

蛋白质：是随机产生的还是被设计的？

“倒置”的视网膜是“糟糕的设计”吗？

DNA条码显示物种之间近期的瓶颈效应

创造研究

Creation
Research

2020年 第1卷

生命的起源

若没有智慧，单凭物理过程和化学反应，会自然产生生命？

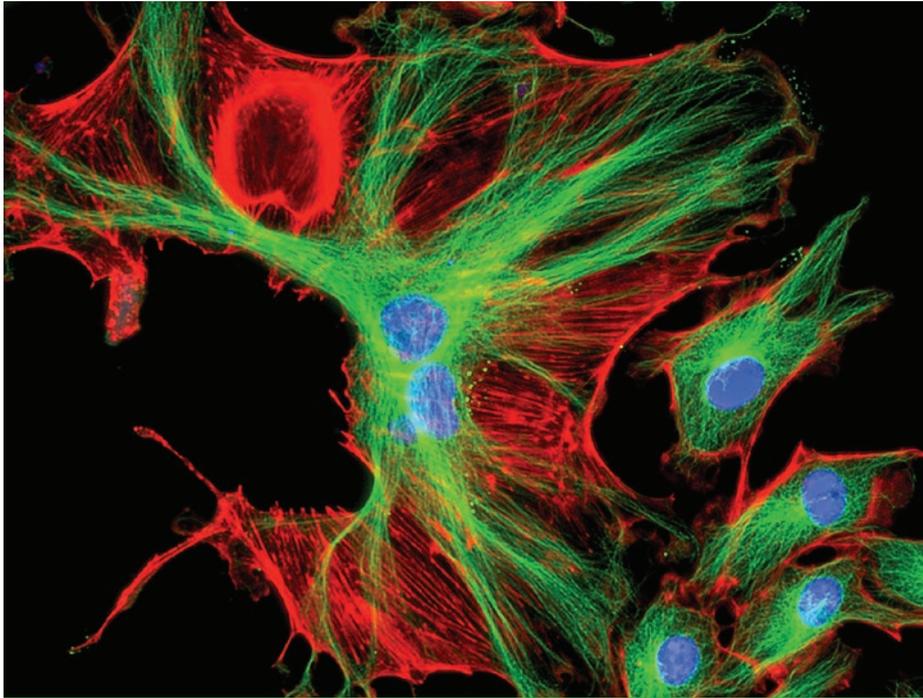
揭露 用基因组来骗人的把戏

使用被抛弃的数据对“人与黑猩猩的DNA几乎完全相同”的故事进行重新评估



目录

生命的起源：何以会自然产生生命？



03-13

随着我们研究发现的事实和规律越来越多，生命起源的“问题”变得更加困难；答案不是越来越接近，而是越来越遥远。

我们“倒置”的视网膜真的是“糟糕的设计”吗？ 22

DNA 条码显示物种之间的差距和近期瓶颈效应 29

蛋白质家族：是随机产生的还是被设计的？ 32

人类与黑猩猩基因组的相似度不超过 87%，甚至可能小于 81%。

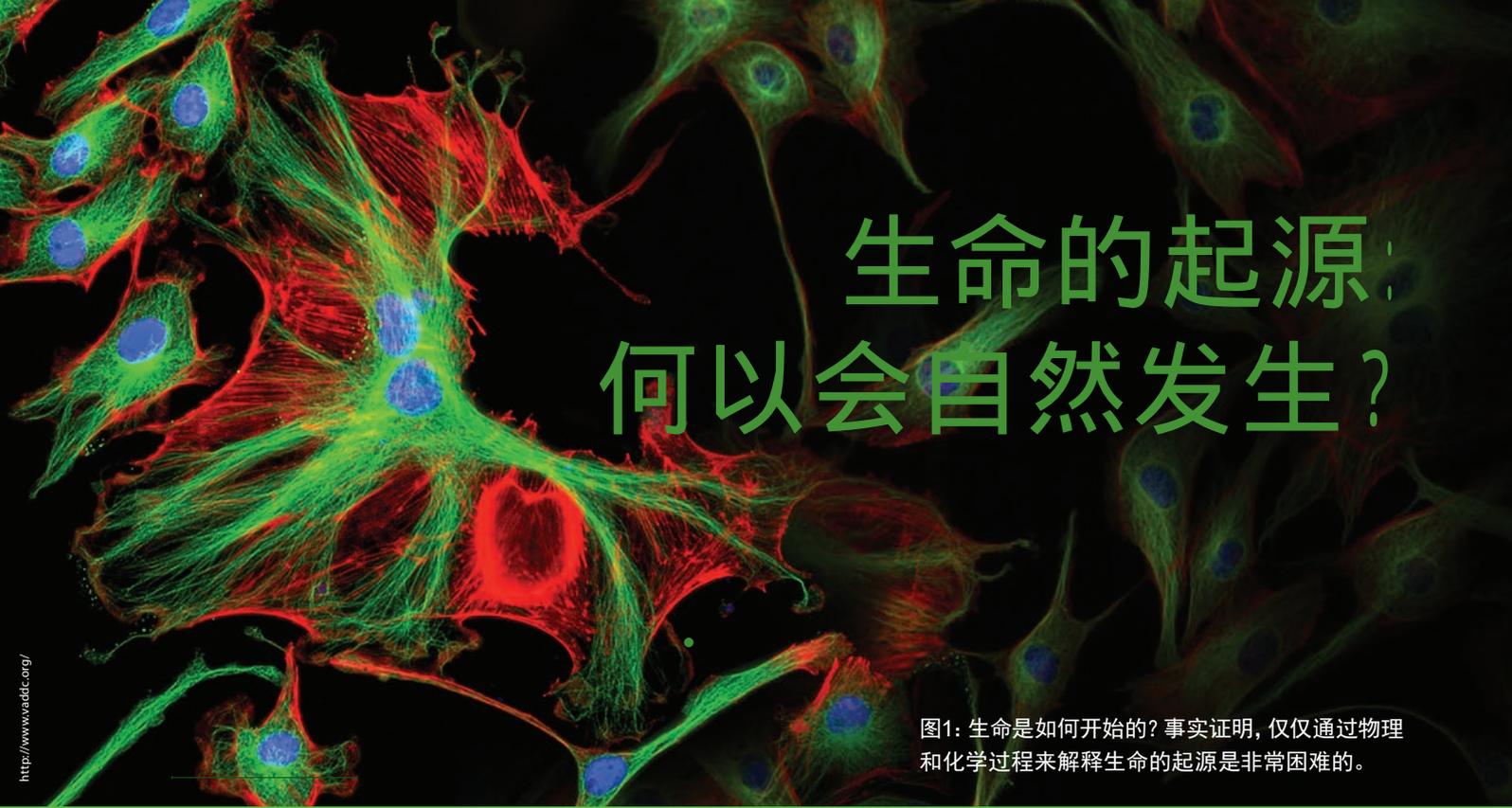
揭露用基因组来骗人的把戏



14-21

如何免费订阅

- ◎ 《创造研究》为电子版期刊，是创造论科学家近期发表的论文，为创造科学领域的科研人士提供展现和交流的平台。
- ◎ 请提前订阅，方便安排为您发送。
- ◎ 请通过封底页提供的方式联系我们。



生命的起源： 何以会自然发生？

图1: 生命是如何开始的? 事实证明, 仅仅通过物理和化学过程来解释生命的起源是非常困难的。

文/ 唐·巴顿 (Don Batten)

生命

的起源也称为生命的自然发生, 有时也称作化学进化。生命的基础是富含信息的长链分子, 例如 DNA 和 RNA。这些分子中带有制造蛋白质的指令, 而蛋白质则是生命的砖瓦。但要读取 DNA / RNA 中的信息来制造蛋白质, 要复制 DNA 或 RNA 以制造新细胞 (繁殖是生命的标志), 都需要大量的由 DNA / RNA 编码的蛋白质。因此, 在生命一开始, 就必须同时有 DNA / RNA 和蛋白质: 这是生命自然发生假说一个无法解决的“先有鸡还是先有蛋”的难题。

因此, 对于那些坚持仅仅通过自然的过程 (物理和化学) 就能产生生命的人来说, 生命的起源是一个棘手的问题。

一些进化论者声称生命的起源不是进化的一部分。然而, 几乎每一本进化生物学教科书都有关于生命起源的章节。加州大学伯克利分校的“进化 101”课程中就包含生命起源的内容, 其标题为“从原始汤到细胞——生命的起源”。¹“万物源于进化”理论的高调捍卫者们, 例如迈尔斯 (P.Z. Myers) 和尼克·马茨克 (Nick Matzke) 等人都同意生命的起源是进化的一部分, 理查德·道金斯 (Richard

Dawkins) 也是如此。²

之前著名的进化论者克尔库特 (G.A. Kerkut) 的确曾将进化论分为两种: 包含生命起源的广义进化论 (General Theory of Evolution, GTE), 以及仅涉及生命多样性的狭义进化论 (Special Theory of Evolution, STE), 后者据称是达尔文 1859 年的著作: 《物种起源》的主题。³

直到最近, 一些进化论的捍卫者才试图将生命起源与进化分开。

这可能是因为近来的生命科学研究, 即使在最简单的活细胞中, 也发现了一个又一个复杂的分子机器, 使得自然主义 (生命自然产生) 的问题变得越来越困难, 解决这些难题的希望正在迅速消失。

那么, 生命的产生需要什么呢? 我们可以将生命起源的问题分解为许多子话题, 以向大众解释生命起源相关的内容 (尽管这仍然可能是比较烧脑的)。

形成活细胞需要什么材料? 活细胞能够从周围环境中获取所需的所有资源并自我复制。第一个细胞必须能够独立生存: 也就是说, 它不能依赖于其他细胞而存活, 因为其他细胞不存在。寄生生物不能用作“第一个生命”的模型, 因为它们需要其他的细胞才能存活。基于这个原因, 病毒之类也不可能是“生命的先驱”, 因为它们必须寄生于活细胞内才能进行自我复制。朊病毒, 致病性畸形蛋白质, 同样与生命起源无关, 因为它们的“复制”不过就是让宿主细胞产生的蛋白质变形。

首先需要的是合适的材料。这有点像做蛋糕；如果没有香蕉或面粉，你就不能做出香蕉蛋糕。

备齐合适的材料

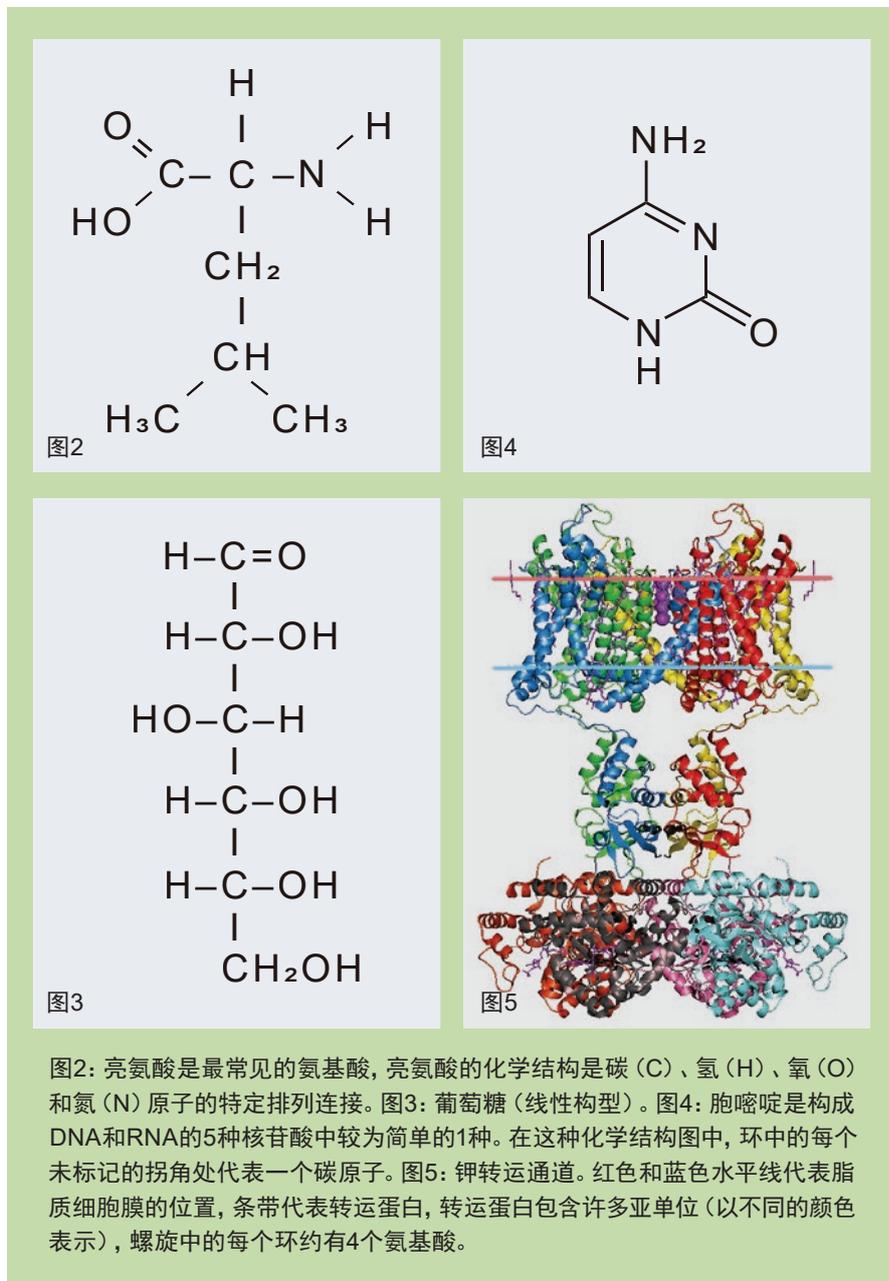
“生命起源于化学汤”的假说中存在一个主要的问题：所有的生命材料必须存在于同一个地方，这样才有可能组装出一个活细胞。但生命所必需的化学材料具有羰基(>C=O)，羰基化合物与氨基酸和其他氨基(-NH₂)化合物会发生破坏性的反应。含羰基的分子包括“糖”，⁴糖也是构成DNA和RNA骨架的材料。在活细胞中，羰基化合物与氨基化合物被分隔在不同的区域内保护着，以防止这种交叉反应，有些细胞可以在交叉反应发生后进行修复，但是化学汤没有这种保护或修复机制。

细胞是由较简单的化学物质组成的极其复杂的结构。本文不打算详述第一个细胞所需要的每种化学物质；这需要一本书的内容或更多的篇幅才能说清。本文重点介绍在任何生命起源模式中都必须存在的一些最基本的生命材料。

a. 氨基酸

生物体中充满了蛋白质，蛋白质是由线性串联的氨基酸组成的。例如，酶是特殊的蛋白质，用于催化化学反应（催化剂）。比如在我们的唾液中的淀粉酶，可以将大米、面包、马铃薯等食物中的淀粉分子分解成较小的分子，然后这些小分子可以被进一步分解成葡萄糖分子（淀粉分子的组成单元）。我们不能吸收淀粉，但我们能够吸收葡萄糖并将其作为能源来驱动我们的身体。

生命所必需的一些反应在没有酶的条件下进行得非常缓慢，即使有数十亿年的时间，也不能产生足



够的产物以供生命的需要。⁵

其他蛋白质组成了肌肉、骨骼、皮肤、毛发以及细胞和身体的各种结构部分。人类可以生产超过100,000种蛋白质(还可能是数百万种，没有人确切知道是多少种)，而典型的细菌可以产生超过一两千种不同的蛋白质。

蛋白质由20种不同的氨基酸组成(某些微生物拥有一种或两种额外的氨基酸)。氨基酸不是简单的化学物质，如果没有酶(酶本身

也是由氨基酸组成的)，它们就不容易正确地合成，见图1。

几乎每一本生物学教科书仍然在介绍1953年发表的米勒-尤里(Miller-Urey)实验，这个实验在没有酶的条件下制造了一些氨基酸，因此它通常被解释为“生命起源”的模拟实验，但这种解释要么是非常无知，要么是存心欺骗。

虽然这个实验制造了微量的正确氨基酸，但是这个实验所设定的

条件从来没有在地球上发生过。例如，烧瓶中的“大气”中如果含有微量氧气，就不能合成任何氨基酸。另外这个实验还同时制造了一些错误类型的氨基酸，以及其他会发生“交叉反应”的化学物质，它们会妨碍任何有用物质的形成。

在自然界中，功能性蛋白质所需的氨基酸绝不可能通过这个实验所模拟的条件来制造。⁶ 当斯坦利·米勒（Stanley Miller）在 1983 年用一种稍微更现实的气体混合物重复做这个实验时，他只得到了微量的甘氨酸，这是制造蛋白质所需的 20 种氨基酸中最简单的一种。⁷

如何才能得到正确氨基酸的混合物仍然是一个未解之谜（而且还有另外一个主要的问题，请看下面“手性”部分）。

b. 糖

有些糖可以仅通过化学方法制造，而不需要酶的参与（请记住，酶只能由细胞制造）。然而，在没有酶参与的条件制造糖需要一个碱性的环境，而在这种环境下氨基酸是不能合成的。

在合成糖的化学反应过程中，要求环境中没有含氮化合物，如氨基酸，因为它们会与制造糖的中间产物——甲醛，以及糖发生反应，产生出没有用（非生物）的化学物质。

核糖构成 RNA 的骨架，核糖经改造后构成 DNA 的骨架，二者都是所有活细胞的重要组成部分。核糖在化学进化中是一个难题。因为在现实世界中，在接近中性 pH（既非酸性也非碱性）的环境中，核糖是一种不稳定的糖（它会快速地分解），半衰期太短。⁸

c. DNA和RNA的部件

在没有来自活细胞的酶的帮助下，我们怎样才能得到 DNA 和 RNA 中的化学“字母”——核苷酸呢？在没有酶参与的情况下，使用化学方法制造核苷酸，就需要甲醛（ $H_2C=O$ ）与氰化氢（ $HC\equiv N$ ）进行反应。然而，甲醛和氰化物都是致命的剧毒化学品（特别是后者）。它们将非常严重地破坏重要的蛋白质（如果有的话）！

胞嘧啶（图 3）是 DNA 和 RNA 中 5 种必需的核苷酸碱基中的 1 种，在生命产生之前的任何现实的情景中都很难以制造，也非常不稳定。⁸

DNA 和 RNA 都具有糖和磷酸基团交替连接构成的主链骨架。上面已经讨论了制造糖的问题。磷酸离子会被海水中丰富的钙离子所沉淀，或粘附在粘土颗粒的表面上，这两种情况都会阻止磷酸被用来制造 DNA。

d. 脂类

脂类（“油脂”）对于形成细胞膜是必不可少的，后者包裹细胞内容物并执行其他细胞功能。细胞膜由几种不同的复杂脂类组成，是独立生存、能自我繁殖的细胞的重要组成部分。

脂类比糖或氨基酸具有更高密度的能量，因此它们在化学汤中的形成是任何生命起源理论的一个难题（从热力学的角度来看，含有高能量的化合物比低能量的化合物更加难以形成）。

即使假设没有氧气存在（在一个“还原性”的大气中），作为所有细胞膜的主要成分的脂肪酸也是很难生产的。即使产生了脂肪酸分子，生命所必需的离子，如镁和钙离子（每个原子具有两个电荷， $2+$ ，即二价），也会与脂肪酸结合，

并使脂肪酸沉淀而失去作用。⁹ 这如同在（富含镁和钙离子的）硬水中使用肥皂（主要成分是脂肪酸盐）进行清洗，就会发生相同的沉淀反应，形成“浮渣”。

有些自然发生假说的倡导人喜欢画一个空心脂球（囊泡）的示意图，这样的简单结构在一定的条件下可以在试管中形成。然而，这种纯脂质的“膜”永远不会形成活细胞，因为细胞需要通过脂膜向内和向外运送物质。这种进出细胞的运输需要非常复杂的“蛋白质—脂质”复合物，称为转运通道，其运行方式类似于电动机泵。出入细胞的各种化学物质都有自己的特异性通道（好比用于抽水的泵不一定适合用于抽油）。许多泵能够借助于高能化合物（如 ATP）沿着与自然梯度相反的方向（从低浓度到高浓度）进行主动运输。即使在顺着自然梯度（从高浓度到低浓度）进行运输时，也通常需要载体蛋白来促进。

细胞膜还维持着细胞内酶活性所需的稳定的 pH，并保持着各种离子的恰当浓度（例如钠不能太多）。这需要运输通道（“泵”）在细胞的控制下特异性地运输氢离子（质子），这些泵具有很高的特异性。¹⁰

细胞膜的运输是如此重要，以至“大多数生物的基因组中，20-30% 的基因所编码的都是细胞膜蛋白”。¹¹ 已知最小的自主生存的生物，寄生性生殖支原体的基因组，在其 482 个蛋白编码基因中，26 个基因用于编码细胞膜运输蛋白¹²。

纯脂质的膜甚至不允许带正电荷的金属离子（例如钙、钾、镁、铁、锰等）或带负电的离子（如磷酸基、硫酸基）被动运动出入这种膜，而这些都是生命必需的营养物质。纯脂质的膜会排斥这些溶

于水而不溶于油脂的带电离子，事实上，这种脂肪膜还会阻止水的通过（试一下将橄榄油和水混合在一起）！因此膜转运蛋白是活细胞所必需的。

在 20 世纪 20 年代，生命起源于肥皂泡（脂质泡）的理论很受欢迎（奥帕林的“凝聚”假设），但那时候人们还未知道生命中 DNA 和蛋白质合成的极度复杂性，也不知道细胞膜奇妙的功能。这些假说极端天真，但是相关的视频今天仍然可以在 YouTube 见到，显示脂质泡，甚至还能分裂，好像这就解释生命是怎样起源的（见视频：能自我创造的细胞？当然不是！）

e.手性

氨基酸、糖和许多其他生物化学物质都有三维结构，通常有两种彼此互为镜像的结构；像你的右手和左手是彼此互为镜像一样。这被称为手性（图 6）。

现有生物体内的生物化学物质都是由一种手性结构成分构成的（纯手性）：例如所有的氨基酸都是左手性（左旋），所有的糖都是右手性（右旋）。这就产生了一个问题，在没有酶参与的化学反应中（如米勒 - 尤里实验），当其生成氨基酸时，会同时产生右手性和左手性氨基酸的混合物。化学方法（例如甲酸反应）合成的糖也有同样的问题。¹³

研究生命起源的科研人员一直在设法解决这个问题。他们提出了各种可能的方案，但此问题仍未得到解决。¹⁴ 即使通过一些完全人工的方法，获得了 99% 手性纯度的氨基酸（这通过“自然”的机制不可能获得），也不能解决这个问题。生命需要 100% 纯左手性的氨基酸。原因是：若是蛋白质中的一个左手性氨基酸被右手性氨基酸替代了，就会使蛋白质产生不同的三维形状。

而要得到生命所必需的蛋白质，这种情况一次也不允许发生。

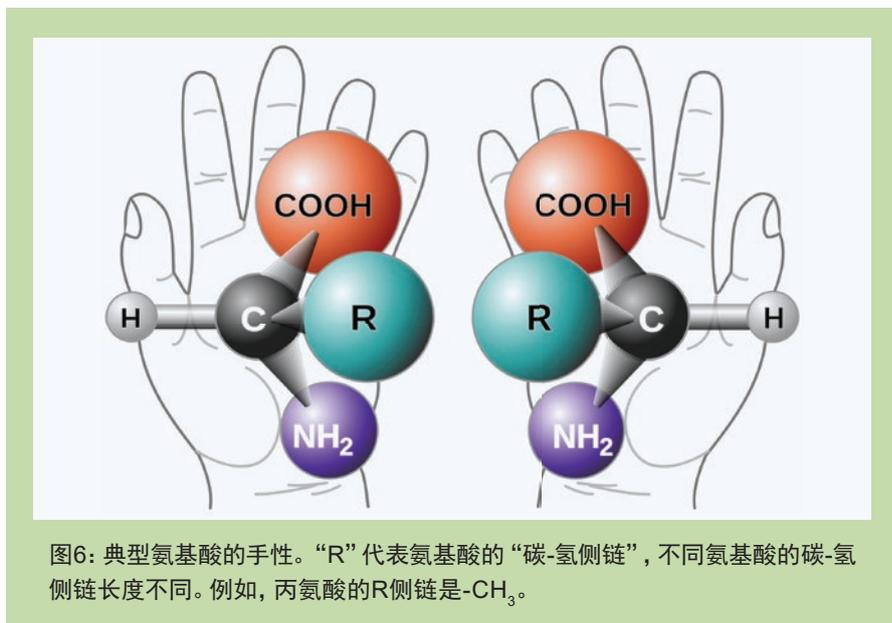
细胞存活的最低要求是什么？

独立生存的细胞若要使用从周围环境中获得的化学物质和能量来制造细胞部件并能自我复制，就必须具有：

1. 细胞膜：该结构能将细胞与环境隔开。细胞膜必须能够维持细胞内与细胞外不同的化学环境（如上文所述）。若没有细胞膜，生命的化学过程是不可能进行的。
2. 一种存储信息或遗传指令的方法，来指导细胞如何复制出另一个细胞（繁殖）以及在不同的情况下如何进行细胞内的运作。唯一已知的信息保存方法是使用 DNA，使用任何其他材料（例如 RNA）的提议尚未被证明是可行的——如果使用其他材料，那就还需要设法将该系统中的信息转化为 DNA，后者是所有已知生命的基础。¹⁵
3. 读取（2）中的信息，并在合适的时间、按照适当的数量制造细胞的构成材料。细胞主要

