展现了这种美丽，所以用通俗的语言展现出分子遗传学与其他学科之间的内在联系，它的逻辑和它的魅力，是这

本书的目的之一。

然而当我真正提起笔来做这件事的时候，才发现它非常困难。首先是变枯燥为有趣又不失其真实和准确。虽然我的研究领域就是 DNA 的组成与结构，加之在北京大学医学部讲授了多年的“生物化学”课程，有些现成的资料，但若想做到让不同知识背景的人都感兴趣，都能理解，都有收获，还得下一番功夫。我尽量将它写得较为浅显通俗，然而在有些章节里仍不可避免地引入一些专业术语和概念，此类地方我均给以解释，细心的读者定可领会。读者也可以选择性地进行阅读。其次是变繁杂为简单又不失其系统。当前，科学的飞速发展迫使每一个研究者只能深入于一个较窄的领域，特别是对于分子生物学和分子遗传学这样每天都有大量新发现的日新月异的领域，能在某种疾病或某个物种的狭窄领域内跟上前沿的进展就已经很不容易，而我却不自量力地想在一个较为广泛的范围里进行介绍，这样就不得不把自己暴露在一个易犯错误的危险境地。然而反复思考之后，我仍不忍放弃这一尝试，决定写这样一个较为宏观的东西。因为科学就像艺术，艺术之美离不开完整，科学的观察就像对任何事物的观察一样，只有观察到它的全貌，才能发现它的美丽与深刻。而若把自己局限在一个微小的局部的精雕细琢，也许会误认为这仅是一个枯燥的、繁琐的技术性的事物，失去对它的感觉，艺术家也可能变成匠人。因此，我想尽量做整体的描述，然而在跨学科领域的描述中若有不妥之处，还望读者

给予指正。

科学与艺术的密切联系使本书能够在科学中引入人文主义的因素。这样，科学也就有了生命力。

本书在论述分子生物学与各基础学科的联系时，无意导入还原主义的立场；相反，本书强调的是整体观、动态观及各个系统之间的联系。正是这种整体性和相关，使我们对科学有了更深刻的了解，并能够领略自然之美与神奇。

人们把原子弹爆炸、阿波罗登月和人类基因组计划并列为 20 世纪三个最伟大的科学进展，而我认为人类基因组计划比起前两者来说，更具有特别重要的意义。当人类可以用威力无比的武器去占领的时候，当人类有了一切物质上的优越，甚至有了飞船，登上了月球的时候，我们的注意力开始从外向的开发和占有逐渐转向我们的自身，我们的细胞的组成和功能。因为只有提高人类自身的健康水平和生活质量才是人类最重要的事情。几乎所有的疾病都与基因有关，都可以从分子水平上进行阐述。因此，若要从根本上提高人类的健康水平必然要研究我们的基因，目前已经有基因诊断的方法和基因治疗的药品出现，基因的开发与利用正在医学领域里展开。几乎所有的生活用品都和基因相关，我们吃的、穿的、用的，离不开植物和动物，也就离不开基因，因此若想要提高人类的生活水平也需要改良植物和动物的基因。目前我们的食物中已经有些是经过基因工程改良了的植物或动物。所以有人预测说，21 世纪将是分子生物学和分子遗传学的世纪。

但是，未来对于基因的不加限制的改造，是否符合人类的道德标准和根本利益，仍是一个没有定论的问题。让民众了解这方面的问题和进展是我们科技工作者的责任。

科学衍生出可造福人类的技术，但技术不等于科学。科学是非功利的，是我们出于道德理念或出于满足好奇，不断追求真理的过程。“Science is the way of seeking truth”我们的人生目的是追求真理，科学为我们提供了方法与途径。我们追求科学的本质是在追求真理，因此利用科学去达到个人的什么利益，利用欺骗去制造伪科学，都是与科学基本精神相违背的。是科学的敌人。

通过沉思真理而得到滋养，这便是
众神的生活。

——柏拉图

目 录

引子	1
第一章 基因的发现——DNA 中的美学	3
第二章 基因的组成——DNA 中的信息学	18
第三章 基因的功能——DNA 中的数学	25
第四章 基因的产物——DNA 中的生物化学	38
第五章 基因的进化——DNA 中的生物物理学	50
第六章 基因的技术——DNA 的利用与论理学的讨论	74
第七章 基因的理论——DNA 研究的社会学提示	89
参考文献	99

引 子

什么是 DNA？DNA 是脱氧核糖核酸(Deoxyribo Nucleic Acid)的英文缩写，它是基因的化学分子名称，是基因的化学本质。

那么，什么是基因呢？不同学科领域的学者可以给出不同的回答：

遗传学家告诉你，基因是遗传信息的贮存形式。

化学家告诉你，基因是大分子的 DNA。

生物学家告诉你，基因是生命得以延续的物质基础。

基因在哲学上的讨论更是五花八门。有人说基因来自上帝，有人说基因来自外星，有人说基因是自私的，有人说基因是无私的。给我留下深刻印象的是一次与基督徒的讨论。我与 Tina 讨论有关人类的起源。为了证明人类是进化而来，我举出 DNA 的例子。

“知道吗？”我以生物化学家的口吻说：“基因的基本组成是 DNA 的三联码，这套基因密码从低等的细菌到人类都是通用的。这说明经过几十亿年的进化过程，密码仍未改变。由此看来，人是由低等生物进化而来。”

“我不这么想”，Tina 眨了眨眼睛，若有所思地说：“我认为这只能证明上帝用同一种原料创造了所有生物。”一时间，我似乎无言以对。

是的，基因就是这样神秘而深奥，对它的解释似乎可以有多种，又似乎只能是一种。

在它被发现后的一百多年里，我们提出了无数的问题，有些找到了答案，而更多的问题仍有待解答。

人类尚在探寻更多，更加完美的答案，然而一个多世纪以来的发现已使我们激动不已。本书力图展示的是那些已经发现的和正在发现的生命原理的答案，以及那些未知的，有待我们去探索的秘密。

今天，我们正处在一个值得庆幸的时代，这是因为生命科学的进展已经使我们可以从单纯的实验资料的收集阶段走向信息的综合分析和生物自然规律的探寻阶段。长期以来积累的大量资料和事实片段像是一抹抹绚丽的色彩，正在逐渐连接融合，使美丽的完整的画面逐渐显现，我们似乎第一次看到这完整的美和隐蔽在这画面里的更深刻的东西。这些东西使我们沉思，使我们检讨过去所得出的结论，也令我们能够将分子生物学——这一实验性学科，推向理论的高度。

愿渴望了解生命奥秘的朋友们与我共同遨游这令人兴奋的领域，共同欣赏这绚丽的画卷，共同领略这发现的乐趣。

第一章 基因的发现

——DNA 中的美学

我想我们追求科学的目的，归根结底是为了追求一种真实的完美，一种对自然的完美性和统一性的发现，一幅完美的，深邃而通透的自然规律蓝图的描绘。

关于生命蓝图，长期以来经许多有关遗传基因的科学家的描画而渐趋完美。描画的第一笔，应该算是拉马克(Lamarck)的获得性遗传思想。拉马克是19世纪的生物学家，他提出的物种可以受环境的影响而发生变化并可遗传和进化的学说，否定了上帝造万物，物种不变的创世论。他所提出的生物可以进化的思想，具有革命性的意义。这是人类历史上第一次对上帝创世论的挑战，第一次以变化的观点看待世界万物。他提出的物种可以随环境而产生变化的观点是很有意义的，这一观点在20世纪50年代被米丘林等人发展，为植物的育种选种提供了理论基础。然而遗憾的是，拉马克又错误地提出了直接适应，过分强调了获得性遗传的作用，未认识到获得性遗传的局限性。