的构成材料是蛋白质，它是由数百至数千个大概 20 种不同的氨基酸分子聚合串联成的大分子（聚合物）。从 DNA 指令中制造蛋白质的唯一已知（或甚至可想到的）方法，需要 100 多种蛋白质和其他的辅助因子共同参与，至少包括：

- a. 纳米机器，例如 RNA 聚合酶（已知最小的 RNA 聚合酶大概含有 4,500 个氨基酸）
- b. 旋转酶：增加或解除 DNA 的螺旋，使 DNA 中的信息能够被“读取”（这些是体积很大的蛋白质）
- c. 核糖体：制造蛋白质的亚细胞“工厂”
- d. 至少 20 种转运 RNA 分子，这些转运 RNA 分子可以选择正确的氨基酸，将其按照 DNA 上指定的顺序排列。已知所有的细胞至少有 61 种转运 RNA 分子，因为大多数氨基酸对应着不止一种 DNA 密码（每个密码由 DNA 上的三个



字母组成)。转运 RNA 具有根据 DNA 密码来确保选择正确氨基酸的复杂机制。

- e. 还需要一些机制以确保所制造的蛋白质以正确的方式折叠成正确的三维结构, 包括以分子伴侣来保护蛋白质避免错误的折叠, 还有称为伴侣蛋白的折叠“机器”, 新制造的蛋白质会进入伴侣蛋白中, 使新造的蛋白能够进行正确的折叠。所有细胞都有这些机制。

(www.youtube.com/watch?time_continue=115&v=LY0hZLDOb00, 英语解说)

这是一个简要的蛋白质合成动画, 其中包括 RNA 聚合酶、核糖体、转运 RNA、伴侣蛋白和分子伴侣的作用。所有活细胞都具有这种蛋白合成系统。

哇! 上述只是最基本的机制。

需要一种从环境中的简单化学物质制造细胞所需要的生化分子的方法。这包括制造 ATP, ATP 是

生命的通用能量分子。已知所有的活细胞都有 ATP 合成酶, 一种非常复杂和高效的电动旋转马达, 可以制造 ATP (这种马达也可以反过来使用 ATP 产生电流, 以驱动细胞内/外的其他反应和运动)。

需要一种复制信息并将其传递给后代的方法 (繁殖)。最近有人模拟了已知最简单的、能独立生存的细菌的一次细胞分裂过程 (这个细菌“只有”525 个基因), 模拟实验需要 128 台台式电脑一起工作 10 个小时。¹⁶

以上列出一些基本条件, 以帮助读者理解第一个活细胞的生存需要什么前提条件。

几年前有人启动了一个有趣的研究项目, 要确定能独立存活的最简单细胞 (就是一个不依赖于另一个生物体就可以生存的最小的细胞) 是什么样的。然而, 这个细胞还是需要富含营养的培养基来为它提供大量复杂的有机化合物, 因此它所需要的许多生物化学物质都不必自己合成。现在已知这个最简单的细胞也需要超过 400 种蛋白质和 RNA 零件,¹⁷ 当然, 这意味着它的 DNA 需要含有制造这些蛋白

质和 RNA 的指令。也就是说, 它的 DNA 需要含有超过 400 个“基因”。我们稍后会继续讨论这个问题。

聚合物的形成 (聚合反应)

生命不仅仅由氨基酸或糖组成, 而且还需要很多聚合物, 它们是由较简单的化合物以串或链的方式连接在一起的。多糖是糖的聚合物; 蛋白质是氨基酸的聚合物; DNA 和 RNA 是核苷酸的聚合物。多糖是三者之中最简单的, 其链中连接着的通常都是相同的糖分子, 如葡萄糖在植物中聚合成淀粉, 在动物中聚合成糖原。蛋白质比多糖复杂得多, 其链中的每个氨基酸可以是 20 种不同的氨基酸中的任何一种。DNA 和 RNA 链由四种不同核苷酸连接而成。

另一方面, 水是活细胞的必需成分, 典型的细菌中的水占其总质量约 75%。水作为“通用溶剂”, 是细胞各种成分的必要载体, 是一切生命活动的环境。

这就为生命起源理论带来一个极大的难题。比如, 当 2 个氨基酸分子聚合时, 会释放一个水分子。这意味着在有水的情况下, 水会推动逆向反应, 造成聚合物的分解 (水解); 也就是说, 除非水能被主动去除, 否则蛋白质会分解而不会形成。细胞之所以能克服这一难题, 是靠着核糖体保护反应位点, 借助于能量不让水进入反应位点, 并提供能量来驱动聚合物的形成。因此, 对于所有的生命起源理论而言, 如何才能形成长度超过几个氨基酸的蛋白质是极为困难的问题 (再多的时间也不能解决这个问题, 延长时间只能导致更多蛋白质被水解)。

聚合物的形成还需要被连接在一起的每一个零件 (单体) 都具备两个官能团。这就意味着用于制造

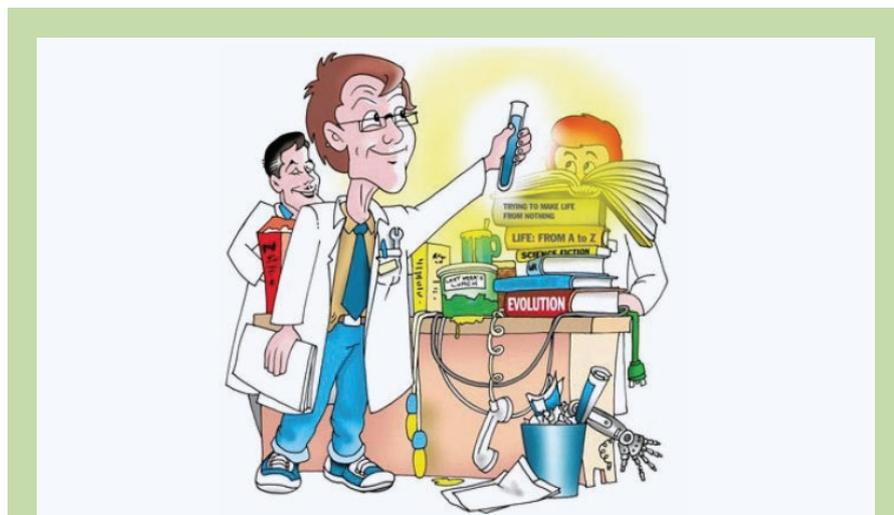


图7: 假如能在试管中制造生命并不证明生命可以在没有智慧参与的情况下自我创造。

蛋白质的氨基酸（或用于制造多糖的单糖）至少具有两个活性位点，以允许另一个氨基酸（或单糖）连接到它的两侧。形成蛋白质的氨基酸须具有至少一个氨基（-NH₂）和一个羧基（-COOH），每一个氨基酸的氨基与另一个氨基酸的羧基连接，从而使氨基酸链不断延长。仅具有一个活性位点（单官能团）的化合物会终止聚合物链的形成。生命起源假说的问题在于，任何可能产生氨基酸的化学反应，也会产生单官能团的物质，后者会阻止蛋白质的形成。¹⁸

核酸（如 DNA 和 RNA）的主链骨架是糖的聚合物。同样，如果存在一些单官能团的糖就会终止 DNA / RNA 的形成，而且在有水的环境中，水分子会促进 DNA / RNA 的分解，而不是聚合。

生命的起源是编程问题，而不仅仅是化学问题

上述科学事实足以完全否定自然主义的生命起源假说，但我们还未提及最重要的问题，即生命程序的起源。生命不仅仅依赖于聚合物，而且更依赖于聚合物中亚单位的特定排列方式；氨基酸的特定排列产生具有功能的蛋白质 / 酶；核苷酸碱基的特定排列产生具有功能的 DNA 和 RNA。

亚利桑那州立大学“超越科学基本概念中心”主任、天体生物学家保罗·戴维斯（Paul Davies）说：

“要解释生命是如何开始的，我们需要了解生命独特的信息管理是如何产生的。”

“生命管理信息的方式有一种逻辑结构，这与单纯的复杂化学有着根本的区别。因此，单纯的化学过程并不能解释生命的起源，就好比对硅、铜和塑料的研究并不能解释电脑是如何运行程序的。”¹⁹

戴维斯明确清晰地提出了这个问题，应该不会让他的进化论同事感到惊讶，因为这与他十多年前的公开言论很相似。例如“真正神秘的是活细胞中的软件，而不是其硬件。”他还说：“无知的原子如何能自发地编写自己的软件呢？……没人知道……”。²⁰

任何试图解释生命自然发生的理论，若没有解释信息处理系统是如何起源的，若没有解释记录在活细胞 DNA 里的信息，那都是在回避问题。我们只需要看看最简单的能独立生存的细胞就可以断定，在仅依赖于物理和化学过程的理论里（自然主义不允许智能设计的解释），信息的起源是不可能解决的。

二十世纪最杰出的科学哲学家之一，卡尔·波普尔（Karl Popper）爵士意识到：

“生命和遗传密码是怎么样自然产生的是一个令人不安的谜团，其原因是：遗传信息没有任何生物学功能，除非它被翻译；也就是说，除非其中的遗传信息指导了蛋白质的合成，即蛋白质的结构由遗传信息确定。但是……细胞（至少是现代细胞——也是我们唯一知道的）翻译遗传信息所需要的分子机器至少由 50 个大分子成分组成，而这些大分子本身也编码在 DNA 中。因此，若不是使用某些翻译产物，就不能翻译遗传密码。这构成了一个令人困惑的循环；对于任何关于遗传密码起源的模型或理论，这都似乎是一个真正的死循环。”

“因此，我们或许要面对一个可能性，即生命

的起源（就如物理学的起源一样）成为科学上一个不可逾越的界限；是在一切试图将生物学简化为化学反应和物理学作用的努力之后，最终残余（无法解决）的一个问题。”²¹

DNA 密码的起源

波普尔所描述的 DNA 信息存储系统不能源于化学过程，而必须有一个智慧的来源。²² 如果我们考虑一下其他的编码系统，例如摩尔斯电码或书面的字母信息，其中的符号用于表示语言的发音，这种编码系统只能来源于智慧。例如，“a”在英语中通常如它在“cat”中的发音，这是一个随意形成的惯例；字母的形状并没有表明它应该如何发音。同样，从物理和化学定律解释 DNA 编码系统是不可能的，因为代码和代码记录的信息之间没有物理或化学关系。

用 DNA 序列来编码蛋白质可以有亿万种方式。任何一套 DNA 密码的起源都是一个大难题，更何况现有的遗传密码恰恰是“达到或非常接近于错误最小化的最优总体方案，即所有可能的代码方案中最好的。”²³ 这是因为，用 64 个“密码子”²⁴ 来编码 20 种氨基酸，几乎所有的氨基酸都对应不止一个密码子（有几个常见的氨基酸，如亮氨酸，有六个密码子）。²⁵ 这些多重密码子有时被称为“重叠”，通常被认为是“超出了需要”或者“多余”的设计。然而，额外的密码子是经过优化的，使得 DNA 编码中最容易发生的单字母错误（突变）一般地不至于造成氨基酸水平上的改变；即使有改变，也是变成化学结构相似的氨基酸（因此对制造出的蛋白质的结构破坏性不大）。

通过额外的密码子还能进行

“翻译水平控制”，以实现蛋白质合成数量的复杂控制。该控制系统在细菌和高等生物中都会起作用。²⁶

密码系统不可能分阶段出现，逐步优化。假设通过某种令人难以置信的侥幸，出现了一个可以运行的编码系统，那么在此系统出现之后，基本代码就不会发生重大的变化，因为代码和解码系统（读码机器）必须同时改变，（例如，在一些细菌的基本代码中存在一个非常微小的变化，即使用三个正常的“终止”密码子中的一个，编码了一个额外的氨基酸²⁰）。所以除非在生命出现之初又发生了一系列难以置信的“自然”侥幸事件，否则最优化的密码系统无法解释。

除了编码系统，还有信息

化学进化不仅需要解释信息存储系统是怎么来的，还要解释存储在 DNA 上的蛋白质信息或指令。我们再来看最简单的细胞，科学家通过从自主生存的微生物中敲除基因来确定哪些是“必需的”，得知这种最简单的细胞需要超过 400 种蛋白质和 RNA 零件。所有这些信息必须在 DNA 上编码，否则假想中的第一个细胞就不会制造这些零件，更不能自我繁殖以产生另一个细胞。若将 DNA 的四个“字母”所编码的、最简单的细胞所必须的信息打印出来，会是一本很厚的书。

按照保罗·戴维斯的比喻，遗传信息的问题类似于电脑的程序。我们如何解释电脑程序的存在？首先要有编程语言（Python、Fortran、C++、Basic、Java 等），然后是用该编程语言撰写的一套实际指令。DNA 起源的问题同样是双重的：编程语言的起源，以及在 DNA 中记录的程序的起源。

从更简单的东西“进化”到最

简单的细胞，如此提议的人需要显示从他们假设的更简单的东西进化到第一个活细胞的途径。热衷于生命自然发生假说的人经常使用（假想的）“数十亿年”作为解决难题的忽悠方法，但并不给出怎么样进化的机制。即使加添再多的时间，朝着错误方向（如分解）进行的反应也不会逆转，不会自己朝着正确的方向进行。

生命也需要纠错系统

分子生物学揭示，即使是最简单的细胞也是极度复杂和精妙的。如上所述，信息存储在 DNA 中，然而 DNA 是一种非常不稳定的分子。一份报告说：

“‘人们普遍认为 DNA 是坚如磐石——非常稳定的’，范德堡大学（Vanderbilt University）的生物科学副教授勃兰·艾希曼（Brandt Eichman）指出，‘但实际上 DNA 反应性很高。在普通的一天，人体细胞内的 DNA 中约有一百万个碱基被破坏。’”²⁷

因此，所有细胞必须具有纠正 DNA 结构改变或编码信息错误的系统。如果没有这些纠错系统，DNA 序列中的错误就会累积并最终导致细胞死亡（称为“错误灾难”）。所有活细胞都有这一系统，这又让生命（自然）起源的理论雪上加霜。

假设原始汤中出现了一个带有信息的 DNA 分子，这个 DNA 分子必须被准确地复制，否则信息会因复制错误和化学损害而丢失。如果没有有效的 DNA 修复机制，遗传信息将迅速被降解。然而，构建这种修复机器的指令是在它要修复的 DNA 分子上编码的，这是生命起源理论的另一个死循环。²⁸

当科学家发现生活在极端条件下的细菌时，例如海洋中的深海热液喷口，喷口附近的细菌被称为“原始生命”，因为一些研究生命起源的人提出生命可能是在这些地方开始的。然而，这些被称为“极端微生物”（因为喜欢极端环境）具有相当复杂的 DNA 纠错系统。例如，耐辐射奇球菌是一种可以耐受极端剂量电离辐射的细菌，这个剂量会杀死你、我或其他细菌。奇球菌的 DNA 也会受到损伤，甚至会被打成许多碎片。但是在损伤后的数小时内，细菌能激活大约 60 个基因来修补断裂的 DNA，重建基因组。²⁹

热液喷口是不适宜生存的高温环境，生活在那里的微生物的 DNA 不断受到破坏，因此它们必须具有复杂的 DNA 保护和纠错系统才能生存。它们一点都不简单，对于生命起源也没有提供任何可行的模型。³⁰

此外，所有细菌，不仅仅是“极端微生物”，都必须拥有复杂的纠错系统，这些系统涉及许多基因，当纠错系统因突变而失效时，细菌将无法生存。这又为生命起源提出了另一个难题。

生命起源的环境

据称生命是起源于一个温暖的池塘（如达尔文所推测的）、深海热液喷口附近、粘土颗粒表面？还是在别的地方以其他方式起源的？这么多环境被提出来，但又没有定论，就表示这些环境都存在重大缺陷。

温暖的池塘和深海热液喷口的一个主要问题是水的存在，水会阻碍许多必需的反应，如聚合反应。此外，深海热液喷口的热量会使任何侥幸形成的化学分子被加速分解。

由于水的存在造成这许多问

题，致力于研究生命起源的物理化学家格拉哈姆·凯恩斯-史密斯（Graham Cairns-Smith）提出，粘土表面能促进一些所需的反应。

然而，在火山边温暖的水池中进行的实验表明，粘土颗粒与氨基酸、DNA 和磷酸基等形成生命的组分结合得太紧密，因此粘土会阻止这些物质之间发生必要的反应。³¹

如上所述，含有 DNA、蛋白质和 RNA 的、能够繁殖的细胞绝不会在化学汤中偶然起源。所以，倡导生命自然发生的学者们又想象出另一种情况：刚开始时生命的需求比较简单，然后逐步进化到如我们今天所知的生命。

先出现蛋白质？

大多数研究集中在“先出现蛋白质”的起源方式上。根据这种途径，蛋白质首先出现，然后出现了 DNA 序列并能够产生所需的蛋白质，随后从 DNA 序列制造了合成蛋白质所必需的 RNA。然而，除了获得一整套 100% 手性正确的氨基酸材料的难题，以及氨基酸聚合制造蛋白质链的难题之外，几乎没有蛋白质可以作为模板来复制自己。³² 此外，正如信息理论家休伯特·约克奇（Hubert Yockey）所指出的那样，一个基本难题是没有任何机制能够从蛋白质制造编码这个蛋白质的 DNA 序列。³³

先出现 RNA？

在 20 世纪 80 年代，研究者发现了一些能够催化某些化学反应的 RNA 分子，称之为“核酶”（核糖核酸酶）。这一发现令人振奋，导致很多人将精力花费在“RNA 第一”或“RNA 世界”的理论上。至少有一些酶可以根据 RNA 信息制造 DNA 信息，也就是说，如

果你能得到 RNA，就可以想象出获得 DNA 的途径。然而，能够从 RNA 序列来制造 DNA 拷贝的酶复合物非常复杂，这些酶复合物本身绝不可能通过自然过程而出现。在“RNA 第一”的理论中还有许多看似无法克服的其他困难，凯恩斯-史密斯列举了其中的 19 个难题。³⁴ 此外，如上所述，DNA 非常不稳定，但 RNA 的稳定性比 DNA 差得多。

人们提出了许多种生命起源可能发生的环境，这反而强化了一个结论：研究人员其实不知道生命是如何“自我创造”的。对于生命是如何从简单状态开始并逐步进化成一个真正的活细胞，没有人能提出可行的假说。新达尔文主义（即突变和自然选择）常被用来“跨越无望巅峰”，但是在一个能够存活且能自我繁殖的实体（即细胞，我之前在“细胞存活的最低要求是什么？”一节中已经列举了它的最低需求）出现之前，突变和自然选择从理论上就完全不起作用。

生命来自外太空？

弗朗西斯·克里克（Francis Crick）是 DNA 双螺旋结构的发现者之一，也是“生命源自外太空”的著名支持者。³⁵ 他提出外星人将生命送到地球，称为“定向泛种论（directed panspermia）”。这种想法的另一种形式，简称“泛种论（panspermia）”，就是生命在宇宙中的其他地方出现，并以微生物的形式，通过陨石或彗星来到地球，以这种方式向地球“播种”了生命。两种版本的泛种论都有效地将生命起源的难题推出了科学的范围以外。关于泛种论假说，唯一可测试的环节是微生物在冲向地球的陨石上或在陨石中的存活能力。这个问题已经过测试，结果对泛种论不利，微生物无法存活。³⁶

人们进行外星文明探索（SETI）和太阳系外行星探索的原动力在很大程度上是为了发现生命在外太空形成的证据。但即使将整个宇宙作为实验室也无法解决生命起源的问题：生命永远不会自发形成，请见下面的概率计算。

生命起源的概率计算

对于化学物质演化成生命的概率，已经有许多人进行了尝试性的计算，但所有这些尝试都使用了简化的、让生命自发起源更有可能性（即概率 > 0）的假设。

数学家弗雷德·霍伊尔（Fred Hoyle）爵士以各种方式说明了生命自然形成的极端不可能性，甚至连一种有功能的聚合物分子，如蛋白质，也不可能自发形成。霍伊尔说，“现在想象有 10^{50} 个盲人 [编辑补充：如果他们肩并肩站立，整个行星系都容不下他们]，每个人都转动着一个被打乱的魔方，试着想象他们同时让魔方各面都转到相同颜色的几率。如果上述目标能够达到的话，你才有机会通过随机排列的方式，获得生命所依赖的众多生物聚合物中的一个。除了生物聚合物之外，就连活细胞的操作程序（遗传信息）也要在地球上的原始汤中随机形成，这个说法简直是痴



图8：魔方

<http://www.wonderopolis.org>

人说梦。生命必定是一种宇宙现象。”³⁷

实际上，我们可以计算出只获得一个长度为150个氨基酸的小蛋白质的概率，假设只存在（手性）正确的氨基酸，并假设它们能以某种方法聚合在一起。有20种不同的氨基酸，150个氨基酸的可能排列的数量是 20^{150} ，或者说得到正确的蛋白质的概率约为 10^{195} 分之一。为了避免有人反对说：不是每个氨基酸都必须按照确切的顺序排列，这只是使用一种小蛋白质进行计算，并且只计算几百种生命必须蛋白质中的一种（现实中许多蛋白质的氨基酸长度比150要大得多），如果同时计算DNA序列自发形成的概率，会使这个问题严重地复杂化。实际上，有些蛋白质的序列即使发生了少量改变，也会失去全部功能。³⁸

那时霍伊尔认为生命必须来自外太空。后来他意识到，在没有指导（没有智慧参与）的条件下，即使将整个宇宙作为实验室，生命也不可能 anywhere 通过物理和化学的过程来形成：

“从无生命物质中形成生命的可能性是 $10^{4,000}$ 分之一……它足以埋葬达尔文和整个进化论。没有原始汤，在这个星球上和任何其他星球上都没有，如果生命的起源不是随机的，那么生命肯定是有目的的智慧产物。”³⁹

如果没有目的和智慧，那么 $10^{4,000}$ 分之一这个数字能不能完全排除生命在宇宙中任何地方自然发生的可能性？我们可以这么说吗？

麻省理工学院研究员塞思·劳埃德（Seth Lloyd）计算了所谓的宇宙大爆炸（据称是在137亿年前发生的）之后的一切可能发生的事

件（或称“基本逻辑运算”）的总数量，结果是不超过 10^{120} 。⁴⁰这就设定了理论上可能进行的实验数量的上限。此上限意味着概率为 $10^{4,000}$ 分之一的事件将永远不会发生，即使是我们计算的一种150个氨基酸的小蛋白质在整个宇宙中也不可能自发形成。

然而，生物物理学家哈罗德·莫罗维茨（Harold Morowitz）⁴¹得出的概率要低得多，为 $10^{10,000,000,000}$ 分之一。这是从含所有细菌基本零件的肉汤（例如，理论上可以通过加热含活细菌的培养液，杀死它们并将它们分解成基本成分而获得）自发组装而得到最简单细菌的机会。

莫罗维茨是无神论者，但他认为生命不是偶然因素的结果，而是出自可用之能量的某种属性，该属性驱动了能够使用能量的实体（即“生命”）的形成。这听起来很像地母盖亚的神话，因为它为宇宙赋予了泛神论的神秘属性。

最近，无神论哲学家托马斯·内格尔（Thomas Nagel）提出了类似的旨在解释生命和思想之起源的说法。⁴²

只要不用相信一个超自然的创造者，好像什么理论都可以提出来。

由于概率计算的困难，使用不同的假设，就会计算出不同的概率。如果我们使用了最有利于生命自然发生的假设进行计算，其结果仍然是极端的不可可能，那么这就是一个比使用更现实的假设更有力的论证。更现实的假设虽然让唯物主义的说法更加不可可能，但唯物主义者会质疑我们的某些假设。

然而，所有关于生命之化学起源的概率计算都会使用不切实际的、有利于生命化学起源的假设，否则概率就会为零。例如，莫罗维茨计算中使用的含有活细胞的所有成分的肉汤是不可能存在的，因为

如上所述，这些化学成分会互相反应，使得它们不能用于制造活细胞所需的复杂聚合物。

著名信息理论家休伯特·约克奇（加州大学伯克利分校）认识到了这个问题：

“原始汤中生命随机起源在概率上是不可能的，就像永动机在概率上不可能一样。虽然本章计算的极小概率并不会令真正的信徒沮丧……[然而]一个明白人必会得出结论：生命并非偶然形成的。”⁴³

请注意，在约克奇的计算中，他慷慨地假设生命原料可以在原始汤中获得。但是在他书的前一章中，约克奇证明原始汤本身就根本不存在，所以相信原始汤是一种“信仰”。他后来得出结论，“原始汤理论是卫道士们基于其意识形态的自我欺骗。”⁴⁴

更多的坦承

请注意，约克奇并不是唯一一位在这个问题上明确发表言论的知名学者：

“如果任何人告诉你，他或她知道34亿年前地球上的生命是如何开始的，这个人不是一个傻瓜，就是一个骗子。没有人知道。”——斯图亚特·考夫曼（Stuart Kauffman）教授，生命起源研究者，加拿大卡尔加里大学（University of Calgary）。⁴⁵

“……我们必须承认，关于任何生化系统或细胞系统的演变，目前尚没有详细的达尔文主义的解释，而只有各种一厢情愿的推测。”——弗兰

克林·哈罗德 (Franklin M. Harold), 科罗拉多州立大学生物化学与分子生物学荣誉教授。⁴⁶

“关于从非生命到生命的进化途径, 达尔文曾写道: ‘目前生命起源只是垃圾思想, 还不如想想物质的起源。’ 我们今天几乎处于同样的黑暗中” —— 亚利桑那州立大学超越科学基本概念中心主任保罗·戴维斯。⁴⁷

“细胞的创造性和复杂性远远超过当今世界上任何无生命的东西, 以至于我们不知道它是如何实现的。” —— 克斯纳 (M.W. Kirschner, 哈佛大学医学院系统生物学系主任、教授) 和格哈特 (J.C. Gerhart, 加州大学研究生院教授)。⁴⁸

“结论: 生命起源的科学课题可以归结为寻找一个化学机制, 能够实现从第一个自动催化的自我复制周期开始, 最后演化到一个共同祖先。目前所有的理论都远远未能完成这项任务。虽然我们仍然不清楚这种机制, 但我们现在已经理解了这个问题的广度。”⁴⁹

“进化论的最大缺口仍然是生命本身的起源……从这一堆分子 (氨基酸和 RNA) 得到即使是最原始的细胞, 差距仍然是巨大的。” —— 克里斯·威尔斯 (Chris Wills), 加州大学生物学教授。⁵⁰

即使是唯物论教条主义者理查德·道金斯 (Richard Dawkins) 也

向本·斯坦因 (Ben Stein, 见电影纪录片 *Expelled*) 承认没有人知道生命是如何开始的:

道金斯: “我们知道生命起源中必定发生的一件事, 就是出现第一个自我复制的分子。”

斯坦因: “那是怎么发生的?”