达尔文提出的进化论的核心是遗传、变异、选择三要素。就是说生物的性状是可以遗传的，在遗传的过程中是可以发生变异的。这些变异的保留与否需要由自然的压力来选择，只有那些适合自然条件的变异性状可以保留下来，代代遗传。虽然当时由于条件所限，达尔文无法得知遗传的是什么，怎样遗传等问题，但他总结了遗传的一些现象，提出了生物

必须适应自然，从简单走向复杂的进化规律。

达尔文无疑是最伟大的生物学家，他的进化论可以说在世界上首先创立了生物学。然而由于历史条件所限，严格地说，达尔文时期的生物学尚不能算是完整意义上的科学 (popper)，这是因为：

其一，科学的基本模式是对假说的实验验证。达尔文关于生物进化过程的假说却验证很困难。

其二，科学的基本原则是科学原理的普遍适用法则，即所提出的自然科学的假说或理论(被验证了的假说)应该能够被基础的化学和物理学的基本原理所贯通。而达尔文的进化论却似乎与自然科学，例如化学和物理学脱节。

其三，衡量一门自然科学是否成熟的重要标志之一，是看其是否能够被量化，即以数学对其掺入的多少为标准，而当时的生物学乃停留在表面现象信息的收集和描述，尚无数学的分析和归纳。

几乎与达尔文同时代，有一位叫孟德尔 (Mendel) 的生物学家，他是奥地利的一位教士。在 19 世纪的欧洲，大多数人尚生活在贫困中，然而教士的生活却是有保障的。这种平静安稳的生活给了孟德尔一颗安祥平和的心，使他可以耐心细致地去观察自然界的种种微妙的变化。同时，奥地利优美的自然景色又给了他丰富的想像力和灵感，使他能够深入地思考，能从某些表面现象中总结出内在的规律。他在教堂的后院种了一些豌豆，每天都充满兴趣地去观察豌豆的开花情况，他发现开红花的豌豆与开白花的豌豆相杂交，子代总是开红花，也就是说红色是显性遗传，白色是隐性遗传。用这个子代自交之后，则产生了 $\frac{3}{4}$ 开红花的和 $\frac{1}{4}$ 开白花的子二代，而不会产生类似粉花的子代这种融合的表象，因此他

认为基因是独立遗传的。他指出，基因是一个个独立的单位，互不融合。他将这种可携带遗传性状的物质叫做遗传因子，即基因。设 C 与 c 分别为代表红色和白色花色的不同基因，对于豌豆这样一个二倍体的高等生物，则有：（图 1）

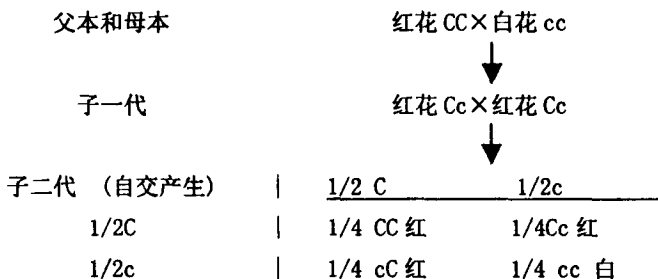


图 1 等位基因的遗传(分离定律)

孟得尔利用简单的概率分析，得出了等位基因的分离定律。什么是等位基因呢？高等生物大部分是二倍体生物，它们的细胞中含有的基因是来自父本和母本的两套基因的杂合，也就是说每种基因有两个拷贝可供选择。对于豌豆来说，表现花色的一对红白基因位于一对同类染色体（又叫姊妹染色体或同源染色体）的相同位置，表达同一类性状，例如花的颜色，故称为等位基因。豌豆通常只能表现等位基因的两个性状其中之一，所以称为分离定律。例如红花或白花，而不表达两者相间的性状，例如粉花。对于那些在等位基因中总是得到表达的基因叫显性基因，例如子一代表现的红花基因 C 及子二代 $3/4$ 表现的红花基因，而那些在有别的等位基因存在时总是保留不表达的基因称之为隐性基因，例如白花基

因 c。从上面的例子中可以看到，子二代仅有 $1/4$ 开白花，这是因为两个白花基因碰到一起 cc 的概率是 $1/4$ 。也就是说红花的数目总是 3 倍于白花。这是由于白花基因和红花基因这两个等位基因在豌豆中表达的情况下，总会表现显性基因红花的特征。并非所有生物的等位基因都是分离表达的。例如黑色的狗与白色狗相交，有可能产出黑白相间毛色的狗后代。这是基因的某种融合现象。基因的表达产物主要是蛋白质，是那些不同颜色，不同功能的蛋白质。人体的一切形态和功能都是通过蛋白质来体现的。例如，肌动蛋白使我们的肌肉可以收缩，酶蛋白使我们的食物可以被消化。人体不同部位的不同颜色通常也是由于蛋白质的颜色不同。例如血液的红色来自血红蛋白，毛发的黑色来自黑色素等等。

非等位基因是指不同源染色体上的基因，或者是同源染色体上不同部位上的基因，它们表达不同的信息，产生不同蛋白质，表现了不同的性状，例如豌豆的颜色和豌豆的饱满度的表达就是互为非等位基因的两个基因的表达性状，这两个基因在染色体上的位置是不同的。互为非等位基因的两个基因可以共同表达（总是共同表达的基因称为基因连锁，它们通常是位于染色体上相邻部位的基因）或进行自由组合式遗传表达，以豌豆为例，虽然一株豌豆中的每个细胞中均含有整套 DNA 的几百个基因，每个细胞却不必完全表达父本的全套或母本的全套染色体，而是选择性地表达父母双方一对等位基因其中的某一个基因。例如，它可以表达一对 1 号染色体中的父本和一对 2 号染色体中的母本，这样就有了不同的组合。例如豌豆颜色有黄色 A 和绿色 a 两种性状，豆粒有饱满 B 和皱折 b 两种性状。若黄色 A 饱满 B 的豌豆和绿色 a 皱折 b 的豌豆杂交，子代可出现黄色饱满 AB，黄色皱折 Ab，

绿色饱满 aB ，绿色皱折 ab 这样 4 种不同性状。黄色基因可以和饱满基因共同表达，产生黄而饱满的豌豆 AB ；也可以和皱折基因共同表达，产生黄而皱的豌豆 Ab 。绿色基因也可以有这两种组合，产生绿而饱满的 aB 或绿而皱的 ab 豌豆。

孟德尔发现的这两条定律（图 2），等位基因的分离定律和非等位基因的自由组合定律，奠定了遗传学的基础。

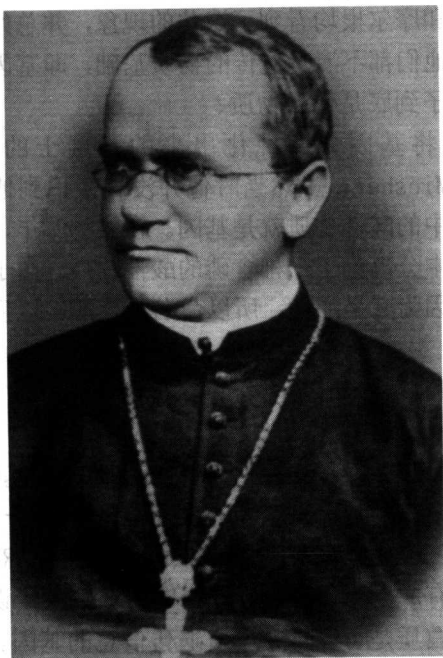


图 2 孟德尔

摩尔根 (Morgan) 的研究则第一次将某个基因定位在一个特定的染色体上。果蝇的眼睛通常为红色。有一次他发现了一只白眼果蝇，将这只白眼果蝇与红眼果蝇进行交配，子