道金斯: “我告诉过你, 我们不知道。”

斯坦因: “所以你不知道它是如何开始的?”

道金斯: “不知道, 没人知道。”⁵¹

“我们永远不会知道生命最初是如何出现的。然而, 对生命起源的研究是一个成熟、完善的科学研究领域。正如进化生物学的其他领域一样, 对于最初生命形式的起源和本质问题的答案, 只能算是探索和解释, 而不是明确和肯定的。”⁵²

结论

若没有智慧, 单凭物理过程和化学反应就不能产生生命。创造生命, 即使是最简单的生命, 所需的智慧都远远超过人类的智慧; 人类仍然在摸索着试图完全理解最简单的生命形式是如何运作的。即使是最简单的细菌中也还有很多东西需要我们去研究。事实上, 随着我们研究发现的事实和规律越来越多, 生命起源的“问题”变得更加困难; 答案不是越来越接近, 而是越来越遥远。但真正的问题在于: 生命的起源大声告诉我们有一个超级智慧的创造者创造了生命, 而这正是今天的世俗头脑所不能接受的。

若是科学能“证明”上帝的存在, 生命的起源可以说是最好的证据。 ■

参考文献和注释

1. <http://evolution.berkeley.edu/evosite/evo101/IIE2aOriginoflife.shtml> (accessed 17 October 2013) .1.
2. Myers, P.Z., 15 misconceptions about evolution, 20 February 2008, scienceblogs.com; Matzke, N., What critics of neo-creationists get wrong: a reply to Gordy Slack, pandasthumb.org. 道金斯在他自称“证明了进化论”的著作《地球奇观》



图9: 如果我们能够在这里合成生命, 那么我就可以证明在起初不需要有智慧参与就能形成生命了!

- (*The Greatest Show on Earth*) 中试图解决生命起源的问题。参见 Sarfati, J., *The Greatest Hoax on Earth?* ch. 13, 2010, Creation Book Publishers.
3. Kerkut, G.A., *Implications of Evolution*, Pergamon, Oxford, UK, p. 157, 1960 (在线阅读 ia600409.us.archive.org/23/items/implicationsofev00kerk/implicationsofev00kerk.pdf); creation.com/evolution-definition-kerkut.
 4. 糖具有含羰基的线性结构 - 参见图 2. 核酸中出现的环状结构的糖在溶液中也占优势, 但与线性结构处于平衡状态。当某些物质与醛基强烈反应时, 会再生出更多的线性糖以代替反应消耗掉的那些线性糖, 最终所有的糖分子将被消耗掉。
 5. Sarfati, J., World record enzymes, *Journal of Creation* 19 (2) :13–14, 2005; creation.com/world-record-enzymes-richard-wolfenden.
 6. Bergman, J., Why the Miller-Urey research argues against abiogenesis.
 7. Truman, R., What biology textbooks never told you about evolution.
 8. Sarfati, J., Origin of life: instability of building blocks.
 9. Chadwick, A.V., Abiogenic Origin of Life: A Theory in Crisis, 2005; origins.swau.edu/papers/life/chadwick/default.html.
 10. 作为一例, 见 Potassium ion channel, hydrated ionic radii, creation.com/ionic-error, 21 August 2010.
 11. Krogh, A. et al., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, *Journal of Molecular Biology* 305 (3) :567–580, 2001; dx.doi.org/10.1006/jmbi.2000.4315.
 12. 生殖支原体 G-37 中的转运蛋白; 见 membranetransport.org/index.html, 2013 年 10 月 11 日。
 13. 手性中的“右”和“左”是指氨基酸的(费歇尔投影)标准化结构图上显示的氨基(NH₂)处于右边和左边的位置。
 14. Sarfati, J., Origin of life: the chirality problem; creation.com/origin-of-life-the-chirality-problem (updated 2010).
 15. Cairns-Smith, A.G., Evolutionist criticisms of the RNA World conjecture, from *Genetic Takeover and the Origin of Life*, 1982; creation.com/cairns-smith-detailed-criticisms-of-the-rna-world-hypothesis.
 16. Stanford researchers produce first complete computer model of an organism; news.stanford.edu, 19 July 2012.
 17. Sarfati, J., How simple can life be? <https://creation.com/how-simple-can-life-be>. 对人工合成细胞 JCVI-syn3.0 的研究显示, 473 个基因是必需的, 其中 65 个功能不明: C.A. Hutchison III et al., Design and synthesis of a minimal bacterial genome, *Science* 351:1414, March 25, 2016; doi: 10.1126/science.aad6253.
 18. Sarfati, J., Origin of life: the polymerization problem.
 19. Davies, P., The secret of life won't be cooked up in a chemistry lab: Life's origins may only be explained through a study of its unique management of information, *The Guardian*, Sunday 13 January 2013; guardian.co.uk/commentisfree/2013/jan/13/secret-life-unveiled-chemistry-lab. Return to text.
 20. Davies, P., Life force, *New Scientist* 163 (2204) :27–30, September 18, 1999.
 21. Popper, K.R., “Scientific reduction and the essential incompleteness of all science”; in Ayala, F. and Dobzhansky, T., (Eds.), *Studies in the Philosophy of Biology*, University of California Press, Berkeley, p. 270, 1974.
 22. Smith, C., Lost in translation: The genetic information code points to an intelligent source, 6 May 2010; creation.com/genetic-code-intelligence.
 23. Freeland, S.J., et al., Early fixation of an optimal genetic code, *Molecular Biology and Evolution* 17 (4) :511–18, 2000; mbe.oxfordjournals.org/content/17/4/511.full.
 24. DNA 有 4 个核苷酸“字母”(A、T、G、C), 分子阅读机器一次读取 3 个字母(即一个“密码子”), 如此就有 4x4x4 = 64 种可能的密码子排列。
 25. 通常有 3 个“终止”密码子作为蛋白质编码序列结束的标记, 因此正常有 61 个密码子用于氨基酸编码。
 26. Novoa, E.M. and de Pouplana, L.R., Speeding with control: codon usage, tRNAs, and ribosomes, *Trends in Genetics* 28 (11) :574–581, November 2012; www2.biol.sc.edu/~elygen/biol655/translation%20speed.pdf.
 27. Newly discovered DNA repair mechanism, *Science News*, sciencedaily.com, 5 October 2010.
 28. Sarfati, J., New DNA repair enzyme discovered, 13 January 2010; creation.com/DNA-repair-enzyme. Tomas Lindahl, Paul Modrich, 和 Aziz Sancar 因发现三种 DNA 修复机制获 2015 年诺贝尔化学奖: Batten, D., DNA repair mechanisms ‘shout’ creation, *Creation* 38 (2) :56, April 2016.
 29. Cox, M.M., Keck, J.L. and Battista, J.R., Rising from the Ashes: DNA Repair in *Deinococcus radiodurans*, *PLoS Genetics* 6 (1) :e1000815, 2010; doi:10.1371/journal.pgen.1000815.
 30. Catchpoole, D., Life at the extremes, *Creation* 24 (1) :40–44, 2001; creation.com/extreme-and-Sarfati, J., Hydrothermal origin of life? *Journal of Creation* 13 (2) :5–6, 1999; creation.com/hydrothermal.
 31. Morelle, R., Darwin’s warm pond idea is tested, 13 Feb. 2006; news.bbc.co.uk/2/hi/science/nature/4702336.stm.
 32. 朊病毒有时被认为是能复制的蛋白质, 但朊病毒是引起既有的蛋白质变形: 它们不会通过使氨基酸按朊病毒自己的顺序排列来复制朊病毒(朊病毒被认为会导致“疯牛病”)。
 33. Yockey, H., *Information Theory, Evolution and the Origin of Life*, Cambridge University Press, 2005, pp. 118–119.
 34. Evolutionist criticisms of the RNA World conjecture; Quotable Quote by Cairns-Smith; creation.com/cairns-smith-detailed-criticisms-of-the-rna-world-hypothesis. See also, Mills, G.C. and Kenyon, D., The RNA World: A Critique, *Origins & Design* 17 (1) ; am.org/docs/odesign/od171/rnaworld171.htm.
 35. Bates, G., *Designed by aliens?* *Creation* 25 (4) :54–55, 2003; creation.com/aliens.
 36. Sarfati, J., Panspermia theory burned to a crisp: bacteria couldn't survive on meteorite, 10 Oct 2008; creation.com/panspermia-theory-burned-to-a-crisp-bacteria-couldnt-survive-on-meteorite.
 37. Hoyle, Fred, The Big Bang in Astronomy, *New Scientist* 92:521–527, 1981.
 38. 例如, 罗尔·杜鲁门(Royal Truman) 研究了存在于真核生物中的泛素蛋白, 结果表明, 要保持其功能, 蛋白质序列几乎不允许改变, 因此这种蛋白质自然发生的可能性被排除了; 见 Truman, R., The ubiquitin protein: chance or design? *Journal of Creation* 19 (3) :116–127, 2005; creation.com/the-ubiquitin-protein-chance-or-design.
 39. Sir Fred Hoyle, as quoted by Lee Elliot Major, “Big enough to bury Darwin”. *Guardian* (UK) education supplement, Thursday August 23, 2001; education.guardian.co.uk/higher/physicalscience/story/0,9836,541468,00.html.
 40. Lloyd, Seth, Computational capacity of the universe, *Physics Review Letters* 88:237901, 2002; <http://arxiv.org/abs/quant-ph/0110141v1>.
 41. Morowitz, H., *Energy Flow in Biology*, Academic Press, NY, 1968.
 42. Nagel, T., *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*, Oxford University Press, 2012.
 43. Yockey, H., *Information Theory and Molecular Biology*, Cambridge University Press, 1992, p. 257.
 44. Ibid. p. 336; see Quotable quote: Primeval soup—failed paradigm.
 45. Stuart Kauffman, At Home in the Universe: *The Search for the Laws of Self Organization and Complexity*, Oxford University Press, p. 31, 1995.
 46. Harold, F.M., *The way of the cell: molecules, organisms and the order of life*, Oxford Uni. Press, New York, p. 205, 2001.
 47. Davies, Paul, The Cosmos Might Be Mostly Devoid of Life: We still have no idea how easy it is for life to arise—and it may be incredibly difficult, *Scientific American*, 1 September 2016; www.scientificamerican.com/article/the-cosmos-might-be-mostly-devoid-of-life.
 48. Kirschner, M.W. and Gerhart, J.C., *The plausibility of life: Resolving Darwin's Dilemma*, Yale University Press, New Haven and London, p. 256, 2005.
 49. Watchershauser, G., Origin of life: RNA world versus autocatalytic anabolism, *The Prokaryotes*, Vol. 1, 3rd edition, chapter 1.11, pp. 275–283, p. 282, 2006.
 50. 被引于 Evolution’s final frontiers, *New Scientist* 201 (2693) :42, 2009.
 51. *Expelled: no intelligence allowed*, Premise Films, 2008.
 52. Lazcano, Antonio, Historical Development of Origins Research, *Cold Spring Harbor Perspectives in Biology* 2 (11) :a002089, November 2010; doi: 10.1101/cshperspect.

唐·巴顿 (DON BATTEN)
农理学士荣誉学位、博士

唐·巴顿曾从事植物研究, 也曾任植物生理学顾问。多年全职从事创造事工, 目前任国际创造事工 CMI (澳大利亚) 总干事。更多作者相关信息, 参见: creation.com/batten.



揭露

用基因组来骗人

的把戏

Illustration: ©iStock.com/BlackJack3D

文/ 杰弗里·汤姆金斯 (Jeffrey Tomkins) 和 杰瑞·伯格曼 (Jerry Bergman)

本 研究汇总了多个关键研究发表的论文，分析文中概述的方法和数据，发现人类和黑猩猩的基因组所谓几乎相同的这个主流说法是非常值得怀疑的。他们报告中所谓的人类和黑猩猩的基因组 DNA 序列有很高相似度的估算结果主要基于预先筛选的生物样品和 / 或数据。而那些不相似而难以匹配的 DNA 序列通常会被抛弃、掩盖和 / 或不报告。此外，最终比对时出现（不一致）的间隙序列也经常被抛弃，以进一步夸大相似度的最终结果。这种高度选择性地抛弃数据的方法正是由达尔文教条所驱动的，这才产生了常被吹捧的骗人把戏：“人类与黑猩猩的 DNA 相似度达到 98%”。本研究对发表的各项论文，以及经常被引用的 2005 年黑猩猩基因组报告中提供的数据进行分析，可以保守地作出以下结论：人类与黑猩猩基因组的相似度不超过 87%，甚至可能小于 81%。这些修正后的相似度估算结果是基于（之前学者的）最终相似度评估中通常被抛弃的相关数据。

近来的一些研究报告证实了上述结论，最近人类和黑猩猩 Y 染色体之间的比较虽然不能作为全基因组相似度的估算，但发现人类和黑猩猩 Y 染色体之间的极大差异，是灵长类动物由共同祖先进化而来教条的一个不可逾越的困难，因为 Y 染色体是至今已知人类基因组中变异最小的染色体。最后，最近一次大规模的人类与黑猩猩基因组比较的研究报告极大地证实了本报告中提供的数据。因此所谓人类和黑猩猩拥有共同祖先的范式显然更多地基于神话和进化论信仰的宣传，而不是基于事实。

经常听到一个普遍说法，称黑猩猩（*Pan troglodytes*）和人类（*Homo sapiens*）大约有 98 % 的 DNA 是相似的。这种过于简化且经常被吹捧的估算结果实际上可能涉及两个完全独立的概念：

1. 基因含量，比较和计算在不同物种之间是否存在相似类型的基因编码序列的数量。
2. 比对不同物种之间 DNA 序列实际碱基对的相似度。

在大多数情况下，现代相似度比较范式是指 DNA 序列的比对研究。

生物体的 DNA 序列数据通常要先经过几个级别的预筛选，过滤和选择，（抛弃不理想的数据）后再进行汇总和讨论。正如下面几个例子，在整个比较遗传学研究领域中，主要问题之一就是大多数研究中，在最终分析之前，先对可用的生物样本和数据进行大量的预筛选。从大的数据库中通常只挑选最有希望的数据用于最终分析。当然，你可以只比对你所知道那些高度匹配的数据，否则在大多数情况下，序列难以进行比较。生物体的 DNA 序列数据经常经过几个级别的预筛选，过滤和选择后再进行汇总和讨论。在序列比对中完全不匹配的序列区域以及序列的间隙通常在最终比较结果中被抛弃，或其影响被掩盖。如下所述，这种抛弃或掩盖数据的方式有多种，因此必须根据每个已发表的研究论文进行逐个评估。

人类和黑猩猩相似性的早期研究使用了重组动力学方法

最初报告的人类和黑猩猩 DNA 的极高相似度的估算结果来自一个称为重组动力学的研究领域。这些初步的报告推动了牛津大学教授理查德·道金斯等受欢迎的进化论名人的早期主张，他说“黑猩猩

使用被抛弃的数据对

“人与黑猩猩的DNA几乎完全相同”

的故事进行重新评估

和我们人类有 99 % 以上的基因是相同的。”¹ 当时，这种说法是毫无根据的，因为在那时人类和黑猩猩的基因数量尚未清楚。人类和黑猩猩的基因组序列初稿直到 2001 年和 2005 年才公布。^{2、3、4、5}

而道金斯在 1986 年提及的基因相似度是基于人和黑猩猩 DNA 混合物的重组动力学实验的间接估算，而不是基于明确定义的基因。¹ 在重组动力学研究中，使用热和/或化学方法将双链的 DNA 分离成为单链。当 DNA 链以受控方式进行重新结合时，可以使用各种方法对其进行分析。重新结合得越慢的 DNA，被认为是越复杂且基因比较密集。通常，可以区分出三种类型的 DNA：高拷贝（高度重复，基因含量极少）、低拷贝（中等重复，基因较少），以及单拷贝（富含基因）。在比较研究的时候，从两个物种中收集 DNA 的单拷贝部分混合在一起，再使 DNA 的双链分离，然后使其重新结合，使人和黑猩猩的单链 DNA 有机会结合。结合后，人和黑猩猩的 DNA 链之间的互补碱基的匹配水平可以通过各种方法进行间接评估，包括测量 DNA 链重新结合的速率或水平。

需要注意的是，只有人类和黑猩猩基因组中的单拷贝部分片段被用于进行早期的相似度估算。由于基因含量高，因此科学家们专注于单拷贝的 DNA 区域。然而，位

于基因组的其他部分的许多基因，却被排除在相似度分析之外。另一个问题是，现在已经知道事实上几乎整个基因组（的每一部分）都是具有某些方面的功能，而且非编码区域已被证明有重要的作用，包括具有许多关键控制功能以及核苷酸模板。^{6、7、8}

第一个 99 % 的相似度宣称（科恩称之为“1 % 的神话”）是由艾伦·威尔逊和玛丽·克莱儿·金在 1975 年使用单拷贝 DNA 的重组动力学方法制造出来的。⁹ 其他类似的研究的单拷贝 DNA 重组动力学实验结果报告的 DNA 平均差异约为 1.5 %，因此就产生了广泛流传的“人与黑猩猩 DNA 序列相似度为 98.5 %”的说法。¹⁰⁻¹² 虽然实际上绝大部分人类和黑猩猩基因组被排除在早期的 DNA 比对实验之外，但在基因组中占比较小的单拷贝 DNA 区域存在所谓的“高度相似”也使研究人员感到惊讶。正如吉本斯评论的那样，进化论者得到的最终共识是：人类与黑猩猩在解剖和行为上的巨大差异是由于所谓的“很小的”遗传（DNA）差异所造成的。

基因组学研究——神话被肯定

后来的研究应用了 DNA 测序方法，并使用了之前建立了“高度相似教条”的重组动力学所用的单

论文作者	所分析的基因组 碱基对总数	能匹配的 碱基对数量	报告的 DNA 相似度	实际的 DNA 相似度 *
布里顿 (Britten), 2002	846,016	779,132	95.2%	~ 87%
埃伯斯贝格尔 等人 (Ebersberger et al.), 2002	3,000,286	1,944,162	98.8%	< 65%
刘教授 等人 (Liu et al.), 2003	10,600,000 (人, 黑猩猩, 狒狒和狨猴 DNA 分析的总数)	4,968,069 (人 - 黑猩猩)	98.9% (不计算 插入 / 缺失)	?
威德曼 等人 (Wildman et al.), 2003	~90,000 (97 个基因的 外显子)	?	98.4-99.4%	?
黑猩猩 22# 染色体协会 (Chimp. Chrom. 22 Consort.)	32,799,845	?	98.5% (不计算 插入 / 缺失)	80-85% (计算 包括插入 / 缺失)
尼尔森 等人 (Nielson et al.), 2005	?	?	99.4% (选择基 因的区域)	?
黑猩猩基因组测序协会 (Chimp. Seq. Consort.) 2005	全基因组 (五倍的冗余 测序)	2.4 Gb	95.8%	81%**

* : 基于比对中被抛弃的 DNA 序列的量

** : 与国际人类基因组测序协会 (2004 年) 的数据相比—— $[(0.9577 \times 2.4 \text{ Gb}) / 2.85 \text{ Gb}] \times 100$

? : 由于未提供 (被抛弃的) 数据, 因此无法计算实际相似度百分比。

表1: 人和黑猩猩基因组比较论文的汇总。在条件允许的情况下, 使用来自报告的排除在比对之外的数据来计算实际的DNA相似度百分比。

拷贝 DNA 序列。在另外一篇文章中, 我们讨论了在涉及关于人类与黑猩猩相似度时, 重组动力学的初步研究已经建立了一个不言而喻的、基于教条的 (不能被挑战的) “黄金标准”。¹⁴

创造论者托德·伍德撰写了一篇关于人类和黑猩猩之间生物相似度的评论文章, 强调并据称证实了进化主张的相似度, 但忽略了由于普遍存在的抛弃数据和选择性分析而导致的严重的生物信息学问题。¹⁵ 伍德的评论几乎没有支持创造论者的主张: 人类是唯一按照上帝的形象被创造的, 而且与黑猩猩的差

异远不止在于少量 DNA 碱基对的不同。因此, 我们对 DNA 序列相似度的研究将包括伍德的评论中列出的论文, 以及最近的几篇论文。本文使用包括被抛弃的数据, 对相似度进行重新计算, 结果见表 1。

在黑猩猩基因组计划开始时, 发表的第一篇人类和黑猩猩 DNA 序列的论文也许是最客观的论文之一。DNA 重组动力学的早期开拓者之一罗伊·布里顿使用非典型的、未公布的基于“公式翻译程式语言”的计算机程序将源自 5 个黑猩猩的大插入 DNA 克隆 (细菌人工染色体, 或称为 BACs16) 的

基因组序列与人类基因组序列进行比较, 选择这五种黑猩猩的 BAC 序列是因为它们是当时唯一可用的。¹⁸ 研究人员按照常规流程选择初始 BAC 种子进行基因组测序, 因为它们含有单拷贝 DNA, 这使得它们更容易结合并与其他物种进行比较。所有 5 种 BAC 的 DNA 序列的总长度为 846,016 个碱基。然而, 其中只有 92% 的碱基可与人类的 DNA 碱基匹配, 因此最终的统计数据仅报告了 779,132 个碱基。值得称道的是, 布里顿的论文中包含了插入和缺失的 DNA 比对数据, 并报告了人和黑猩猩 DNA 相

似度大约为 95%。然而，更真实的结果应该包括所有五个 BAC 的完整高质量序列，这与对比结果应当包括配对中的插入和缺失同样合理；最终 DNA 相似度为 87%（表 1）。参见图 1，用图示阐明了两个 DNA 链之间的碱基序列比对的插入 / 缺失和取代的概念。

在布里顿发表那篇论文的同一年，埃伯斯贝格尔等人发表了另一个重要的研究论文，他对黑猩猩基因组序列进行随机剪切，选择长度为 300 至 600 个碱基的 DNA 片段进行研究¹⁹。这些 DNA 序列使用 BLAT (Blast-Like Alignment) 计算方法与人类基因组早期建立的数据库进行比对。研究人员选择了三分之二的 DNA 序列进行更详细的分析，因其它三分之一的黑猩猩 DNA 序列不能与人类基因组匹配而被抛弃。在论文的第 19 页的实验方法部分描述了如何进一步对预筛选实验产生的数据组进行过滤，以便获得最符合（研究者）期望的比对结果。然后对得到的数据进行各种比较分析，但这些分析实际上是完全没有意义的，因为用于分析的序列数据是经过极严重的选择、屏蔽和过滤的。结果毫无悬念，他们报告人类与黑猩猩 DNA 的高度相似匹配的区域仅有 1.24% 的差异。基于研究人员声称在比对中抛弃的数据量，更切合实际情况的 DNA 序列相似度不超过 65%（表 1）。