代仅有少数白眼果蝇，他们均为雄性。这说明眼睛的颜色基因总与性别基因雄性联系在一起遗传，摩尔根认为果蝇的眼色基因与果蝇的性别基因不可分割地位于同一基因单位（当时并不知道就是染色体）上，共同遗传，共同表达。他将这种位于同一基因单位上的基因倾向于集体遗传的现象称为基因连锁。

孟德尔和摩尔根均看到了遗传的现象，并总结出了一些规律，然而他们都不清楚遗传的物质基础，即基因的物质基础，基因分子到底是什么物质？

第一次将基因分离纯化出来的是瑞士的一名学者 Friedrich Miesher，虽然当时他并没有认识到他分离出的这种细胞核中的酸性物质就是基因。在 1868 年，他从脓细胞的细胞核中分离到了一种含磷的酸性物质。由于存在于细胞核中，又因为其呈酸性，所以命名为核酸。以后证明，核酸有两种，根据核酸所含核糖的不同区别为核糖核酸 (Ribonucleic Acid, RNA) 和脱氧核糖核酸 (Deoxyribonucleic Acid, DNA)。

当 20 世纪初，人们试图从细胞核中纯化出基因分子时，发现细胞核中有两种可供选择的大分子化合物，一种是我们提到的核酸，另一种就是蛋白质。核酸只有两种，RNA 和 DNA，而蛋白质却有成千上万种。因此当时的多数人错误地认为，多样化的蛋白质是可以携带丰富遗传信息的基因。

那么，到底是核酸还是蛋白质起了传递遗传信息的作用，尚需要经过实验的证实。核酸包括 DNA 和 RNA，由于 DNA 不像 RNA 那样易遭自然界核糖核酸酶的破坏，一些研究者找到了分离提纯 DNA 的方法。认识到 DNA 的遗传功能的研究发生在 1928 年。当时 Griffith 的研究发现，从一种致病的肺

肺炎球菌细胞核中提出的物质，可以使另一种无致病性的肺炎球菌转变成具致病性的肺炎球菌。1944 年 Avery 证明了用这种方法抽提出的物质正是 DNA。由于是 DNA 导致了肺炎球菌遗传性状的变化，因此得出结论为 DNA 是遗传物质。

别的研究组也做了类似的研究。噬菌体是噬食细菌的病毒，它可以感染细菌并改变其遗传特性。为了证明到底什么是遗传物质，1952 年，Hershey 等设计了这样一个试验。他用放射性硫 ^{35}S 标记噬菌体的蛋白质部分而用放射性磷 ^{32}P 标记噬菌体的核酸部分。用这样的噬菌体感染细菌，结果发现放射性磷 ^{32}P 进入了细胞核内，并将细菌的遗传性状改变，使细菌带上了噬菌体所携带的某些遗传性状，而带放射性硫 ^{35}S 的蛋白质则被留在了细菌的外面，并未进入细胞。因而他发现核酸才是遗传分子。此实验进一步证实了核酸为遗传信息的携带者。

核酸为什么可以携带遗传信息？它是如何组成的？带着这些问题，Chargaff 用酚和氯仿除去细胞核中的蛋白质，用乙醇将核酸沉淀，他发现细胞核内的核酸主要为 DNA。而长链 DNA 可被水解为四种单核苷酸。单核苷酸是由磷酸，脱氧核糖和碱基三部分所组成。四种单核苷酸的磷酸部分和脱氧核糖部分均是相同的，只有碱基部分有所不同，碱基分为腺嘌呤(A)，鸟嘌呤(G)，胞嘧啶(C)，和胸腺嘧啶(T)四种：经过对 A, C, G, T 的定量分析，他发现了几个有趣的现象。

——同一物种的各种组织细胞中的 DNA 的碱基组成比例均相同。

——相近物种的 DNA 碱基组成比例相近。

——在四种碱基含量的比较中， $A = T$ ， $G = C$ 。在各种生物中均如此。