在这些初期的人类和黑猩猩 DNA 比较论文发表之后不久，一个令人担忧的趋势迅速发展，就是论文中仅报告最终的比对结果，并且不再提及如何对数据进行过滤、屏蔽和选择的具体细节。让批判性的读者能够对人类和黑猩猩的整体相似度进行（合理地重新）计算的关键数据开始都被全部省略。例如，刘教授等人发表了人类基因组序列与黑猩猩、狒狒和狨猴的

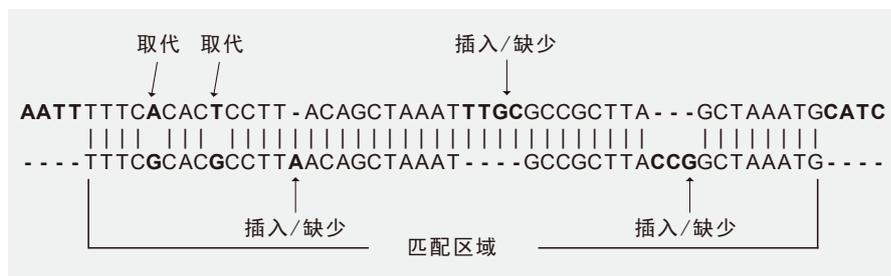


图1: 显示来自两个不同物种的两条同源序列（序列上和序列下）之间假设的比对的说明图示。首先需要说明的是，在几乎所有序列比对中，在匹配的区域之外通常存在不匹配的部分，这些序列会被忽略，但是这些被忽略的序列却反映了两个物种间真实的差异。比对中的“插入和缺失” (indels) 代表这个物种与另一个物种的 DNA 序列相比，存在插入或缺失。插入或缺失的 DNA 序列大小范围从单个碱基到数千个碱基不等。这些插入或缺失在有些序列比对中被计算，在某些比对中会被忽略。“取代”是两个 DNA 序列之间存在不同的碱基，在大多数情况下“取代”会反映在比对结果中。

比对结果²⁰，其中省略了关于序列的起始组和比对具体数据的重要信息。他们只说对上述所有物种总共对比了 10.6 Mb 的 DNA 序列。他们相似度比对的最终估计结果为 98.9%，此结果忽略了插入 / 缺失和非匹配区域的数据。若将插入 / 缺失计算在内，我们得出的相似度估计结果为 95.6%，与布里顿的研究结果相近。由于省略了非匹配区域的重要数据，因此无法对其进行评估。

另一个令人担忧的趋势是，通常只有高度保守的蛋白质编码序列（外显子）才被用于全基因组相似度的评估。我们现在知道，基因组中超过 95% 的序列都是不编码蛋白质的序列，它们对于遗传和基因组各个方面的功能都是至关重要的。⁸ 这个趋势的典型操作就是仅比对外显子的 DNA 序列，威德曼等人发表了一项研究论文，该研究仅对人和黑猩猩共 97 个蛋白编码区的外显子片段进行了比较，合计共 90,000 个碱基。²¹ 选择这样的外显子是基于它们都存在于人类和黑猩猩基因组中，并且已知它们是高度匹配的。由于这些偏见以及论文中缺乏对实验材料和方法细节的描述，在这种情况下不可能算出被抛弃的数据，以及可靠的真实相似度结果（表 1）。

在 2004 年渡边等人使用多种细菌人工染色体 (BAC) 文库来进行黑猩猩代表性的 22# 染色体 DNA 测序克隆。²² 然后将该序列与对应的人类同源序列进行比对。需要注意的是，这个研究只选择含有 6 到 10 个人类 DNA 标记的黑猩猩 BAC 克隆进行比对。我们又一次看到了在相似度比对的开始，预先进行带偏见的筛选。而且在这个研究中，这个带偏见的筛选甚至发生在 DNA 序列对比数据产生之前。遗憾的是，文章或其补充信息中并未提供关键的整体 DNA 比对统计数据。作者指出在匹配区域的核苷酸取代率为 1.44%，但没有提供包含插入 / 缺失的相似度估计结果。虽然在相似度的比对中抛弃了插入 / 缺失的数据，但作者给出它们的数量是 82,000 个碱基，并提供了一个直方图，该图表显示了各类数据的大小分布。奇怪的是，图中没有提供插入 / 缺失的大小或插入 / 缺失总长度的平均数据。同样，只给出了序列间隙的数量，但没有提供关于序列间隙总大小的数据。尽管对相信是测序较准确的种间同源染色体区域进行比对，但是让人能够计算总体的 DNA 相似度的特定数据显然是不存在的。基于图表所提供有限的碱基取代和插入 / 缺失的数据进行总体相似度的评估，

粗略且相当保守的相似度估计结果是 80-85% (表 1)。

在所有人类与黑猩猩相似度研究中,最模糊不清的研究之一是由尼尔森等人发表的论文。²³为了与已形成的误导公众的趋势保持一致,他们只使用了高度保守的外显子进行比对,并且没有给出可以用来计算任何真实整体相似度的数据。在分析开始时,基因序列的总数量(20,361个碱基)中,研究人员决定抛弃其中的 33%(6,630个碱基)的序列,含糊其辞地说是为了“非常保守地进行质量控制”。换句话说,在黑猩猩与人类的基因序列比对的一开始,就有三分之一的序列不能匹配,因此它们就被抛弃了。实际上,报告中没有给出具体的数据可用于评估之后比对的三分之二黑猩猩 DNA。作者仅报告了“沉默位点”之外的序列中取代的差异。这些被抛弃的“沉默位点”序列数据代表遗传信息差异的区域,进化论者称它们是对基因组功能几乎没有影响的区域。由于现在证明基因组的大多数不编码蛋白质的非编码区域是具有活性功能的,因此这个错误预设是需要被重新检讨的。重要的插入/缺失数据也被完全忽略。遗憾的是,在这个非常模糊不清的报告中没有提供足够的数据来进行真实相似度的粗略估算。

黑猩猩初步的基因组数据与人的基因组有 81%相似吗?

在人类与黑猩猩基因组研究中,里程碑式的重要论文是 2005 年国际黑猩猩基因组测序协会发表在《自然》杂志的论文。不幸的是,这论文也是遵循了先前形成的趋势,即大多数比较数据是经过高度选择的,并以模糊不清的方式描述数据,且没有关于比对的详细信息。论文中的大部分内容主要涉及对各种变异率和选择力进行各种假设的进化分析,并且小心避开了整体相似度

的关键问题。

然而,基于黑猩猩基因组论文中提供的数据,我们可以通过包括来自人类基因组计划公布的信息来粗略计算人类和黑猩猩总体基因组之间的相似度。作者指出,关于总体基因组的比对,“黑猩猩基因组和人类基因组的高质量测序的 DNA 序列,在核苷酸水平上大约有 2.4 Gb 的序列是能较好地匹配的”。²⁴此时,人类的常染色体基因组数据库估计完成 99%,核苷酸数量为 2.85 Gb,测序的错误率为 100,000 个核苷酸中有 1 个错误。黑猩猩基因组论文的作者表示,“基因组之间的插入/缺失总共约为 90 Mb,这个差异占两个物种基因组的 3%,而核苷酸取代导致的差异占基因组的 1.23%。”²⁶

总的来说,只有 2.3 Gb 的黑猩猩 DNA 序列与高度准确和完整的人类基因组 DNA 序列(2.85 Gb)匹配,这一比对方法屏蔽了低复杂度 DNA 序列数据。对于能与人类基因组匹配的黑猩猩序列,如计算取代和插入/缺失的相似度结果为 95.8%,这是一个不包含被屏蔽区域的、带偏见的结果。使用这些数字进行保守估计,黑猩猩全基因组与人类全基因组相比,总体相似度为 80.6%。2005 年,已经完成了黑猩猩基因组五倍的冗余测序,得到的 DNA 数据应该占整个基因组序列的 95%以上。

伍德的报告主要为分析和验证 2005 年整个黑猩猩基因组数据库的数据。²⁷伍德使用人类与黑猩猩中已知比较相似的种间同源基因所翻译的氨基酸序列,因此可进行比对。然而,使用已知种间同源基因的计算机翻译的氨基酸序列的比较结果,几乎不能作为全基因组 DNA 序列相似度的准确指示数据。“种间同源基因”是不同物种中的基因,假定它们是从共同的祖先基

因进化而来的,主要是因为它们在两个物种中具有相同的功能和相似的序列。已知种间同源基因的计算机翻译氨基酸序列比较并非全基因组序列相似度的准确指示,首先因为事实上只有少于 5%的人类基因组含有蛋白质编码序列。使用计算机翻译产生的氨基酸序列数据进行比较的第二个问题是,大多数哺乳动物基因会进行选择性转录,选择性的翻译起始/终止位点,另外还有外显子剪接、基因调控 RNA 编码区域、增强子元件和许多其他复杂的编码转录和剪接机制。^{28、29}鉴于我们目前对基因组实际功能的了解,使用计算机翻译的氨基酸序列数据进行基因组间的比对方法需要被进化论者和创造论者认真重新检讨。

人类和黑猩猩相似度的范式开始崩溃

在伍德对人类与黑猩猩相似度神话相关的一些主要论文的总结¹⁵发表之后,出现了几个重要的论文,这些论文质疑人类—灵长类动物进化范式的教条。第一个是埃伯斯贝格尔等人的研究,其中大量的人类,黑猩猩,猩猩,恒河猴和大猩猩基因组序列被用于构建系统发育树(以进化树形式进行的多重比对分析)。³⁰最初的 DNA 序列库实际上经过了几个级别的选择来进行预分析,修剪和过滤,以实现最佳匹配。首先,一组 30,112 条序列的数据被选择,这些序列在五个物种之间具有同源性(重叠相似度)。对这些 DNA 序列进行比对时,仅保留那些匹配度 ≥ 300 个碱基的序列用于另一系列的比对,并且在最终分析时仅使用那些能够产生优异统计结果(相似度 $> 95\%$)的序列。这个预先选择的过滤方法去除了超过 22%已知的同源序列。虽然为了得到最好的匹配度并建立

(最好看的)进化树,已经对所有这些数据都进行过滤,但比对结果并没有清晰显示黑猩猩或任何类人猿是人类的亲缘祖先。结果反而显示人类和灵长类动物 DNA 序列非常独特,没有任何明确的共同祖先的进化路径。也许最好的研究总结正如作者自己所说的:

“根据我们基因组中大约 23% 的序列比对结果,我们与最亲近的黑猩猩之间没有直接的亲缘遗传关系。”

“因此,在三分之二的家谱研究案例中,人类和黑猩猩不是彼此最亲近的遗传亲属。相应的家谱分析结果与物种进化树(所表示的进化关系)并不一致。实验数据表明人类基因组并没有单一进化历史。相反,人类基因组好像是将不同的,按照自己谱系发展的动物拼凑而成的。”³¹

作者补充说,人类和其他灵长类动物之间的基因组相似度不支持一致和清晰的进化树的原因是“比对没有包含清晰的系统发育信号”,³²这是一个重要的陈述,表明他们使用极高水平的数据过滤和选择的目的是为了提供极高水平的“系统发育信号”。

令人震惊的 Y 染色体

最近几年发表的最能打破进化教条的论文之一是人类和黑猩猩之间的 Y 染色体比较。³³在这项研究中对人类和黑猩猩的 Y 染色体的一个大区域——男性特异区域(MSY)进行了比较。为了实现这一目标,必须进行大量的重新测序,因为黑猩猩该区域的 DNA 序列数据是零散且不完整的。最终测序结果得到了在 8 个连续片段



图2: 染色体

中分布的高度准确的黑猩猩 Y 染色体 DNA 序列,合计 25,800,000 个碱基。黑猩猩 Y 染色体与人类 Y 染色体相比,差异是巨大的。作者指出,“大约一半的黑猩猩扩增子序列在人类的 MSY 中没有同源的、对应的可匹配区域,反之亦然。”³⁴扩增子序列包含丰富的重复单元(称为回文单元),它们向前和向后读取都能读到相同信息。一些基因家族分散在这些回文序列中,它们主要在雄性睾丸中表达。在 Y 染色体中不但有 50% 的扩增子序列不能在人类和黑猩猩之间匹配,而且人类的总基因数量超过黑猩猩的两倍(人类为 60 个,黑猩猩为 25 个)。在人类中发现了 3 个完整类型的基因(基因家族)在黑猩猩中甚至不存在。对于这种基因含量的巨大差异,作者指出,“尽管黑猩猩的 MSY 结构精细,其基因库确实比人类的 MSY 小得多,而且简单得多,”³⁵另外“与人类 MSY 含有的基因相比,黑猩猩 MSY 中仅有三分之二的独特基因或基因家族,只有人类的一半的蛋白质编码转录单元。”³⁵

除了这些雄性特有的基因外,

还有一个称为“X 简并基因”或“X 染色体退化基因”的包含基因的特征区域,这是一个带误导性的术语,因为它是源于一个进化论的假设,即 X 简并基因与雌性 X 染色体有同源性,假设它是由 X 染色体演化而来的。在对人和黑猩猩之间的 X 简并基因区域进行比较时,除了基因含量的差异之外,基因的组织结构和位置也明显不同。事实上,人类有三种类型的 X 简并基因,黑猩猩中甚至不存在。

在人类和黑猩猩 MSY 区域除了基因的含量存在巨大差异外,整体的结构差异也是巨大的。请注意作者的一些补充意见:

“而且,保留在两个物种谱系中的 MSY 序列已经被反常地重新排列:黑猩猩和人类 MSY 的整个染色体的点图比较显示出它们的总体结构有显著的差异。”

“黑猩猩扩增区域特别庞大(比人类的大 44%),而且结构上更加丰富,有 19 个回文序

列（人类中只有 8 个），并且在染色体的短臂和长臂之间存在精致的核苷酸镜像序列，这个特征结构在人类的 MSY 中是不存在的。”

“在 19 个黑猩猩的回文序列中，在人体的 MSY 中也发现了 7 个（与其相似的回文序列）；另外 12 个是黑猩猩特有的。与人体 MSY 不同，几乎所有的黑猩猩 MSY 回文序列都存在多个拷贝。”³⁴

在对人和黑猩猩 Y 染色体研究中，发现了其 DNA 特征性的结构排列和基因含量均存在巨大差异，这个发现对人类和黑猩猩 DNA 相似度的神话，以及灵长类动物进化的教条产生了巨大的打击。事实上，作者令人震惊地指出“……600 万年前分开进化的黑猩猩和人类，今天它们之间的 MSY 基因含量的差异程度，接近于 3.1 亿年前分开进化的鸡和人类中的常染色体基因含量的差异程度。”³⁵

主要的问题在于进化论教条无法解释人类和黑猩猩 Y 染色体之间的巨大差异

主要的问题在于进化论教条无法解释人类和黑猩猩 Y 染色体之间的巨大差异。一项关于人类基因组遗传变异的大型研究结果表明，Y 染色体非常稳定，遗传变异比常染色体少 5 倍。³⁶ 这一结果非常合理，由于 Y 染色体在基因组中没有相似的同源染色体，所以它在减数分裂过程中与 X 染色体发生的重组极少。由于 Y 染色体缺乏重组和序列多样性，灵长类动物的进化模型遇到了严重的问题，因为（进化模型预测）人类和黑猩猩的 Y 染色体的相似度应该更高。进化论者认为 DNA 序列的高水平变异

表明基因组中该区域快速进化。而 Y 染色体应该具有这种快速进化的特征，因为人的 Y 染色体与黑猩猩的明显不同，但实际上它并没有快速的变异。相反，它事实上是非常稳定的，全世界男人的 Y 染色体的结构差异很小，序列多样性很少。与人类基因组的其他染色体相比，Y 染色体的稳定性已经得到证实，再结合人类和黑猩猩 Y 染色体之间的巨大差异，这问题对于人类与黑猩猩拥有共同祖先的范式来说，是一个无法解释的谜团。

一些高度相似的结果可能是由于污染

在人和黑猩猩相似度讨论中还要考虑的另一个因素就是一些序列高相似度的结果可能是由于污染造成的。不仅因为黑猩猩的基因组的构建仍主要以人类的基因组框架为基础，另外已知，非灵长类动物基因数据库大范围受到人类 DNA 的污染，成为一个严重的问题，在某些情况下可能高达 10%。³⁷ 这是由于实验室 DNA 克隆片段的测序过程中，污染了人类的 DNA 导致的，例如实验室空气中含有来自咳嗽，打喷嚏以及受污染手指的接触带来的人体细胞，会产生人类 DNA 的污染。而灵长类动物基因数据库中的人类 DNA 污染的检测和确定可能是一项困难而且非常主观的工作，因为灵长类动物的进化理论是高于一切的教条。还有一点值得注意，黑猩猩基因组的测序工作是在基因组数据库大范围受到人类 DNA 污染的问题未被充分暴露的时期进行的。使用人类基因组框架进行黑猩猩基因组序列拼接和标注也加重了污染问题的复杂性。

实际上，污染不仅可能由实验室错误产生，而且在黑猩猩基因组拼接和基于进化论教条标注基因组期间被故意引入的。在最近的 Ensembl 数据库网站（EMBL-EBI

和 Wellcome Trust Sanger 研究所之间的生物信息学合作项目）上，一个名为“Chimp Genebuild”的网页提供了以下一种使用人类基因组作为参照，来拼接和标注黑猩猩基因组的指南：

“由于蛋白质数量较少（其中许多分布在与人基因组相同的位置），因此可将人类基因的位置点投射到黑猩猩的基因组上。人类基因组高质量的标注，以及人类与黑猩猩基因组之间的高相似度使我们能够通过将人类基因投射到黑猩猩基因组中的相应位置来识别黑猩猩的基因。”

“人类蛋白质编码基因的结构通过 WGA [全基因组拼接] 投射到黑猩猩基因组的染色体上。通过在结构中插入‘移位框’内含子来纠正那些干扰转录阅读框的小插入/缺失。”³⁸

不仅黑猩猩的基因组使用人类基因组作为构建的框架，在黑猩猩的基因组中的人类序列污染也被承认是存在的，因为人类基因组序列的计算机数据被填充在假定缺失的黑猩猩序列中。基于进化论的神话和教条，人类的 DNA 被认为几乎与黑猩猩的相同，人类 DNA 的碱基和片段已被整合到黑猩猩基因组的空隙和各个区域中，使它看起来更像人类。因此，当研究人员下载黑猩猩基因组序列的数据库和标注用于进行独立研究时，没有像正常所期望的那样得到 100% 无偏见的黑猩猩基因组序列。相反，得到的是基于人类基因组进行构建、对齐和调整后的，人类 DNA 和黑猩猩 DNA 序列的拼凑体。

结论：人类和黑猩猩的DNA并不那么相似

黑猩猩基因组数据库的最终标注和构建结果显然是一种偏见的产物。此外，几乎所有人类与黑猩猩DNA相似度相关的研究论文都省略了大量不匹配的，以及代表序列间隙的数据。事实上，大量的论文甚至没有包含足够的信息，让独立的读者能够在最终的经过高度过滤的相似性结果给出之前，计算数据被最初筛选前存在多少的不相似性。从已公布的论文中提供的信息（但经常被掩盖），可以肯定地说，估计人类与黑猩猩基因组的相似性不超过81-87%，甚至可能更低。

为了支持这一结论，一项大规模的人类与黑猩猩基因组比对研究项目最近刚刚在一份独立的期刊上发表。³⁹ 这项研究的结果完全支持并证实了本报告中提供的信息。在这项研究中，作者汤姆金斯使用了30种不同参数组合下进行blastn比对，得到了40,000个随机黑猩猩基因组序列与4种不同版本的人类基因组的两两比对数据。这项研究工作总共进行了120万次尝试性比对——如果以四种不同的人类基因组版本的比对来计算就是480万次。不包括大量无法匹配的黑猩猩序列数据，汤姆金斯报告了能匹配区域中的人类与黑猩猩DNA相似度的非常保守估计的结果为86-89%（使用不同的算法和参数结果有所不同）。通过这项广泛且非常客观的研究，结果明确地表明人类和黑猩猩的基因组的相似性与通常宣称的相似性相比，至少低10-12%。人类与黑猩猩拥有共同祖先的进化范式宣称：“人类与黑猩猩的DNA序列几乎完全相同”，显然更多是基于神话和进化信仰的宣传，而不是基于真实的事实数据。 ■

参考文献和注释

1. Dawkins, R., *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe Without Design*, W.W. Norton, New York, p. 263, 1986.
2. 国际人类基因组测序协会，人类基因组的初步测序和分析，《自然》杂志，409: 861-920, 2001。
3. Venter, J.C. et al., The sequence of the human genome, *Science* 291:1304-1351, 2001.
4. 黑猩猩测序与分析协会，黑猩猩基因组初始序列与人类基因组比较，《自然》杂志 437: 69-87, 2005。
5. 在近期的一份基因组测序综述中解释了各种用于人类和黑猩猩基因组测序的各种技术，参看：Tomkins, J., How Genomes are sequenced and why it matters: implications for studies in comparative genomics of humans and chimpanzees, *Answers Research J.* 4:81-88, 2011, www.answersingenesis.org/articles/arj/v4/n1/implications-for-comparative-genomics.
6. Bergman, J., The functions of introns: from junk DNA to designed DNA, *Perspectives on Science and Christian Faith* 53(3):170-178, 2001.
7. Woodmorappe, J., *Junk DNA indicted, J. Creation (formerly Creation Ex Nihilo Tech. J.)* 18(1):24-33, 2004.
8. Wells, J., *The Myth of Junk DNA*, Discovery Institute Press, Seattle, WA, 2011.
9. Cohen, J., Relative differences: the myth of 1%, *Science* 316:1836, 2007.
10. Hoyer B.H. et al., Examination of hominid evolution by DNA sequence homology, *J. Human Evol.* 1:645-649, 1972.
11. Sibley, C.G. and Ahlquist, J.E., The phylogeny of the hominoid primates, as indicated by DNA-DNA hybridization, *J. Mol. Evol.* 20:2-15, 1984.
12. Sibley, C.G., DNA hybridization evidence of hominoid phylogeny: a reanalysis of the data, *J. Molec. Evol.* 30:202-236, 1990.
13. Gibbons, A., Which of our genes make us human? *Science* 281:1432-1434, 1998.
14. 参见 Bergman, J. and Tomkins, J., Is the human genome nearly identical to chimpanzee?—a reassessment of the literature, *J. Creation* 26(1):54-60, 2012.
15. Wood, T.C., The chimpanzee genome and the problem of biological similarity, *Occasional Papers of the BSG* 7:1-18, 2006.
16. 要理解更多BACs的信息，参见Tomkins, ref. 5, p.1.
17. Britten, R.J., Divergence between samples of chimpanzee and human DNA sequences is 5% counting indels, *Proc. Nat. Acad. Sci.* 99:13633-13635, 2002.
18. Britten, personal communication, 15 July 2011.
19. Ebersberger, I. et al., Genomewide comparison of DNA sequences between humans and chimpanzees, *American J. Human Genetics* 70:1490-1497, 2002.
20. Liu, G. et al., Analysis of primate genomic variation reveals a repeat-driven expansion of the human genome, *Genome Res.* 13:358-368, 2003.
21. Wildman, D.E. et al., Implications of natural selection in shaping 99.4% nonsynonymous DNA identity between humans and chimpanzees: enlarging genus *Homo*, *Proc. Nat. Acad. Sci.* 100:7181-7188, 2003.
22. Watanabe, A.F. et al., DNA sequence and comparative analysis of chimpanzee chromosome 22, *Nature* 429:382-388, 2004.
23. Nielson R. et al., A scan for positively selected genes in the genomes of humans and chimpanzees, *PLoS Biology* 3(6):e170. doi:10.1371/journal.pbio.0030170, 2005.
24. 见参考文献4中的第71页。
25. 国际人类基因组测序协会，完成人类基因组的全基因组测序，《自然》杂志，431:931-945, 2004。
26. 见参考文献4中的第73页。
27. Wood, ref. 15, p. 3.
28. ENCODE 项目协会，通过ENCODE试点项目识别和分析人类基因组中1%的功能元素，《自然》杂志，447:799-816, 2007。
29. Barash Y. et al., Deciphering the splicing code, *Nature* 465:53-59, 2010.
30. Ebersberger, I. et al., Mapping human genetic ancestry, *Molec. Biol. Evol.* 24:2266-2276, 2007.
31. Ebersberger et al., ref. 30, p. 2266.
32. Ebersberger et al., ref. 30, p. 2269.
33. Hughes, J.F. et al., Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content, *Nature* 463:536-539, 2010.
34. Hughes, J.F. et al., ref. 33, p. 537.
35. Hughes, J.F. et al., ref. 33, p. 538.
36. 国际SNP图谱工作组，包含142万个单核苷酸多态性的人类基因组序列变异图谱，《自然》杂志，409: 928-933, 2001。
37. Longo, M.S. et al., Abundant human DNA contamination identified in non-primate genome databases, *PLoS ONE* 6(2): e16410, 2011.
38. 截至2012年1月13日，参见：ensembl.fugus.org/Pan_troglydotes/chimp_build.html。如果这个URL或网页失效，可以联系作者 Tomkins(jtomkins@icr.org) 以获得网页文章的PDF文档。如前所述，Ensembl代码可供公众下载和使用。
39. Tomkins, J.P., Genome-Wide DNA Alignment Similarity (Identity) for 40,000 Chimpanzee DNA Sequences Queried against the Human Genome is 86-89%, *Answers Res. J.* 4:233-241.

杰弗里 P. 汤姆金斯 (Jeffrey Tomkins)

拥有克莱姆森大学遗传学博士学位、莫斯科爱达荷大学植物科学硕士学位，以及华盛顿州立大学农业教育学士学位。他曾在克莱姆森大学遗传与生物化学系任教10年，在同行评审的科学期刊上发表了58篇世俗的研究论文，在科学书籍中发表了7章内容——涉及遗传学、基因组学和蛋白质组学。在过去的三年里，Tomkins博士一直是ICR的一名研究科学家，他在ICR杂志上发表了9篇关于创造科学的论文，14篇半技术性的文章，并有两本书。

杰瑞·伯格曼 (Jerry Bergman)

拥有9个学位，其中包括2个博士学位。主要研究领域是生物、化学和心理学。毕业于底特律的韦恩州立大学、托莱多的俄亥俄医科大学、托莱多大学和鲍林格林州立大学。伯格曼博士是一位多产的作家，发表过许多著作，他在俄亥俄州阿奇博德的西北州立大学教授生物学、化学和生物化学超过24年。他也是托莱多大学医学院的副教授。



我们[倒置]的 视网膜

真的是[糟糕的设计]吗？

文/ 彼得·W.V·格尼

对于脊椎动物，光线到达眼睛的感光细胞之前，必须先经过几层视网膜神经层，这种“倒置的”视网膜结构，长期以来一直是进化论者的嘲笑对象，他们声称这种结构效率低下，因而是反驳上帝设计的证据。本文研究了我们的视网膜倒置的原因，并讨论视网膜正置（即感光细胞位于最内层，也就是接受入射光线的第一层）可能会让这些动物的视觉效果更差。我认为，在脊椎动物中，视网膜需要被保护免受光的损伤，特别是波长较短的光，并免受会聚光线产生热量的伤害，这就需要生物拥有“倒置”的视网膜结构。

前言

进化论者经常认为，脊椎动物眼睛视网膜的特征表明它不是被设计而来的，因为它的结构似乎不合理。他们指出，光在到达感光细胞之前必须穿过视网膜大部分神经纤维层，并假定在感光细胞水平上形成的图像会因此减褪或扭曲。这种视网膜结构的生物学术语是“倒置”了，因为感光细胞的感受端远离入射光（图1）。这种“倒置的”视网膜在脊椎动物中是常见的，但在无脊椎动物中比较罕见，在一些软体动物和蛛形纲动物中才有发现¹。

在无脊椎动物中，更常见的是“正置”的视网膜结构，即是感光细胞正面接受光的信号（称为垂直结构）。许多进化论者认为这种视网膜结构更有效，嘲笑脊椎动物的视网膜结构“安装错误”。然而，著名的无神论进化论者道金斯，承认在光穿过倒置的视网膜进入感光细胞的过程中没有受到明显的干扰，他写道：

任何一个生物学家都会自然而然地认为感光细胞应该向着光线，它们的传入神经应向后延伸到大脑，若感光细胞向着

远离光的方向，它们的传入神经向着离光最近一侧的结构会遭到嘲笑。然而，这正是所有脊椎动物视网膜的情况。实际上，每个感光细胞都是倒装的，其传输神经伸向光的一侧，神经纤维汇聚后再从视网膜表面一个称为盲点的孔穿过，然后连接视神经。这意味着光不是直接到达感光细胞，而必须先穿过一个传输神经网络，在此过程中，应该至少出现一些损耗和扭

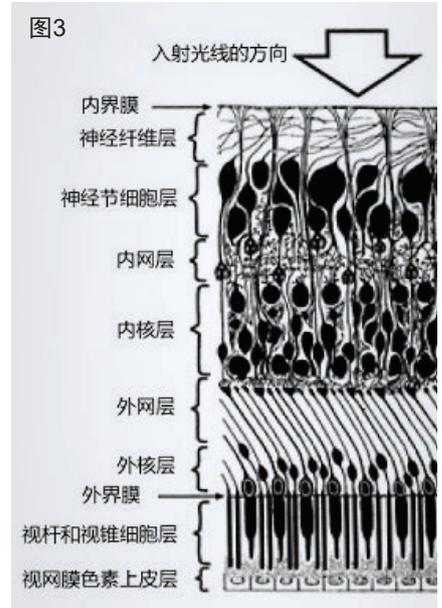
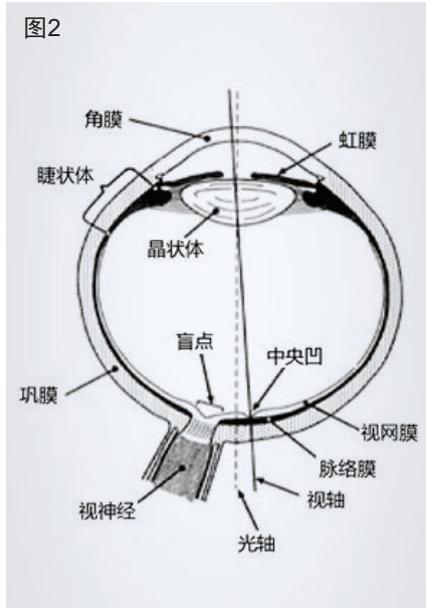
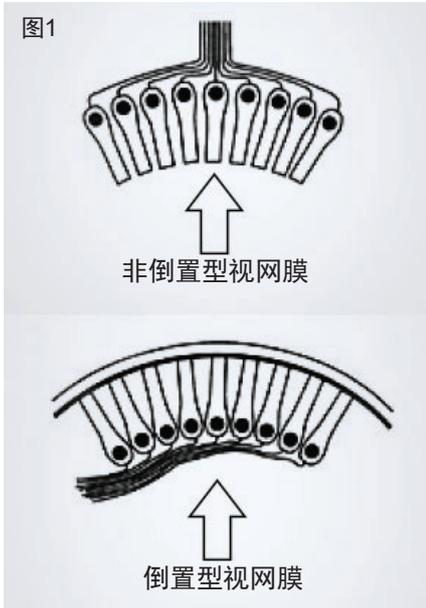


图1: 眼睛中感光细胞的两种排列方式。图中箭头指示入射光线的方向。图2: 眼睛截面的示意图。图3: 视网膜层次结构的示意图。

曲（实际上，可能不太多，但仍旧成为可能冒犯头脑精明的生物学家的重大问题）。我不知道如何确切地解释这种奇怪的状况，因为相关的进化时期是很久以前的事了。²

在讨论进化论的观点之前，我们先简要回顾一些眼睛解剖结构和相关术语（图2）。

光通过眼睛的前窗户——透明的角膜进入眼内，角膜相当于一个强力的凸透镜。光通过瞳孔（虹膜形成的光圈）后，会被晶状体进一步折射。由此，外界物体的像聚焦在视网膜上，视网膜是眼后部三层结构的最内层（最靠近眼球几何中心的那层），能将光信号转化为神经信号。眼后部的另外两层膜中，最外层的是白色坚韧的纤维巩膜，其向前连接于角膜。脉络膜是夹在视网膜和巩膜中间的一层，有色素沉着，富含血管。

视网膜由10层组成（图3），最外层（靠近脉络膜）是黑色的视

网膜色素上皮层（RPE层），含有黑色素，所以不透光。RPE细胞内表面有细小的毛发状突起，形成微绒毛，位于感光细胞外段的顶端之间，并覆盖顶端。因此，视网膜色素上皮层和感光细胞之间存在一个潜在的裂隙，当视网膜神经感觉层和色素上皮层分离，即视网膜脱落时会出现（损伤可引起这种情况）。

感光细胞包括视杆细胞和视锥细胞，每一个感光细胞都是由一个内节和一个外节组成的，内节具有制造视色素的细胞器（细胞内机器），视色素存在于外节中。视杆和视锥层及向内的其它八层视网膜结构层（不同于RPE层）组成了几乎透明的视网膜神经感觉层，光传到感光细胞，产生电脉冲，电脉冲经处理后，由视网膜神经感觉层繁复的神经连接传输到视网膜神经纤维层，然后再经过视神经传到大脑。

对于许多需要在光照水平非常低的情况下运用视力的动物，在RPE或脉络膜中有一层叫“绒毡

层”的反射晶体材料十分重要。³当光线在感光细胞之间通过后，绒毡层像一面镜子一样反射光线，从而增强了感光细胞接受的光线。因此，眼睛在黑暗中能看到微弱的光线，俗称“猫眼”。

视网膜色素上皮层

了解RPE的关键作用是认识视网膜倒置结构的基础。RPE的许多重要功能现在已经广为人知^{4,5}。每一个RPE细胞都与20个或更多的感光细胞外节的顶端紧密接触。感光细胞外节数量超过1.3亿，如果没有RPE，感光细胞和视网膜神经感觉层的其余部分就不能正常工作，并会最终萎缩。⁶因此，如果视网膜神经感觉层与RPE分离，受影响区域的视网膜视力将逐渐恶化，并可能最终失明。

感光细胞外节由一堆含有感光色素的膜盘组成。这些膜盘不断地由感光细胞内节产生，并不断向外移动到外节朝向RPE的部位，被RPE中的吞噬细胞吞噬并回收其化学成分（图4）。⁷

RPE 储存视色素的前体维生素 A，并参与其再生。有四种视色素在光照下都被脱色：视紫红质（存在于视杆细胞中，用于夜间视力），其他三种视色素分别存在于三种不同类型的视锥细胞，分别对应三原色（红绿蓝）中的一色。它为感光细胞间的基质合成粘多糖，该基质即位于感光细胞之间并分隔感光细胞的物质。

除氧气外，RPE 还选择性地从脉络膜输送营养物质，以供应视网膜的外部的三分之一，从感光细胞运走代谢产物至脉络膜循环并将其清除。通过选择性地泵送代谢产物以及细胞间存在的紧密连接，RPE 充当一个屏障，称为血视网膜屏障，可以阻止大的或有害的化学物质进入视网膜组织，从而有助于维持视网膜稳定并维持最佳的视网膜环境。^{8、9、10}

RPE 有复杂的机制来处理由光作用产生的有毒分子和自由基。超氧化物歧化酶、过氧化氢酶和过氧化物酶等特异酶可催化潜在有害分子（如超氧化物和过氧化氢等）的分解。抗氧化剂如生育酚（维生素 E）和抗坏血酸（维生素 C）可用于减少氧化损伤。

因此我们的感光细胞从已被 RPE 消化的旧膜盘中回收物质，不断合成新的外节膜盘和特定的视色素。这就产生了一个问题：“为什么会有如此复杂的过程？”因为这是一个生物更新的过程，通过这种方式，暴露在有害的化学物质、辐射、机械创伤等环境中的组织才能够保存下来。如果没有自我更新，皮肤、肠道内壁、血细胞等组织会迅速积累致命缺陷。同样，通过不断更换膜盘，感光细胞可以对抗由有毒物质，特别是短波光产生的破坏性分解作用。^{11、12}

RPE 细胞含有可以吸收分散的和过量的光线的黑色素颗粒，从

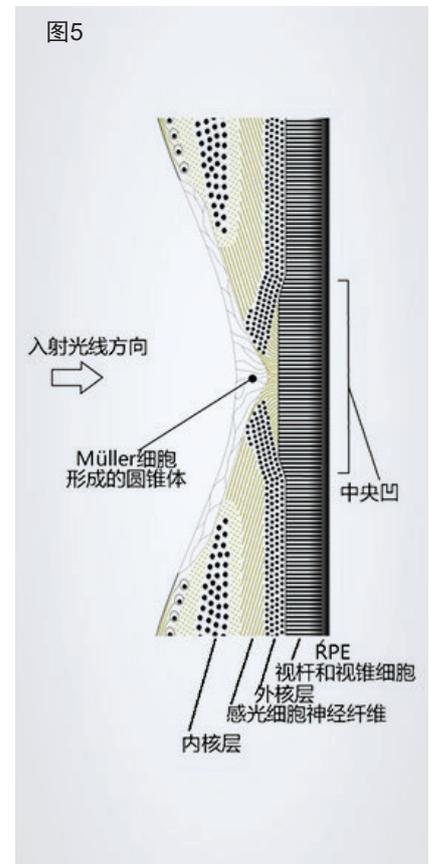
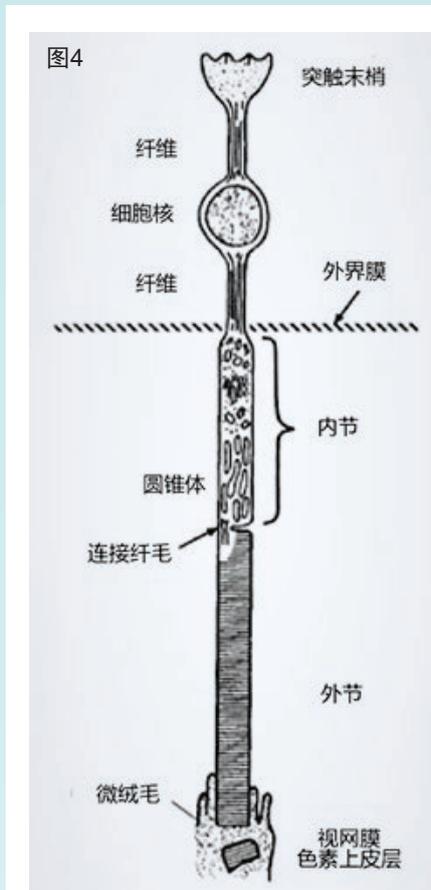


图4: 视杆细胞示意图。图5: 人类眼睛中央凹的截面图。

而提高视力（指视觉敏锐度或分辨两个相近点的能力）。25%-33% 进入眼内的光线被 RPE 和脉络膜中的色素颗粒吸收。¹³ 光波波长越短能量越大，而黑色素对光谱的吸收随波长的减小而增大。¹⁴ 因此，它也能保护感光细胞免受光损伤，可起到抑制有害的光敏化分子（包括单线态氧¹⁵）和自由基的作用。¹¹ 此外，RPE 和脉络膜的黑色素还能屏蔽通过巩膜进入眼睛的光，避免感光细胞受到干扰。

RPE 中如此密集的代谢活动需要充足的血液供应，与之接触的脉络膜则起到了这一作用。

脉络膜散热器

已经观察到，在实验模型中感光细胞的损伤与温度密切相关¹⁶，其他研究也证实热量加剧了光化学

损伤。因此，任何要防止光化学损伤的机制，都要防止温度过高。¹⁷

1980 年的一篇文章首次对脉络膜的一些已知情况做出了解释。¹⁸ 例如，尽管视网膜的代谢活性很强，但其极高的血流速度远远超过了视网膜的营养需求。文中提到了早期的实验，这些实验表明，与活的动物相比，少量的光就会导致已死动物的视网膜烧伤。

论文接着描述了在动物身上的实验，实验证明减少脉络膜血流会使视网膜更容易受到光产热的损伤。脉络膜血流量占眼血流的 85%，在体内所有组织中，脉络膜每克组织的血流量竟然是最高的，是肾皮质的 4 倍。作者还注意到，在流经脉络膜的血液中，只有很少的氧气被使用。

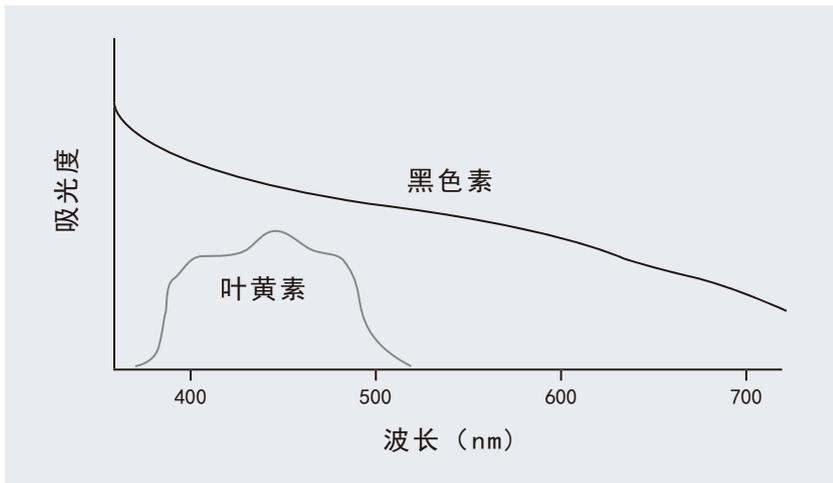


图6: 色素的吸收光谱

脉络膜毛细血管形成一个丰富的血管丛，位于 RPE 的正外部，主要是其中心区域，并仅由一层很薄的膜（玻璃膜 / 布鲁赫氏膜）与之隔离。RPE 吸收多余的光会在外层视网膜产生热量，为避免对其自身和邻近的精密复杂的生物组织造成热损伤，必须进行散热。

这项研究的作者有说服力地指出，脉络膜具有快速的血流，并接近 RPE，它的一个重要功能是充当散热器和冷却装置。更令人惊叹的是，同一批工作人员的进一步研究表明，有些中枢神经（通过大脑）在光照的条件下能提高脉络膜的血流，这些中枢是按照光强度调节脉络膜血流的神经反射。^{19、20}

因此，很明显，RPE 和脉络膜对于保护人类视网膜的功能是很重要的。但是 RPE 由于其黑色素、脉络膜由于其血液和黑色素，这两种结构都是不透明的。因此，为了让光能到达感光细胞，RPE 和脉络膜都必须位于视网膜神经感觉层的外部，因此我们可以得出结论，人和脊椎动物视网膜的倒置结构有合理的原因。^{21、22} 接下来，我们讨论与视网膜倒置结构相关的其他两个设计特征。

黄斑中央凹

除极细血管中的血液之外，视网膜神经感觉层几乎是透明的，但在其中心区域（称为黄斑）还有一种精细结构。视网膜向枕部大脑皮层（称为视觉皮层）²³ 传输视觉信息，由二者的组织结构决定视觉在视轴上达到最佳效果（视轴即视线见图 2）。视轴直穿黄斑中心带有倾斜壁的圆形凹坑——中央凹，（图 2 和图 5）。离中央凹越远，越靠视网膜边缘，视力越差。因此，红色、绿色，可能还有蓝色²⁴ 的色觉感光细胞——视锥细胞在直径仅为 300-330 微米²⁶ 的黄斑中央凹的密度最大，为每平方毫米 150000 个。²⁵

此外，中央凹的视锥细胞与其他地方的不同，它更高、更细、完全笔直，并且相对于入射光而言，精确地轴向排列，以获得最佳的视力。这个区域无血管，视网膜更薄，只有感光细胞（视锥细胞）和最小的支持组织。视网膜神经感觉层内部的神经元件从中央凹向外放射状排列，以使光线畅通无阻地进入（黄斑的位置），并消除其他地方出现的少量光散射（图 5）。中央视网膜主要作用是感知颜色和形

状，而周边视网膜主要感知光、运动以及夜视。与这些功能一致的是，中央凹处的感光细胞与神经节细胞的比例很低（ $\geq 1:1.2$ ）²⁷，在这里光学像差也很小。视网膜的其他部位的比率与此相反，为每 4-6 个视锥细胞或高达 100 个视杆细胞对应 1 个神经节细胞。^{28、29}

叶黄素

人眼的光学系统构造让落在视网膜黄斑的光线峰值最强，而视网膜边缘光的峰值强度则要小得多。因此，重要的是，因黄斑区 RPE 细胞比其他地方更高，单位面积的数量也更多，黑色素更丰富，不仅如此，视网膜的中心区域还有黄色色素：叶黄素。³⁰ 在视网膜的中心区，叶黄素在其两个限制膜之间（视网膜内界膜和外界膜之间）渗透到神经感觉视网膜的所有层，并集中在视网膜细胞中，包括神经元和支持组织细胞。³¹ 最近，人们注意到在中央凹的内表面上有一组视网膜支持组织细胞（以首次描述它们的人命名为缪勒细胞（Müller cells）），它们形成一个圆锥体，锥尖扣在中央凹处（图 5）。³² 除了为中央凹提供结构支持外，缪勒细胞还被认为充当叶黄素的贮存器，特别是在中央凹的缪勒细胞。

视网膜叶黄素是一种类胡萝卜素，在化学上与维生素 A 有关，其吸收光谱范围从 480nm 到 390nm，峰值约为 460nm（图 6）。^{33、34、35、13、36、37} 它通过吸收许多有潜在破坏性的较短波长可见光（即蓝色和紫色）来保护视网膜神经感觉层，而蓝色和紫色的光更容易被小的分子和结构散射。³⁸ 研究表明，视网膜对光损伤的敏感度随波长的减小呈指数级增加，对紫外线（uvr）敏感度是对蓝光敏感度的 6 倍。³⁹ 然而，几乎所有波长小于 400 纳米的非电离辐射都被角膜

和晶状体联合阻挡，⁴⁰ 留下光谱中蓝色部分波长为 420-450 纳米的危险波段，而叶黄素是剩余波段的有效屏蔽层。⁴¹

这种色素的作用在实践中得到了很好的证明，当眼科医生最初使用氩激光时，发出的蓝光没有被减去，叶黄素因吸收这些蓝光并在视网膜神经感觉层上产生有害的烧伤。因此现代的眼科氩激光被制作成只发射绿光。³⁵

盲点

由于视网膜的倒置结构，将数据传输到大脑的神经纤维——轴突穿行于视网膜内表面下方并汇聚到一个小区域，即视乳头，在那里作为视神经一起离开眼睛。视乳头没有感光细胞，因此无感光能力，从而在视野中产生一个小的盲点。这也是进化论者批评的一个对象，进化论者认为这明显使生物处于劣势。正如威廉姆斯所说：

“我们的视网膜盲点很少会造成任何困难，但‘很少会’与‘从不会’不一样。当我用手遮盖一只眼睛以抵挡一只昆虫时，一个重要事件的影像可能正好投射在另一只眼睛的盲点上。”⁴²

尽管如此，这一问题还是必须对比来看待：盲点的中心位于距视轴 15° 的位置（距中央凹 3.7 mm），相对于眼睛的视野来说非常小，占不到 0.25%。⁴³ 如前所述，视网膜上离中央凹越远的点，其视力和灵敏度就越低。视乳头周围的视网膜，在光适应状态下，其视力仅为中央凹的 15%。⁴⁴ 我们可以保守推断，威廉斯所提到的单眼人盲点产生的理论风险可以忽略不计，因此，单眼人驾驶私人汽车（即非职业目的）被认为是安全的。

由于两眼的视野有很大程度的重叠，所以一只眼睛的生理盲点被另一只眼睛的视野所补偿。确实，当一只眼睛被遮挡或缺失是一种障碍，但这种障碍并不是主要由于上述视力盲点所导致的。在视力起重要作用的情况下，相比之下，更会威胁生存的是失去视觉的立体感（双眼视或深度知觉）及周边视野的减少。⁴⁵

然而，假设创造者为了避免威廉姆斯所描述的情况，而制造了一个非常大的单眼，即两个相同的光学系统能够会聚视轴，并且两个相距足够远的中央凹，以达到我们所满意的立体感水平，显然这一尝试是不切实际的。事实上，如果此时一只眼睛受伤或被遮住，这种假想的生物会更脆弱。

人与动物视觉

人眼的视力虽然很好，但不如某些动物（如鸟类）的视力高，它们的视网膜也是倒置的。但不同的视力能力与动物的生活方式所需相匹配。

人类的视觉系统不能像苍蝇那样准确、灵敏地跟踪运动，但如果人有那样的视觉系统，我们就会看到所有的荧光灯和电视都在不停地闪烁。我们不能像猫那样在黑暗中

看到东西，但在其他某些方面我们超过了猫，例如，猫没有色觉。人眼的功能在数量和质量之间表现了出色的平衡，这使得人类在历史上取得了惊人的技术成就。

近代，人类制造望远镜观察遥远的地方，并制造显微镜观察微观的世界，以及在最黑暗的夜晚能看得见东西的设备，扩大了我们的视野，使人类的能力超过任何其他生物的能力。

无脊椎动物的垂直观网膜

一些进化论者声称，头足类动物（如鱿鱼和章鱼）的垂直观网膜比脊椎动物眼睛的倒置视网膜更有效。⁴⁶ 但这个观点能成立的前提是倒置的视网膜首先被证明是低效的。如上所述，进化论者未能证明倒置的视网膜是一个糟糕的设计，也不能证明它的功能很差；他们忽略了很多优秀的设计因由。

而且，他们从未证明头足类动物的视力确实更好。相反，它们的眼睛的功效仅仅“接近一些低等脊椎动物的眼睛”，⁴⁷ 而且它们可能是色盲。⁴⁸ 此外，头足类动物视网膜除了“没倒置”外，实际上比脊椎动物的“倒置”视网膜简单得多，正如布德尔曼所说，“头足类动物视网膜的结构比脊椎动物的眼



图7：苍蝇的视觉系统准确而灵敏

睛简单得多，只有两种神经成分，即感光细胞和传出纤维”。⁴⁹它是由“带微绒毛的感杆长柱状感光细胞组成的波状结构”，⁵⁰因此头足类动物的眼被描述为“单透镜复合眼”。⁵¹感杆起着导光的作用，它们的微绒毛的排列使动物能够检测到偏振光的方向（以此克服本来会因动物基于反射的伪装机制造成的干扰）。

最后，与大多数脊椎动物相比，头足类动物生活的自然环境中，光强度要低得多，而且它们通常最多只能活两三年。人们对巨型乌贼的寿命一无所知，但无论如何，人们都认为它经常出现在光线很少的深海。⁵²因此，对于头足类动物来说，保护其免受光照损伤的需要较少。不同的眼睛设计适用于不同的环境，因此可以很好地与“没倒置”的视网膜配合使用。⁵³

总结与结论

乍一看视网膜的倒置设计有缺点且效率低下，但事实上，这些反对意见毫无分量。甚至进化论者也承认，倒置的视网膜很好地服务于那些拥有它的生物⁴¹，给它们提供极好的视力。我们已经回顾了这种结构的必要性，并揭开了感光细胞的秘密。

总结

不同波长的光对生物体可能产生非常有害的影响。视网膜，不仅是一个非常复杂的传感器和图像处理器，也明显地是为了可以承受光的伤害和热效应而设计的。眼睛有保护视网膜免受日常生活中常见的辐射伤害的精妙设计。除了角膜和晶状体几乎完全排除紫外线辐射之外，视网膜本身还被赋予了許多额外的机制防止这种损害：

- 视网膜色素上皮产生的物质可以对抗光辐射产生的有害化学物质。
- 视网膜色素上皮是维持感光细胞的重要组成部分，通过回收和代谢它们的产物，使它们在受到光刺激的不断磨损时得以更新。
- 中央视网膜充满了过滤并吸收短波长可见光的叶黄素。

因此，感光细胞需要与不透明的视网膜色素上皮密切接触。反过来，视网膜色素上皮需要与同样不透明的脉络膜密切接触，以满足其营养需求，并通过脉络膜大血流量产生的散热效应，防止视网膜因为会聚光而过热。

如果人类的视网膜是以另一种

相反方向的结构排列（正置，即“未倒置”），如进化论者道金斯提出的，²这两个不透明层将不得不挡在光到达视网膜感光细胞的路径中，这将使感光细胞无法接触光！

因此，我认为，保护视网膜免受光损伤是视网膜倒置的主要原因，而在我们的自然环境中，正置设计的视网膜无法提供相同程度的保护机制。

更新

本文的更新工作只是为了证明其论点的正确性。例如，其他研究人员从分析婴儿斑马鱼的眼睛中引用了一句话：“有证据表明倒置的视网膜实际上是一种更好的节省空间的解决方案，特别是在小眼睛中。”⁵⁴而在感光细胞前面的缪勒胶质细胞起到了纤维光学板的作用，吸收更多的光线，提高图像清晰度，并减少来自杂散光的扭曲，因此是“最佳结构”。^{55, 56}这种结构也有助于分离颜色，使绿光特别是红光指向对这些波长更敏感的视锥细胞，而高能蓝光则指向对它们反应更强烈的视杆细胞，从而优化我们的色觉。⁵⁷

致谢

我感谢乔纳森·萨法蒂博士在我准备这篇文章时的建议，感谢他提供的物理化学信息，以及关于头足类动物和其他动物眼睛的信息。

参考文献和注释

1. DukeElder, S., System of Ophthalmology, Henry Kimpton, London, vol. 1, p. 147, 1958.
2. Dawkins, R., The Blind Watchmaker: Why the evidence of evolution reveals a universe without design. W.W. Norton and Company, New York, p. 93, 1986.
3. DukeElder, Ref. 1, pp. 608–609.
4. Hogan, M.J., Alvarado, J.A., Weddell, J.E., The Retina. In Histology of the Human Eye, pp.393–522, 1971. W.B. Saunders, Philadelphia. As cited in Tasman W., Jaeger E.A. (eds), Foundations of Clinical



图8: 章鱼的眼睛

- Ophthalmology, Lippincott-Raven, New York, vol. 1, ch. 21, 1998.
5. Zinn, K.M., BenjaminHenkind, J., Anatomy of the human retinal pigment epithelium, 1979. In Zinn, K.M., Marmot, M.F. (eds.), The Retinal Pigment Epithelium, Harvard University Press, Cambridge, MA, pp. 3-31. As cited in Tasman W., Jaeger E.A. (eds.), Ref. 4.
 6. 感光细胞有两种类型: 视杆感光细胞的数量大约有 1.25 亿, 而视锥感光细胞的数量约为 650 万。
 7. LaVail, M.M., Outer segment disc shedding and phagocytosis in the outer retina, Trans. Ophthalmol. Soc. UK 103:397, 1983.
 8. Steinberg, R.H., Research update: report from a workshop on cell biology of retinal detachment, Exp. Eye Res. 43:696-706, 1986.
 9. Törnquist, P., Alm, A., Bill, A., Permeability of ocular vessels and transport across the bloodretinal barrier, Eye 4: 303-309, 1990.
 10. Grierson, I., Hiscott, P., Hogg, P., Robey, H., Mazure, A., Larkin, G., Development, repair and regeneration of the retinal pigment epithelium, Eye 8: 255-262, 1994.
 11. Kennon Guerry, R., Ham, W.T., Mueller, H.A. Light toxicity in the posterior segment, 1998. In Tasman W., Jaeger E.A. (eds.), Clinical Ophthalmology, LippincottRaven, New York, vol. 3, ch. 37.
 12. Young, R.W., The Bowman Lecture: Biological renewal: Applications to the eye, Trans. Ophthalmol. Soc. UK 102:42-67, 1982.
 13. Geeraets, W.J., Williams, R.C., Chan, G., Ham, W.T., Guerry, D., Schmidt, F.H., The loss of light energy in retina and choroid, Arch. Ophthalmol. 64:158, 1960. As cited by Parver, L.M. et al., Ref. 18.
 14. 光子的能量 (E) 与波长 (λ) 成反比, 其公式为: $E = hc/\lambda$ 。公式中的 h 是普朗克常数, 而 c 是光在真空中的速度。
 15. 光动力学的术语诸如“单线态”、“双线态”, “三线态”等, 是指原子带不同的电子数量和能量的状态。氧分子的基态 (能量最低) 的是三线态 (3Σ), 带有两个未配对的电子。但当它被一个光子激活时, 它成为一个不带未配对电子的状态, 即具有更高能量的单线态 (1Δ), 因此反应性更强 (能产生破坏性)。
 16. Noell, W.K., Walker, V.S., Kang B.S. et al., Retinal damage by light in rats, Invest. Ophthalmol. 5:450, 1966.
 17. Friedman, E., Kuwubara, T., The retinal pigment epithelium: IV. The damaging effects of radiant energy, Arch. Ophthalmol. 80:265-279, 1968.
 18. Parver, L.M., Auker, C., Carpenter, D.O., Choroidal blood flow as a heat dissipating mechanism in the macula, Am. J. Ophthalmol. 89:641-646, 1980.
 19. Parver, L.M., Auker, C., Carpenter, D.O., Choroidal blood flow: III. Reflexive control in the human, Arch. Ophthalmol. 101:1604, 1983.
 20. Parver, L.M., Temperature modulating action of choroidal blood flow, Eye 5:181, 1991.
 21. Wieland, C., Seeing back to front: Are evolutionists right when they say our eyes are wired the wrong way? *Creation* 18(2):38-40, 1996.
 22. Anon., An eye for creation: an interview with eye disease researcher Dr G. Marshall, University of Glasgow, Scotland, *Creation* 18(4):19-21, 1996.
 23. 大脑皮层是覆盖在哺乳类动物大脑的整个表面的灰色细胞层 (1-4 毫米厚)。在人与灵长类动物的大脑皮层中, 一部分枕叶皮层 (在每个半球的后端) 专门接收两个视网膜发来的信号, 视网膜的信号传递到这里才能产生视觉感知。视网膜中央的 1.5 毫米区域 (黄斑) 不成比例地使用了视觉 (枕叶) 皮层总共约一半的面积。
 24. 接收蓝色光的视锥细胞比接收红色和绿色光的视锥细胞少很多, 并且现在它们被认为可能不存在于中央凹内。
 25. Osterberg, G., Topography of the layer of rods and cones in the human retina, Acta Ophthalmol. (suppl.) 6:1, 1935. As cited in Tasman W., Jaeger E.A. (eds), Ref. 4, vol. 1, ch. 21.
 26. 中央凹的直径弧长约为 20 角分 [按照眼睛中心点测算, $60 \text{角分} = 1^\circ$ (度)], 而眼睛正常的分辨能力, 或可分辨的 2 个点的最小距离所对应的弧度是 1 角分。
 27. 感光细胞产生视觉信号后, 首先通过视网膜内核层的双极细胞, 然后传递到神经节细胞的轴突或神经纤维所形成的视网膜神经纤维层。
 28. Curio, C.A., Allen, K.A., Topography of ganglion cells in human retina, J. Comp. Neurol. 300:5, 1990. As cited in Tasman W., Jaeger E.A. (eds), Ref. 4, vol. 1, ch. 19.
 29. Schein, S.J., Anatomy of macaque fovea and spatial densities of neurons in foveal representation, J. Comp. Neurol. 269:479, 1988. Cited in: Tasman W., Jaeger E.A. (eds), ref. 4, vol. 1, ch. 19.
 30. Streeten, B.W., Development of the human retinal pigment epithelium and the posterior segment, Arch. Ophthalmol. 81:383-394, 1969.
 31. 要确定叶黄素的精确位置, 即要确定叶黄素存在于视网膜感觉神经的哪些层次和结构, 研究人员已经证明这个课题是比较困难的, 但这里所述的是研究人员已达成共识的叶黄素位置。
 32. Gass, J.D.M., Müller Cell Cone, an Overlooked Part of the Anatomy of the Fovea Centralis, Arch Ophthalmol. 117:821-823, 1999.
 33. 这张图是基于多个研究结果制作的, 正如参考资料所示。某些黑色素吸收曲线显示黑色素对短波长的光谱吸收明显的下降, 但这只是由于短波长辐射的传输被眼睛中的介质 (角膜和晶状体) 减少所导致的, 而不是黑色素颗粒对短波光谱的吸收降低。
 34. DukeElder, Ref. 1, pp. 118-125.
 35. Sabates, F.N., Applied laser optics: Techniques for retinal laser surgery, 1997. In Tasman W., Jaeger E.A. (eds), 1997. Clinical Ophthalmology, LippincottRaven, New York, vol. 1, ch. 69A.
 36. Nussbaum, J.J., Pruett, R.C., Delori, F.C., Historic perspectives. Macular yellow pigment. The first 200 years, Retina 1:296-310, 1981.
 37. DukeElder, S. (ed), Ref. 1, vol. 2, p. 264, 1961.
 38. 散射的程度与波长的四次方成反比。
 39. Ham, W.T. Jr., Mueller, H.A., Ruffolo, J.J. Jr. et al., Action spectrum for retinal injury from nearultraviolet radiation in the aphakic monkey, Am. J. Ophthalmol. 93:299, 1982. (术语“aphakia”是指眼睛没有晶状体。猴子眼睛的晶状体被切除后进行实验。)
 40. Boettner, E.A., Wolter, J.R., Transmission of the ocular media, Invest. Ophthalmol. Vis. Sci. 1:776, 1962. Cited in: Tasman W., Jaeger E.A. (eds), Ref. 11, vol. 5, ch. 55.
 41. 以下的解剖特征能使眼睛进一步防止暴露在光和紫外线辐射所造成的伤害, 例如: 人在直立的姿势, 眼睛向着的正常水平方向时, 眉毛 (特别是那些深陷的眼睛), 鼻子和脸颊都有助于保护眼睛。但任何天然的防御系统都可能无法完全保护, 所以使用额外的阻隔蓝光和紫外线辐射的防护装备是明智的 (就像在寒冷的天气中穿更多的衣服一样), 特别是在我们地球的臭氧层不断较少的情况下。
 42. Williams, G.C., Natural Selection: Domains, Levels and Challenges, Oxford University Press, Oxford, pp. 72-73, 1992.
 43. Traquair, H.M., An Introduction to Clinical Perimetry, The C V Mosby Co., St Louis, 1938.
 44. Wertheim, Z., Psychol. Physiol. Sinnes. 7:172, 1894. Cited in: DukeElder, S. (ed.), Ref. 1, vol. 4, p. 611, 1968.
 45. 按照两眼直视前方来计算, 当失去或遮挡一只眼睛时, 整体视野区域大约减少 20 ~ 25% (具体比例主要受到鼻子的形状所影响)。每只眼睛的视野范围通常受到面部轮廓的限制 (详见参考文献 40 中所提到的)。由鼻子造成的视野损失可以通过向盲点稍微转头得到很大程度的恢复。当一个独眼的人定睛细看一个物体时, 他会不自觉地做这个动作。
 46. Diamond, J., Voyage of the Overloaded Ark, Discover, June, pp. 82-92, 1985.
 47. Mollusks, Encyclopædia Britannica 24:296322, 15th ed., 1992; quote on p. 321.
 48. Hanlon, R.T., and Messenger, J.B., Cephalopod Behaviour, Cambridge University Press, Cambridge, New York, p. 19, 1996.
 49. Budelmann, B.U., Cephalopod sense organs, nerves and brain, 1994. In Pörtner, H.O., O'Dor, R.J. and Macmillan, D.L., ed., Physiology of cephalopod molluscs: lifestyle and performance adaptations, Gordon and Breach, Basel, Switzerland, p. 15, 1994.
 50. Sensory Reception, Encyclopædia Britannica 27:114-221, 15th ed., 1992; quote on p. 147.
 51. Budelmann, Ref. 48, p. 15.
 52. Mollusks, Ref. 46, p. 319.
 53. Wieland, Ref. 21, endnote 6.
 54. Kröger, R.H.H. and Biehlaier, O., Spacesaving advantage of an inverted retina, *Vision Research* 49(18):2318-2321, 9 September 2009 | doi:10.1016/j.visres.2009.07.001.
 55. Franze, K. et al., Müller cells are living optical fibers in the vertebrate retina, Proc. Nat. Acad. Sci. USA, 10.1073/pnas.0611180104, 7 May 2007 | doi:10.1073/pnas.0611180104.
 56. Labin, A.M. and Ribak, E.N., Retinal glial cells enhance human vision acuity, Physical Review Letters 104, 16 April 2010 | doi:10.1103/PhysRevLett.104.158102.

彼得·W·V·格尼 眼科医生

1960年在英国布里斯托尔大学取得医学资格。回英国专攻眼科之前, 他花了近6年的时间在巴基斯坦、亚丁、埃塞俄比亚和厄立特里亚的穆斯林群体中进行医疗宣教。他在伦敦的 Moorfields 眼科医院完成了培训, 是皇家外科医学院 (Royal Colleges of Surgeons) 和眼科医师协会 (Ophthalmologists) 的会员, 自1980年以来一直在西米德兰兹郡 (West Midlands) 作为顾问眼科医师执业, 1998年退休。

他是医疗宣教士罗伯特·格尼的兄弟, 他的兄弟也相信创造论。



DNA条码 显示

物种之间的差距和近期瓶颈效应

文/ 流萤

2018年斯多克尔（Stoeckle）和泰勒（Thaler）在《人类进化》杂志（Human Evolution）上发表了一篇文章，题目是“为何线粒体能定义物种？”（Why Should Mitochondria Define Species?）。¹文章挑战了人们对地球生物历史的认知。两位作者分析了近十年累积的DNA条码资料，得出以下结论：

1. 现存动物物种都是“在最近几十万年内从小群体发展起来的”；
2. 动物界各个物种之间在基因序列上有空白区。该文引起了大众媒体的注意，我们作为创造论者如何看待这些发现？

该基因是否有代表性？

上述文章研究的DNA序列比较短，本不是用来做谱系（所谓的“进化”关系）分析的，而是用来鉴定物种的，类似于商店里识别商品用的条码，所以被称为“DNA条码”。然而随着DNA条码数据的日益增多，斯多克尔和泰勒得以从中寻找谱系发生方面的踪迹。

动物DNA条码都取自同一个

基因，即线粒体内的细胞色素氧化酶第一亚单位（COI）。细胞色素氧化酶是细胞在分解食物的时候直接与氧分子发生作用的蛋白质。虽然动物细胞中大部分的DNA保存在细胞核内，但是线粒体内也有少量DNA，即线粒体基因组。该DNA分子是通过母系遗传的，没有父母之间的混合，其序列比核内DNA稳定，而且物种之间也有

较强的一致性（同源性），不同物种的COI序列可以用同样的试剂盒来分析。线粒体DNA的同源性也方便进行谱系分析。尽管COI基因只占线粒体基因组的5%，不到人体总基因组的百万分之一，但是根据该基因建立的谱系树与根据整个线粒体基因组建立起来的谱系树良好地吻合。根据COI序列进行的动物分类也与分类学专家根据总

体性状的分类结果一致。创造论遗传学家纳撒尼尔·詹森（Nathaniel Jeanson）也认为根据线粒体DNA建立的谱系树是准确的、有意义的。²

动物都经历了近期的族群瓶颈？

这篇文章报告的一项重大发现，就是不同物种内部的个体之间在COI基因都是惊人地相似，平均差异在0.0%到0.5%之间（人与人之间的平均差异是0.1%）。种内变异主要是突变造成的。斯多克尔和泰勒认为这些突变并不影响繁殖能力（中性突变），不受选择，所以能够随时间而累积。线粒体基因的突变比核内基因突变更不容易被自然选择所识别，因为线粒体内的基因读码系统比较简单，同义tRNA比较少，而且没有选择性剪接。无论这些突变是中性还是近中性，自然选择都不会终止突变的累积。³

随机突变的累积可用作一种分子钟。虽然历史上不同物种的突变率可能有所变化，但是突变的数目还是和时间相关的。既然所有的动物种群都累积了相同数目的突变，这就强烈提示所有动物群体都是在同一时间内发展起来的。所以，斯多克尔和泰勒得出如下结论：

“就是说，现存的动物族群，不管它们目前的规模有多大，也不管它们与哪一个时代的化石相似，都是在过去20万年以内发展起来的。20万年前，线粒体DNA没有种内差异。”

这两位作者提议，线粒体DNA的种内一致性可以用“瓶颈、始祖效应、世系重整或基因取代”来解释。所谓“世系重整”，就是一种基因形式通过随机漂移排除了其他形式（类似于人类社会中某

些姓氏因偶然因素而消失），“基因取代”则是强力的正向选择的结果。“世系重整在小群体中的效果最强”，而基因取代则需要“群体中所有的线粒体基因组都从同一个母亲重新发源。”

然后作者们推测：

“不知道是否发生了一系列的事件，为动物界各个物种带来了类似的后果，还是只发生了一个事件，影响了整个动物界。根据简约原则（奥卡姆剃刀），应该只用一个过程来解释。”

同样是这两位作者，在几年前还在反对这种想法：

“如果限制变异的是瓶颈效应，那么所有种类动物存在的历史都不长，那就意味着不久前所有的物种都经历了群体数量的锐减。这不太可能——几乎就指向了挪亚方舟的假说——尽管长时期气候变化循环也可能造成广泛的周期性瓶颈。”⁴

他们已经捅开了这层窗纸！数据完全符合年轻地球和全球性洪水，尽管还不能区分这到底是创造的始祖效应还是洪水的瓶颈效应，因为创造和洪水之间间隔的时间并不长。

分子钟可靠吗？

至于线粒体分子钟在历史上是否恒速运作，线粒体DNA的突变率具体是多少，我们无从知晓，但是我们可以根据作者估算时间的方法对这篇文章所给出的时间跨度做出评价。既然现存人类和动物群体是在同一时期起源的，斯多克尔和泰勒便根据公认的人类历史来估算

动物的历史，而人类历史又是基于以下研究，包括“数千人的基因组分析、数万个线粒体基因序列、古生物学、人类学、历史学和语言学”。化石（古生物学）在分子钟校对中起了重大作用，^{5、6}但是对化石的年龄争议很多。相反，詹森只考虑目前观察到的线粒体DNA突变率（每代每人5-8个突变），发现人类线粒体基因组的变异情况更符合圣经记载的六千年左右的历史，而不是主流科学家认为的20万年。^{7、8}

林奈和达尔文

斯多克尔和泰勒的另一个重要发现是动物族群之间缺乏基因序列上的连续性。虽然种内个体之间的变异很小，但是种间的序列差异却很大。他们指出：

“DNA条码的分布有两个重要特点：1. 群内变异低；2. 群间存在序列上的空白区，没有过渡序列。”

这本不足为奇，因为条形码就是用来区别物种的。如果物种之间在序列上有严重的重叠，COI基因就不能当区分物种的条码用了。然而斯多克尔和泰勒强调这种不连续性是很有道理的。分子谱系分析本来就是要将物种间的相似性和差异性做定量分析，而DNA和蛋白质序列一直给人一种印象，就是物种之间只有量的差异。与宏观性状不同，人们很难在分子序列中看出质的区别。由于今天大多数谱系树都是基于序列比较，分子进化论对此方法有强烈的依赖性，尽管这些谱系树与古生物学分类会有冲突，不同的分子谱系树之间也会相互矛盾。⁹然而随着DNA条码资料的大量积累，斯多克尔和泰勒却看出了线粒体基因组的种内变异被“限制在狭窄的范围内”。

他们正确地回答了林奈与达尔文之争，这两位作者引用艾夫斯（Aves）等人的说法：

“艾夫斯和他的同事们在关于谱系地理学的一篇开创性著作中，提到了生物学上的一个长期争论，即林奈学派与达尔文学派之间的争论，林奈学派认为物种是各自独立的，而达尔文学派则强调种内的渐进的变化产生了新物种。”¹⁰

卡尔·林奈（Carl Linnaeus, 1707-1778）从亚里士多德时代流传下来的“自然等级”出发，撰写了他的《自然系统》，开创了生物分类学和双名法，强调物种分别，物种间有不可逾越的鸿沟。同时代的布丰（Comte de Buffon, 1707-1788）则反对这一观点，撰写了《自然史》，倡导自然界有“看不见的过渡”。¹¹关于物种间是连续还是间断的争论一直延续到今天。斯多克尔和泰勒指出：

“种内条码紧凑成簇，种间却一片空白，这些都是动物界的关键事实，进化论必须予以解释。”

他们甚至为此提出了一个假说：

“最亲近的现存物种之间的差距或大或小，这可能反映了不同数量的已经灭绝的过渡序列。”（下划线是本文作者加上去的）

进化论的新定律？

斯多克尔和泰勒认为空白区可以用灭绝来解释，他们引用了1973年范瓦伦（Leigh Van Valen）提出的“进化论的新定律”，又叫“灭绝定律”。范瓦伦写道：“现有资

料表明，所有的族群都以特有的、恒定的速度走向灭绝。”¹²

这一定律主要是基于化石记录。按照范瓦伦的说法，灭绝的主要原因是种间的生态互动，所以灭绝的概率与物种的年龄无关。范瓦伦和他同时代的多数人一样，相信新物种在不断地出现。

既然没有一个物种能长期生存，那么从历史上的任何时候去看，所有的物种必然都是新的。然而，该定律认为不同种群走向灭绝的速度不同，所以不能解释为什么现存动物种群的线粒体年龄都差不多。

值得指出的是，灭绝定律得到了当代生物学的支持。植物遗传学家圣弗德（John Sanford）在他的《基因熵》一书中，提出所有的物种都有一定的寿命，不是由于外在因素，而是由于基因组因随机突变的累积而最终瓦解。³圣弗德不认为新的生命形式会自然发生，所以整个生物界的历史也不是很长。

下一步

斯多克尔和泰勒指出，还有23个小型无脊椎动物门，其DNA条码的分布尚不明了。随着条码分析资料的日积月累，它们的历史轨迹也会越来越清晰。

创造论生物学家一直试图为圣经上讲的“受造类”做出明确的定义。这些“类”是现存的和已经灭绝的一切物种的祖先。斯多克尔和泰勒的研究虽然显示多数动物属内缺乏种间连续性，但是也有一些个别的属，比如熊属，种间变异相对不大。这不难解释，熊属内的各个种可能都是同一个受造类的后代，至今种间还可以杂交。或许DNA条码数据可以帮助我们鉴定受造类并了解大洪水前后受造类的历史。

结论

斯多克尔和泰勒的文章之所以能一石激起千层浪，是因为他们所发现的事实挑战了统治学术界多年的达尔文渐进主义思想。虽然动物种群的绝对年龄值得讨论，但是普遍年轻的线粒体基因是与近期创造或者大洪水所造成的基因瓶颈是一致的。这些发现不仅不支持漫漫长进化的信条，而且DNA序列数据的不连续性也驳斥了新达尔文主义者一贯倡导的突变过渡机制。 ■

参考文献和注释

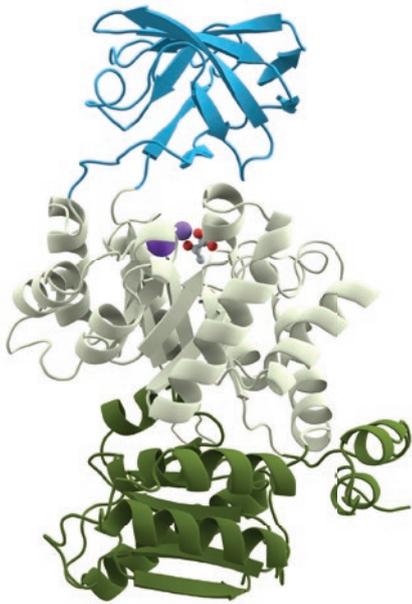
1. Stoeckle, S.M. and Thaler, D.S., Why should mitochondria define species? *Human Evolution* **33** (1-2) :1-30, 2018. 10.14673/HE2018121037.
2. Jeanson, N., *Replacing Darwin: The new origin of species*, Master Books, Green Forest, AR, pp. 248-249, 2017.
3. Sanford, J.C., *Genetic Entropy and the Mystery of the Genome*, FMS Publications, Waterloo, NY, pp. 45-68, 2008.
4. Stoeckle, M.Y. and Thaler, D.S., DNA barcoding works in practice but not in (neutral) theory, *PLOS ONE* **9**:7e100755, 2014.
5. Cann, R.L., Stoneking, M., and Wilson, A.C., Mitochondrial DNA and human evolution, *Nature* **325** (6099) :31-36, 1987.
6. Stringer, C., The origin and evolution of *Homo sapiens*, *Philosophical Transactions of the Royal Society B: Biological Sciences* **371** (1698) :20150237, 2016.
7. Jeanson, ref. 2, pp. 182-186, 2017.
8. 一篇近期的综述总结了分子遗传学的证据，表明创造是在近期发生的，详见：Tomkins, J., Empirical genetic clocks give biblical timelines, *J. Creation* **29** (2) :3-5, 2015.
9. Tomkins, J. and Bergman, J., Evolutionary molecular genetic clocks—a perpetual exercise in futility and failure, *J. Creation* **29**:26-35, 2015.
10. Avise, J.C. Neigel, J.E., and Arnold, J., Demographic influences on mitochondrial DNA lineage survivorship in animal populations, *J. Molecular Evolution* **20**:99-105, 1984.
11. Sloan, P.R., The Buffon-Linnaeus controversy, *Isis* **67** (3) :356-375, 1976.
12. Van Valen, L., A new evolutionary law, *Evolutionary Theory* **1** (1) :1-30, 1973.

流萤

中华创造科学协会副会长；上海医科大学医学硕士，俄亥俄大学生物学博士，Liberty University（利伯缇大学）医学院微生物学副教授。病毒学专家，研究病毒、遗传学、基因编辑。

蛋白质家族：

是随机产生的？ 还是被设计的？



文/ 罗尤·杜鲁门和迈克尔·海西希

摘要：进化的计算机模型假设自然选择可以通过连续地对有利的突变基因进行选择来进化出新的遗传信息。^{1、2} 我们证明新的基因家族必须先解决统计上“不可能”的困难，然后达尔文的自然选择过程才能发挥作用。

如果一个分子机器无法在现实中通过偶然的出现，那么一个由数百个生化过程精妙配合来构成的细胞也就永远不会偶然出现。就如一头牛不能跳过一座大楼，那么它就永远不能跳过月球。

单个存在的基因是没有生物功能的，因为所有细胞都需要多种由

不同基因编码的蛋白质来共同执行细胞的多个流程。当进化论者被问及基因是如何同时出现的时候，他们有时会引用“共同进化”这个概念，即一个基因的拷贝进化出一个新的序列，并且与其他已存在的基因一起发挥作用。由于当前的生物功能A（图1）需要多个基因，其祖先的功能可能是A₁，A₂……现在的基因假设是由之前的功能基因突变产生的新基因。

方框表示特定功能所必需的基因组合。按照进化论的推测，修改基因可以进化出一种新的功能。它们以前的功能是怎样来的（基因是如何起源）这个问题仍然需要回答。（黑色：今天存在的蛋白质部分；深灰色：蛋白质被修改的部分，浅灰色：假设的之前蛋白质功能所需的部分。）

这种解释造成了更大的困境，因为这就需要先解释该单一基因之前的多个有功能的基因是怎样来的。这样一来，就引入了一个更大的难题（图1）。³ 然而，唯物主义的框架假设生物的复杂性是源于更简单的状态。

一个细胞功能往往是由成千上万的蛋白质互相配合完成的，比如特异酶的催化作用。⁴ 因为没有证据表明这一类蛋白质之前具备其它

功能。所以认为自然界中所有的基因或者蛋白质都是源于单个生物体祖先是不合理的，因为蛋白质的序列长度的范围可以从几十个氨基酸到30000个氨基酸，⁵ 这表明蛋白质是源自于许多序列不相关的蛋白质家族。

我们在这里先忽略生命自然发生的问题，假设存在一些能够成功繁殖而且不会自我消亡的简单生命形式。神导进化论者可能会认为上帝用了进化的过程（使万物出现），且在进化过程中没有进行积极的干预。那么自然遗传过程能否在漫长的时间里进化出新的、差异极大的蛋白质呢？

生成一个新的蛋白质家族

让我们只考虑这一难题的一部分：假设产生了新的基因，此基因能够编码具有足够的（任何种类）有生物功能的蛋白质，让达尔文的自然选择可以发挥作用，开始对基因序列进行微调。这是如图2所示（见“功能失效点”）。进化不可能预见未来，不可能像进化论期待的那样，选择带有与有用的基因相似度这么小的随机基因序列的生物体。¹

我们知道的所有进化计算机模型都忽略了随机DNA序列能产

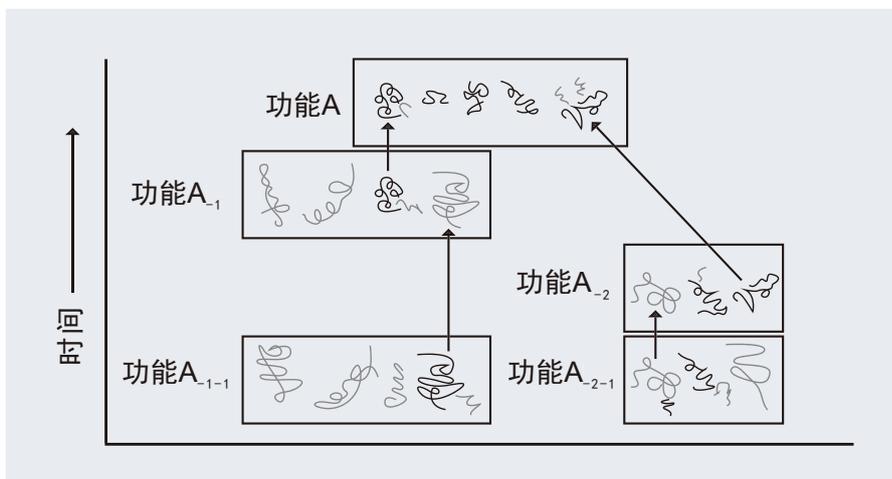


图1：进化论的共同进化产生基因的假设

生最小功能的、正确折叠的蛋白质几率是极低的。^{1、2、6} 进化论者诉诸于“之前的基因”根本不能解释这个问题。新的无相互关联的基因家族必须有个源头。认为大自然通过不断复制一个原始基因从而产生了今天观察到的大量的不相关的基因，这种想法是愚蠢的。针对我们正在研究的问题，使用亚基因部分“蛋白质结构域洗牌”的论点也是没有说服力的。在进化论的理论框架中也没有人会争辩，目前（截至2001年3月1日）所知的31474个已知蛋白质结构域在最早的生物基因组都已存在。⁷ 而且，从中选择一个或多个成员，组织并将它们装配到基因的正确位置，是一个永远不可能完成的任务。⁸

进化论假设我们的祖先可能是类似于细菌的生物。但请注意，这些类生物通常不会有大量多余的可用于进化试验的DNA。

最小功能的多肽

广泛存在于不同生物中的最长序列的蛋白质是细胞色素c。将不同物种的细胞色素c基因序列排列好后，发现在一些位置上某些氨基酸缺失了。假设这缺失的氨基酸是非必要的，在所有生物体中，我们就找到了110个氨基酸的共同序

列。⁹ 一个平均序列长度的蛋白质比细胞色素c要大得多，一般由约350个氨基酸组成。¹⁰ 让我们来看看功能最小的新基因是否能独立进化产生。首先，需要有一些不受关键生物功能限制的基因，用于任意修改，我们暂且不讨论这些基因的起源问题，简单地假设能从已存在的基因复制后产生一个与其简单相似的基因（可用于任意修改），不然我们需要探讨的就是其母本基因的起源。

让我们将这个DNA新的部分视为随机碱基对序列。在可以使用达尔文的自然选择来加速进化到新的基因之前，这个序列的改变需要一个反复试错的过程，以产生一个可接受的最小功能基因序列。

生殖选择是取决于即时的生物学价值，而不是取决于这个序列是否像一个遥远的（进化）目标基因。^{1、2、7}

进化到第一个细胞色素c的机会

在附录1中，我们总结了约克奇（Yockey）基于细胞色素c的概率计算。由于一些氨基酸在自然界中很少使用，因此，要产生功能蛋白质的真实范围比产生特定长度的多肽的全部可能性要小。推理结果是，由氨基酸随机组成有功能的多

肽的可能性极低，几乎不会随机产生。

由于进化论者认为这个基因家族超过十亿岁，¹² 所以有无数机会生成各种不会影响生存的突变基因。约克奇按照由Borstnik和Hofacker开发的模型，^{13、14、15} 慷慨地扩展已知细胞色素c序列的清单甚至假设许多在自然界中找不到的序列也是可以接受的。估计可接受的细胞色素c蛋白序列如附录1（20）所示。

最小功能的细胞色素c多肽与长度为110个氨基酸的所有序列（不包括那些概率非常低的）¹⁶ 的数量比例，使我们估计进化和选择可以发挥作用之前的最小功能的细胞色素c的比例：

$$\frac{2.3 \times 10^{93}}{1.15 \times 10^{137}} = 2.0 \times 10^{-44} \quad (1)$$

另外两种合理评估的计算方式

另外还有两种方法可以估算这个比例，都比较容易理解。第一种方法，我们先确定获得一个在蛋白质中每个位置编码的可接受的氨基酸密码子的几率，然后将每个几率相乘。回顾一下遗传学的知识，每种氨基酸可以由1至6个遗传密码子代表。从附录2的表2底部的第二栏，我们得到：

$$2.7 \times 10^{-44} \quad (2)$$

另外一种算法是忽略同义密码子的比例并使用每个氨基酸的通用遗传密码子，只考虑从20种候选氨基酸中可接受的氨基酸数量来进行估计（见附录2，表2底部的最后一栏）的

$$6.9 \times 10^{-45} \quad (3)$$

后面这种估算方法对于自然发生的假说而言是更加合理的，这种假说假设氨基酸是存在的，并且（在没有遗传密码的情况下）随机连接起来，并且假设那些存在的氨

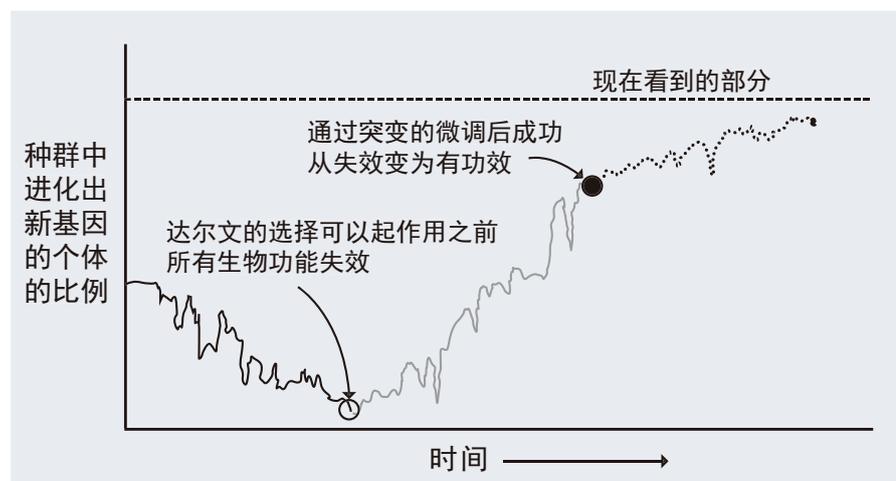


图2：假设从重复基因中进化出新基因群体的基因频率变化。

基酸都只为L-型的氨基酸。

随机试错的可行次数

最有利的进化情景是在无性繁殖的生物中发生，例如通过出芽生殖或分裂机制繁殖。为确保细胞色素c能够转移到整个生物圈，我们的进化场景假设了一种类似细菌的生物。这场景必须发生于寒武纪生物大爆发之前的至少十亿年前（进化论者认为寒武纪生物大爆发在大约5.5亿年前发生），以确保所有生命形式都可拥有细胞色素c。在附录3中，我们提出了以十亿年进化出一个新基因的最大可能的进化随机试验次数： 2×10^{42} 次（25），但是我们不得不使用不切实际的、数量巨大的、同源的生物群体；快得难以置信的平均每代繁殖时间；以及一个非常快速的突变频率，并能以某种方式避免突变过快而失控造成自我毁灭。

最大试错次数（25）和110个氨基酸序列长度能产生最低功能细胞色素c的比例（1）让我们可以计算碰巧随机产生有用的细胞色素c突变体的概率：

$$1 - (1 - 2 \times 10^{-44})^{2 \times 10^{42}} = 0.04 \quad (4)$$

讨论

按照在（4）中使用的对进化模型有利的假设，提出：在10亿年间一个数量大得荒唐的生物群体，进行高速的繁殖，并有很高的突变率。在进化出最小生物功能的细胞色素c之前，此时不能假设存在任何繁衍优势。在达尔文的进化论能派上用场前，从无到有以产生单个体的生化过程的最低要求究竟能否得到满足，目前的分析结果对此提出很大质疑。

我们将在后面的文章中讨论群体遗传学和基因序列微调的细节。

从（4）看来，即使是一种非常小的蛋白质在所描述的最佳条件下也不太可能自发地偶然产生。这种违背统计学几率，即使真的发生了，进化的流程必须在图2中的启动点之后才会开始微调步骤。带着一个小优势的幸运的生物现在与一个极大的生物群体竞争，其群体中个体数量计算见附录3的（22）×（23）：

$$(4 \times 10^{20}) \times (1 \times 10^{11}) = 4 \times 10^{31} \quad (5)$$

费舍尔对有性繁殖生物体的分析结果表明，使用不切实际的高选择系数 $s = 0.1$ 对有利的突变进行选择，在10000或更大的生物群体中，有利突变只有2%的几率能在群体中被固定下来。¹⁷

我们只要求随机突变找到一个允许达尔文的选择能起作用的最小功能的细胞色素c，因此非常慷慨地假设选择系数 $s = 0.01$ 。请注意，在每一代中都至少有一个突变后代必须存活下来，否则这个进化过程将完全失败，我们可以设想一个分裂或出芽生殖的模型：平均每10分钟原始细菌就会分裂或死亡（因为这个种群在其环境中已经达到最大化，不能容下更多的数量）。非突变体复制的概率是 $p_0 = 0.5$ 。而突变体复制的概率是 $p_0 \times (1 + 0.01) = 0.505$ 。假设这个预设的庞大种群数量将允许一个成员的轻微地、局部地（至少是暂时的）增加，即突变体存活1代的概率是 $p = 0.505$ 。

越过了这个障碍之后，新突变体数量 $n = 2$ ，使用二项式概率分布计算死亡的概率，使用 $x = 0$ ：

$$\frac{n!}{x! (n-x)!} p^x (1-p)^{n-x} = \frac{2}{1 (2-0)!} 0.495^2 = 0.245 \quad (6)$$

因此，新突变的生物存活概率仅为

$$(0.505) \times (1 - 0.245) = 0.381 \quad (7)$$

新突变生物经过2代的存活概率是0.381，这个概率在庞大的生物种群中并不是太好。

另一个考虑因素是突变体的积累，开始的时候是非常缓慢和随机的，如图2所示。由于成功地分裂产生2个子细胞，预期经过t代后的突变体数量计算如下：

$$(2 \times p_0) (1 + s)^t \quad (8)$$

其中 $p_0 = 0.5$ 。

若使用 $s = 0.01$ ，经过100代后我们平均预期只能得到2.7个突变体；若使用 $s = 0.001$ ，这是一个更接近现实情况的选择系数，这意味着我们需要经过1000代才能平均得到2.7个突变体。

然而，随着时间的推移生存机会可能会大大偏离平均水平，特别是在局部区域，且由于几个外部因素的原因。在任何一升的水中，在数百万年中，细菌数量将改变几个百分点。假设最初在每升水（23）中有 1×10^{11} 个非突变体，在世界范围内这些细菌不太可能全部死亡。

但由于突变体的数量很少，只有1或2个突变体，它们所面对的局部竞争困难也能在几代时间中将它们轻易全部清除掉，即使我们期待较高的平均s值能使它们不断增加。这种情况只需要在突变体数量累积的时期内发生一次就足以摧毁所有进化的进程。这类似于投资股票，假设多年来的股票市场呈现出上升态势，价值增长约为每年10%。最初投资1美元，每10分钟观察这个股票的表现。该进化就好像如果它只是一次价值低于1美元，你必须（放弃之前的）用新的一代再开始，全部重头再来。事实上，即使是每年增长10%的股票，即使经过无数

股票和很多年的尝试，都并不能保证，不会在每10分钟的观察频率内出现低于原始投资价值（1美元或1个幸存生物）的情况。

我们所知道的所有计算机进化模型简单地默认那些突变体肯定存活，并定能进化出复杂的生物功能^{1、2、6}而忽略了不断出现从零开始进化的要求。

值得注意的是，在新基因的精确位点获得有利突变的概率要远远低于在基因组其它位置累积的破坏性突变的概率。精调的基因被破坏的概率远远大于功能齐全的新基因形成的概率。

图2说明了另一个非常重要的难题，我们从未看到在进化模型中讨论过这个问题：在DNA进化试错的过程中，进化中的带无用基因的生物，在每代时间很短的“简单”生物的群体中占比是不断下降的。我们已经知道没有或减少多余的无用的基因组材料会产生可衡量的繁殖优势。^{19、20、21、22}复制DNA所需的材料和能量更少，这样复制错误的几率更少，繁殖周期较快。对生物体来说，不制造没有价值的多肽是非常重要的：因为能节省材料和能量，而且避免干扰必要的生物功能。如果有20-30%的细胞质由没有正确折叠的蛋白质和多肽组成的话，这些无用的蛋白质会通过疏水相互作用和结合，而使细胞内乱套。²³朊病毒就是一个由缺陷的多肽造成危险的例子。

弗雷德·霍伊尔已经研究出芽繁殖或二分裂繁殖的详细信息：²⁴

$$x = \frac{x_0 e^{(st)}}{1 + x_0 [e^{(st)} - 1]} \quad (9)$$

其中x是尝试进化出一个新基因的生物群体数量的比例；s是选择性因子；t是每代繁殖的时间。

举例说明，假设在十亿年的开始时或过程中，大自然正在寻找最小功能的细胞色素c，而99.99%的生物具有“非必要的”DNA材料可供进化试错：即 $x_0 = 0.0001$ 。在这种“简单”的基因组生物中，不带“非必要”额外基因的生物，与带额外基因的生物相比，前者基因组约小0.01%到1%的数量级。不携带额外的基因组的优势，可以转换为更快的繁殖速率。或许，它们每代的时间从10分钟可以平均缩短0.1秒。我们不能通过实验来确定自然条件下的选择系数s是多少，让我们假设一个非常妥当的优势， $s = 0.000167$ （基于缩短的每代时间，即0.1秒/600秒）。

现在，我们假设一个非常短的每代时间，以增加随机试错的可用次数来提高通过随机突变产生出的最小功能的新基因的几率。但是，供细胞色素c进化所需的非必要DNA占了基因组相当大的比例，因而产生较大的繁殖劣势。事实上，从（9）平均来看，若某一个体的繁殖速度稳增0.01%，经过3年之后，它的后代将在这个种群中占比99.99999%！²⁵此时将没有多余DNA供进化过程试错。无可否认，自然选择并不完美，而且局部幸存者有可能坚持更长时间。新的遗传意外突变可能会偶尔重新引入非必要的DNA。这些突变必须大小合适，而且不干扰有功能的基因。此后，自然选择将继续稳定地选择那些能够一块一块地丢弃新垃圾（无用的DNA）的后代。

我们假设的每代时间缩短了0.1秒，只用一个实验上已知的事实就可以证明。大肠杆菌中有2个（DNA复制的）生长叉，每个生长叉每秒可以复制不到1000碱基对（bps）的DNA。²⁶新的类似细胞色素c大小的基因，包括其调控区域，长度大约400bps。若使用两个复制

叉来复制该基因组，就会多要0.2秒。在其他条件不变的情况下，在快速繁殖的微生物中，携带供新基因进化的额外DNA的突变体，繁殖效率会较低。

在基因组非常小的生物群体中，若某个基因不是快速复制所急需，就会被自然选择淘汰。因此，在我们假定的十亿年进化时间中，供进化细胞色素c的、大小位置都合适的遗传材料，很快就会被淘汰。

这些假设是否切合实际？

我们使用了慷慨的假设，以便做出明确的判断。很容易证明，与小群体相比，在庞大的生物群体中一个新的基因出现并被固定的可能性更大，虽然在较小的群体中新基因能更快速地固定，但不足以弥补小群体的每代的可用试错数量少的劣势。

进化论者将突变视为近似是随机的，以避免目的论的风险。此外，促进“最优”基因从零开始进化的任何因素都会降低其它不相关序列出现的概率。唯一的问题在于（1）计算中得到的最小功能的细胞色素c与无功能的多肽链的比例为 2.0×10^{-44} 是否被低估了。

或者我们可以继续深入地来讨论， 2.0×10^{-44} 的概率是基于蛋白质的序列计算出的，当我们以进化的模式考虑编码这段蛋白质序列的基因时，与实际情况相比，这个概率实在太慷慨了，原因有以下几点：

1. 细胞中的各种蛋白质必须按照合适的比例来合成：单个蛋白质是没有价值的，而过多的蛋白质又是致命的。蛋白质转录的调控涉及激活蛋白或阻遏蛋白，它们会与基因附近特定的DNA序列（即A、C、G、T的不同组合）结合。每

个（激活蛋白或阻遏蛋白的）识别序列通常为5至40个碱基。有些识别序列必须是完全正确的，有些识别序列的某些位置允许2个或更多个碱基替换。由于结合序列的错误识别导致调节蛋白的过多结合或不正确的结合是极其有害的。在进化过程的无数次试错试验中，必须确保基因组的其他位置不会生成太多的调节序列——这种情景下最好的结果是需要过量的调节蛋白，最坏的结果是这些序列会阻止正常的基因表达。除了需要具备合适的调控序列以控制基因转录、以及识别基因开始/结束的位置以外，结合位点相对于基因编码区的位置也是有限制的。

我们假设，新进化的基因附近仅仅需要2个结合位点，每个位点长度是5个碱基，而且不变，以让合适的2种表达调节蛋白质结合。²⁷

我们忽略了基因调控的空间要求、特异性的调节蛋白的来源以及需要从基因组中消除产生错误结合的序列。要计算总几率，仅需要将获得最低功能基因的概率乘以这两个结合序列随机形成的概率：

$$4^{-2(5)} \approx 1 \times 10^{-6} \quad (10)$$

即使有通过之前的一个基因复制产生另一个新的基因，也必须考虑到以下问题。新的基因不但必须通过突变之前存在的序列来创建新的有用功能，而且还需要改变之前基因的调节蛋白的结合序列，以改变基因转录的调控计划。所有这些改变必须通过随机突变，并且需要同时通过共同进化改变调节蛋白相应的三维结构，以便适合我们新的基因所需的

相互物理作用。

2. 现在已经知道在广泛的不同生物体中细胞色素c蛋白质的生物功能。但是否所有生物都可以利用全部这些多样性的细胞色素c是值得怀疑的，约克奇也指出了这个问题。^{28、29}
3. 约克奇进一步假设即使所有种类的细胞色素c理论上是相互兼容的，最后带不同的替代氨基酸的细胞色素c的三维蛋白质结构可能真的是一致的。但并不确定所有这些可能的细胞色素c是否能折叠到产生预期的可接受的蛋白质结构。⁴⁶
4. 虽然在蛋白质中没有“停止转录”或“终止子”的三个密码子表达，但它们必须放置在基因的正确位置。DNA序列转录的多肽长度若不是约110个氨基酸，就不会产生功能最小的细胞色素c，但这也减少了可接受的候选基因的比例。

进化框架的可信度

我们现在对自然或随机过程是否能产生一个非常小的蛋白质的新基因家族已经提出了相当大的疑问。而且还有比细胞色素c变异更小的基因，如组蛋白和泛素，另外大多数基因长度都大于细胞色素c。人们也可以只分析蛋白质的一部分，³⁰如高度保守的蛋白激酶核心部分，其长度为260个氨基酸。这些例子的最小功能成员的比例远远小于在（1）中计算的 2.0×10^{-44} ，并且最小功能成员不可能通过随机试错试验来产生。

当看到细胞的过程需要多种不相关的蛋白质一起才能执行，就需要认识偶然产生这些蛋白质的几率是将每个单独的蛋白质自然出现的概率相乘。例如，细菌操纵子是连续基因的簇，基因簇整个地进

行转录，再从中产生多种蛋白质。色氨酸的操纵子产生五种蛋白质，当需要制造色氨酸时，它们一起由一个细胞流程产生。这个方案确保蛋白质产生的相对比例一样。³¹偶然进化出含n个基因的操纵子大致几率是，进化出一个平均序列长度基因的n次方。³²我们至今还不知道任何宣称：未同时具备所有n个基因成员的操纵子，而能够具有生物学功能。操纵子似乎带有“不可简化的复杂性”，不能逐步进化产生。

由于与其他蛋白质的复杂相互作用，有功能的蛋白质序列的比例在某些情况下是非常小的。³³或者使用以下这个例子更容易理解，成千上万特异的蛋白质——酶：必须有特异的三维腔，形成适合特定化学反应的中间产物的空间和电子结构，好像手和手套那样契合。这降低了生成（决定反应速率的）中间产物所需的能量，因而能加快整个反应的速率达数百万倍。不需要直接参与催化作用的蛋白质部分也是有重要的作用，有些能稳定酶的结构以确保其催化效果；有些部分确保酶随着时间的推移还能正确地折叠，而保持其成熟性；其他部分可能用来确保与其他蛋白质进行正确的相互作用，或将酶定位于细胞的特定位置。

所需基因的最小数量

虽然寄生性支原体是不能独立生活的细胞，但其被用于估计生物体生存所必须基因的最小数量，它至少要在实验室的生命支持系统中进行小心的培养。这些支原体的基因被故意地敲除，看看哪个基因暂时是可有可无的。研究结果表明，生存至少须要250-400个基因。³⁴通过随机（进化的方法）获得多个不相关的基因（这些基因是生物生存的最低要求）的总体概率大约相当于将随机形成每个基因的

概率相乘。如果这些基因具有细胞色素c类似的长度和多样性的特征，要满足最小的勉强生存的要求（必须已在合适的膜中，而且为它们提供能量和营养需求，某种程度上的翻译和转录功能已经具备）的生物随机进化出来的几率是 $1 \times 10^{(-44) \times (300)}$ 的数量级，可以十分有把握地说通过随机试错试验的进化方法，是永远不可能发生的（假设一个最小的细胞必须至少有300种蛋白质）。

给进化论者带来大麻烦的预测

进化论要求如果成千上万的不同基因家族不能同时出现在原始的共同祖先中，那么这些基因必定是在漫长的时间中逐渐进化产生的，“基因正处于产生新功能的进化过程中”的证据应该随处可见。

如上所述，人们不能证明每一个细胞和生物功能都与之前的一个细胞或功能有关。但理论上基因进化应该是这样的，拥有一项现在的功能，并且正在准备下一个新的功能。然后在一眨眼的时间内，这个基因中止了A-2功能，并专注于A-1功能。但是，很多细胞的过程与任何其他已知的过程没有任何相似之处。没有找到这些流程之间的过渡桥梁，同时这些桥梁也是不可能存在的。按照进化理论预测当一位分子生物学家发现一个新的基因时，没有理由期望它已具备生物功能。按照进化论的基因历史观点，基因将主要处于预备新的功能的状态。请留意进化的信念会限制科学的发展。因为，按此观点，大多数的基因都尚未具备功能，科学家就不会有动力去探索新发现的基因的功能了。然而，有趣的是要注意到这一点，生物学家都很自然地假设一个新发现的基因将被证明具有某种功能。

创造论理论模型需要解决各种

科学问题（图3）。在挪亚洪水后，我们尚不清楚在基因家族内和圣经所定义的类之间的基因组存在多少多样性基因。我们认为物种内已经存在很多的多样性，以适应不同的需求和环境。人的堕落会引入基因的缺陷。挪亚洪水产生了严重的瓶颈（它毁灭了许多具有更好基因的成员），并且在洪水期间和洪水之后有害的环境因素也会促进许多破坏性突变的扩散。我们认为生态位的缺失，以及生物种群的迅速增长，使具有缺陷基因组的生物群体可以形成较大的种群。而地理的隔离使遗传库逐渐分化。自然选择只能淘汰较大的遗传缺陷，并将在一个洪水后的世界环境下发挥作用。与均变论所预期的相比，创造论预期上述因素的整体效果是在更短时间内产生更大的基因变异，以及基因缺陷的序列。

自然选择不能确保淘汰每一个缺陷原始基因组。特异性较小的基因序列的比例总是大大超过特异性较大的基因序列。生物能生存下来是一种（受到多因素影响的）非常随机的效应，在我们看来突变净效应是会破坏特异性和基因功能的。按照这个观点来预测，我们确实会找到一个不再执行功能的基因。通常这些基因将不再产生m-RNA。与进化所预测的情况相反，在自然条件下很多基因序列是随机变化的，而非趋向改善的。即使在理想的实验室条件下，使用加速诱导突变，经过无数的果蝇遗传实验和快速复制的大肠杆菌实验，突变都未能产生（增加）有功能的基因信息，突变反而导致了严重的基因退化。

结论

与细胞色素c相近似的，具有最小生物功能的基因序列，与等长度的所有可能的DNA序列比例，

估计为 $1 \text{比} 5 \times 10^{43}$ （来自（1）式的倒数）。因此，我们需要许多突变试错试验来支持进化模型，通过假设单一种类的类似细菌的生物群体，它们每代繁殖时间很短，群体数量众多，无性繁殖，丰富的营养，以及非常高的突变率（我们忽略突变对基因组的其余部分的影响）。理论上前寒武纪的生物，如三叶虫或贝类的祖先，有更大的基因组，它们的每代时间（a）比我们假设的10分钟要长许多个数量级，种群数量（b）会小许多个数量级。这些事实大幅降低了试错试验的数量。所以，如果我们假设的类似细菌的生物都不可能通过这种随机的进化机制进化出新的基因家族（那么前寒武纪的生物就完全没有可能按照进化方式形成新基因）。

我们对细胞色素c这个蛋白质的起源不感兴趣，但试图确定是什么样的多肽可能随机进化出不相关的基因家族。³⁵人们通常认为有功能的蛋白质由超过100个氨基酸组成，大小接近于本研究的蛋白质。^{36、37}平均来说，超过20种氨基酸的一半可能可以用于（蛋白质的）任何氨基酸位点的情况十分罕见。细胞色素c甚至可能是从更简单的祖先基因开始吗？只有在可以减少限制条件的情况下，才可能对进化论者有意义，不然会很快变成荒诞的遐想。如果多个生物的功能允许太多的灵活性的话，自然选择就没有一致的标准来发挥作用了。

此外，假设一个更简单的祖先基因来进化，是否能进化出任何有用的基因是值得怀疑的，产生最小功能的蛋白质的试错试验数量确实会比我们计算的更小。但现在随机进化确实需要很多代的生物才有可能突变进化出这种有用的基因，与其他基因一起产生一种全新的生物功能，如现在的生物所拥有的功

能。与其在无休止地回顾（更简单的基因），不如接受一个最小功能的基因也必须有一个（智慧设计创造的）起点吧。

今天对生命细胞的研究揭示了生命极度的复杂性，这个复杂性是达尔文及其之后的进化理论家无法想象的。例如，DNA分子的阅读机器³⁸执行特定的生物功能，由多个独立的部分以高度协调地方式一起工作。

最近报道了嗜热栖热菌（*Haloarcula marismortui*）的核糖体的大亚基的完整原子结构，它由超过3,000个核苷酸和27个特异蛋白质组成。³⁹它是核糖体发挥作用的不可缺少的组成部分，在每个细胞中有多个拷贝。这个分子机器对每个信使RNA进行多次解码以决定连在一起的氨基酸的顺序以制造蛋白质。此外还需要额外的配件，例如，在翻译一个mRNA的密码子时，^{40、41}必需得到一个可靠的能源供给，在正确的时间，以合适的水平输送到正确的位置。

贝希教授发现了很多生物学功能的例子，都具有“不可简化的复杂性”的特点，^{42、43}即除非所有组成部分都齐备，并且被精细地装配，否则不会具有任何的生物学功能。

然而，相对于长度相近却无功能的DNA的序列，有功能的基因序列的占比非常小。对于细胞色素c，一种远小于平均水平的蛋白质，我们已经计算出有功能的基因序列的比例仅为 2×10^{-44} ，但是对于已知的最大的蛋白质（其几率更低）。萨奥尔使用盒式诱变的方法进行研究，^{44、45}来估算多肽能够正确折叠的比例，正确折叠是一个蛋白质发挥作用的前提条件。⁴⁶其研究结果估计几率大约为 10^{-65} ，但在该研究组（的多肽）中甚至被证明均没有生物学功能。这与本研究

中统计学上的估算一致：其几率就如一个原子与整个银河系的原子相比的几率那么小。

另外，基因表达必须在其附近有特定的碱基序列，以允许预先存在的调节蛋白与之结合。突变试错试验必须随机产生这些结合序列并同时从基因组中消除随机产生的不正确的序列。

即使我们用了慷慨得不切实际的假设，宣称能在类似细菌的生物中通过突变（不论用设计还是用随机的方法）产生大量蛋白质家族都是不合理的。我们的结论仅仅基于在进化的选择机制能够开始发挥作用之前，一个单一的，新的基因不可能自然地发生。我们的计算不是为了优化出最佳表达的良好基因序列，而只是计算在达尔文的微调理论发挥作用之前能够克服障碍的可能性。

我们观察到的细胞内多种没有进化关系的蛋白质精确地配合，产生了数千种不同的细胞功能，最合理的解释为这是被智慧地和有计划地创造而产生的。

致谢

与休伯特·约克奇的深入交流解决了一些困难，我们期待着他第二版的著作。我们感谢与Answers in Genesis合作的匿名审稿人的批评性评论和分析。

参考文献和注释

1. Dawkins, R., *The Blind Watchmaker*, Penguin Books, London, 1986; Dawkins, R., *New Scientist*, 34, 25 September 1986.
2. Schneider, T. D., *Nucleic Acids Res.*, 28:2794, 2000.
3. 在最早期的生物体或新个体中，所有的基因家族是否已经产生和存在？可以在基因组和不相关的物种之间转移的基因因子（如质粒），这可能是一个被设计用于引入多样性的流程，现在的研究可

- 确认这个可能性。但是“所有新的生物功能是从一个早期状态进化而来的”，这并非进化论的一个有说服力的论据。有很多涉及几个基因的过程，这些基因都没有任何其它额外的目的（功能）。一个正在进化的基因需要与其它正在使用的基因进行协作，这就面临一个难题：正在改变的基因如何实现协作呢？特别是所有的改进是随机突变而来的。对正在使用的基因的任何破坏将会让这个生物被自然选择所淘汰。所以在新基因进化的过程中，形成新功能和防止突变所造成的破坏必须能够几乎同时进行。
4. 例如一个动物的细胞，通常含有1000-4000种不同的酶，每种酶催化一个单一的化学反应，或者催化一组紧密相关的化学反应。
 5. Lodish et al., Ref. 4, p. 125.
 6. Küppers, B.-O., The prior probability of the e of life; in: Krüger, L., Gigenrenzer, G. and Mortgan, M.S. (Eds), *The Probabilistic Revolution*, MIT, Press, Cambridge, pp. 355-369, 1987. Discussed by S. Meyer in Ref. 42, p. 80-82.
 7. 蛋白质结构分类 (SCOP) <http://scop.mrc-lmb.cam.ac.uk/scop/> 结构域是蛋白质中具有特异结构（和独立功能）的部分，其大小通常为50-100（参见脚注10），或参考文献30，100-200个氨基酸残基。不同的蛋白质都有这些普遍的特征，且并不需要相同的氨基酸残基。一般可以通过研究蛋白质的结构域来推断蛋白质化学意义和生物学作用。我们将蛋白质看作工程大师的杰作，他将相同的通用组件用于多种的用途。那么，在（细胞膜中）发挥功能的蛋白质有一个附着于疏水表面的结构域（适于细胞膜的内部的环境），这就不足为奇了。
 8. （如要实现蛋白质结构域洗牌）必须精确提取编码蛋白质结构域的那部分基因。蛋白质的N个结构域中，每一个结构域都可以拥有一个或多个拷贝。正确地选择N=2个特定结构域，随机构建新基因的概率为 9.9×10^8 分之1（即 $1/(31,474)^2$ ），（假设两个结构域的基因编码区恰好被适当地提取）。然后这两者的基因必须被引入到染色体中的可接受位置，这个位置不能干扰现有基因。结构域N=5，以及更大的蛋白质并不罕见，那就需要（随机）作出 $(31,474)^5 = 3.2 \times 10^{23}$ 个正确的选择，并且这些基因还要被放置在对于彼此都适当的位置。如果这种混乱和随机的遗传扰乱经常发生的话，那么基因组就不能维持很长时间。
 9. Yockey, H.P., *Information Theory and Molecular Biology*, Cambridge University Press, Cambridge, p. 250, 1992.
 10. Doolittle, R.F., *Annu. Rev. Biochem.*, 64:287, 1995.
 11. 认为基因从一个功能中解放出来，然后假设其没有经过一个非功能阶段就产生另一个新基因是不合理的。复杂的生物学功能并不是通过不间断的拓扑结构（计算机网络的连接）连接起来的。
- 在参考文献1，道金斯教授简单地假设有一个不间断的突变链从一个随机状态形成一个复杂的器官，如眼睛，每个形成的器官在自然选择中都具有优势。
12. Yockey, Ref. 9, p. 135
 13. Yockey, Ref. 9, p. 136
 14. Borstnik, B. and Hofacker, G.L.; in: Clementi, E., Corongiu, G., Sarma M.H. and Sarma, R.H. (Eds), *Structure & Motion, Nucleic Acids & Proteins*, Guilderland, Adenine Press, New York 1985.

15. Borstnik, B., Pumpernik, D. and Hofacker, G.L., Point mutations as an optimal search process in biological evolution, *J. Theoretical Biology* 125, 249–268, 1987.
16. Yockey, Ref. 9, p. 254.
17. Fisher, R.A., *The Genetical Theory of Natural Selection*, 2nd revised edition, Oxford University Press, Oxford, New York, Dover, 1958.
18. Qb 寄生虫基因组使用 4,500 个核苷酸的单链 RNA。Sol Spiegelman (伊利诺伊大学) 将 Qb, 必需的复制酶和某些盐放置在试管中。在这种环境中其基因组非必要的部分丢失了, 使它们的后代能更快地繁殖。在大约 70 代之后, 所有非必要的基因都被去除了, 并且留下了一个具有 550 个核苷酸长的 RNA 单链的物种。在第二个实验 (p.160) 中, 已经最佳地缩短的 RNA 放在能减慢复制的药物环境中进行复制, 该药物附着于特定的三个核苷酸序列。几代之后, 所有成员都拥有相同的新序列, 其中药物的所识别和附着的那 3 个核苷酸突变为其他核苷酸, 这就阻止了药物的结合并允许更快的繁殖。
19. Lodish et al., Ref. 4, p. 187. 遗传上不同的两个细胞的融合产生了一种称为异核体的杂交细胞。随着人-鼠杂交细胞的生长和分裂, 它们逐渐随机地失去了人类染色体, 但保留了小鼠染色体。在能够支持人类细胞和突变小鼠细胞生长的培养基中, 人-鼠杂交细胞最终会失去所有的人类染色体。然而, 在缺乏小鼠细胞必需 (小鼠细胞不能产生的) 代谢物的培养基中, 那一条含有编码所需酶基因的人类染色体将被保留, 因为在有丝分裂后任何失去它的杂交细胞都会死亡。
20. Beaton, M.J. and Cavalier-Smith, T., *Proceedings of the Royal Society (B)*, 2656:2053, 1999. Cryptomonads can remove functionless DNA (pointed out in Ref. 20, p. 8).
21. Petrov et al., *Molecular Biology Evolution*, 15:1592, 1998. Drosophila can remove functionless DNA (pointed out in Ref. 22, p. 8).
22. Darrall, N.M., *Origins: The Journal of the Biblical Creation Society*, 29:2, 2000.
23. Lodish et al., Ref. 4, p. 63: 尽管蛋白质浓度很高 (~100 mg/ml), 但细胞内 95% 以上的蛋白质仍能维持其天然构象, 而在体外这种高浓度条件下通常会导致蛋白质沉淀。
24. Hoyle, Fred, *Mathematics of Evolution*, Acorn Enterprises LLC, Memphis, 1999. Use equation (1.6) on p. 11.
25. For t in equation (9) use: $(6 \text{ generations/hour}) \times (24 \text{ hours/day}) \times (1095 \text{ days/3 years}) = 157680 \text{ generations in 3 years.}$
26. Lodish et al., Ref. 4, p. 456.
27. 除了启动子位点之外, 基因表达的调节和 mRNA 的修饰可能需要多个不编码蛋白质的碱基对序列 (非编码区)。Lodish et al., Ref. 4, p. 295: ‘真核基因中的其他关键的非编码区是指定 3' 端分离的序列和多聚腺苷酸化 [多聚 (A) 位点], 以及初级 RNA 被剪接去除的序列 (内含子)。413 页;’ 几乎所有 mRNA 在多聚 (A) 尾巴的上游的 10-35 个核苷酸中都含有 AAUAAA 的序列。
28. Pennisi, Ref. 42, p. 346: ‘Fitsch & Markowitz (1970) have shown that as the taxonomic group is restricted the number of invariant positions increases.’
29. Yockey, Ref. 9 p. 315.
30. Lodish, Ref. 5, p. 60: ‘A structural domain consists of 100–200 residues in various combinations of a helices, b sheets, and random coils.’
31. Lodish et al., Ref. 4, p. 405. 色氨酸操纵子编码了逐步合成色氨酸所必需的酶。单个操纵子产生的 mRNA 链被翻译成所需的 5 种蛋白质 (E, D, C, B, A)。蛋白质 E 和 D 形成生物合成过程中所需的第一种酶; 蛋白质 C 催化中间体; 蛋白质 B 和 A 结合形成所需的最后一种酶。“如上面例子, 细菌基因组中 (编码蛋白质) 基因的顺序与色氨酸合成过程中所需蛋白质的功能顺序相似。” Ref. 5, p.144.
32. 实际上, 可能性会更低。例如: 获得概率为 0.5 的两个事件的几率为: $0.5 \times 0.5 = 0.25$, 而要获得概率为 0.25 和 0.75 事件的概率更小: $0.25 \times 0.75 = 0.19$ 。要获得概率为 0.01 和 0.99 事件的概率只有 0.009, 即使上述所有事件的概率“平均值”都为 0.5。当被告知一种大蛋白质在其大多数位置具有很大的自由度时, 人们可能会严重高估 (随机形成蛋白质的) 概率。蛋白质中一些不能改变的位点是有重要作用的。检查 Yockey 的数据 (Ref. 9, Table 9.1) 显示, 平均每个位置可以容忍 (替换) 10.9 个氨基酸 (在 20 个可能的氨基酸中)。因此, 这里没有研究过异常限制性或“保守”的蛋白质。然而, 110 个氨基酸位置中的 16 个 (14.5%) 必须是某一个特定的氨基酸。暂不考虑遗传密码有利于某些氨基酸的事实, 获得所有 16 个不可改变氨基酸的位置的几率大致为 $(1/20)^{16} = 1.5 \times 10^{-21}$ 。平均蛋白质比细胞色素 c 约大 3 倍, 具有 330 个氨基酸残基, 相对大约有 14.5% 是不可改变氨基酸的位点, 那么允许任何氨基酸的位点可以有 $(330-48) = 282$ 个, 这显然是一个要求非常低的蛋白质。然而, 通过随机突变形成 48 个不变位点的氨基酸的概率大约为 $(1/20)^{48} = 3.7 \times 10^{-63}$ 。这概率大大低于 Yockey 计算的随机产生整个细胞色素 c 蛋白的可能性。
33. Kirschner, M. and Gerhart, J., *Proc. Natl. Acad. Sci. USA*, 96:8420, 1998.
34. Pennisi, E., *Science*, 272:1098, 1996; Mushegian, A. and Koonin, E., *Proc. Nat. Acad. Sci. USA*, 93:10268, 1996; Bult, C. et al., *Science*, 273:1058, 1996. Discussed in Fisher, Ref. 18, p. 76. 退休的创造论教授 Roland Süßmuth 博士是支原体的专家, 在第 18 届创造生物学大会中就此主题发表了一篇论文: Süßmuth, R., *Die Bakteriengruppe der Mycoplasma*, Tagesband der 18. Fachtagung für Biologie, p. 69, 16–18 März 2001. Studiengemeinschaft Wort und Wissen e.V. Rosenbergweg 29, D-72270, Baidersbronn, Germany.
35. 然而, 大多数蛋白质家族只包括几个到 30 个左右的成员; 常见的例子有细胞骨架蛋白, 70-kDa 热休克蛋白, 肌球蛋白重链, 鸡卵清蛋白, 以及脊椎动物的 a-和 b-球蛋白。来自与顺序无关的基因的蛋白质通常共同协作发挥生物学功能。例如, Ref. 5, p. 300 “几种不同的基因家族编码的各种蛋白质构成了细胞骨架。”
36. “这种基因敲除方法已被用于分析酵母染色体 III。对 DNA 序列的分析表明, 该染色体含有 182 个开放阅读框, 其长度足以编码长于 100 个氨基酸的蛋白质, 在该分析中假设其为自然存在的蛋白质的最小长度 (着重强调)。
37. 一个开放阅读框 (ORF) 是可以分成三联密码子而没有任何干扰终止密码子的 DNA 序列。虽然一些多肽的长度短于 100 个氨基酸, 但这些多肽难以只通过 DNA 的序列来预测, 因为短的 ORF 在长 DNA 序列中是随机出现的。编码 100 个或更多氨基酸的长的 ORF 不太可能随机出现, 且很可能编码一个能表达的多肽 (蛋白质)。
38. Alberts, B., *Cell*, 92:291, 1998.
39. Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A., *Science*, 289:905, 2000.
40. Bergman, J., *CRSQ*, 36 (1):2, 1999.
41. Frank, J. and Agrawal, R.K., *Nature*, 406:318, 2000.
42. Behe, M.J., Dembski, W.A. and Meyer, S.C., *Science and Evidence for Design in the Universe*, Ignatius, San Francisco, 2000.
43. Behe, M.J., *Darwin's Black Box: The Biochemical Challenge to Evolution*, Touchstone, New York, 1996.
44. Reidhaar-Olson, J. and Sauer, R., *Proteins, Structure, Function and Genetics*, 7:306, 1990; Bowie, J. and Sauer, R., *Proc. Nat. Acad. Sci. USA* 86:2152, 1989; Bowie, J., Reidhaar-Olson, J., Lim, W. and Sauer, R., *Science*, 247:1306, 1990; Behe, M., *Experimental support for regarding functional classes of proteins to be highly isolated from each other; in: Buell J. and Hearn, G. (Eds), Darwinism: Science of Philosophy?* Houghton Publishers, Dallas, pp. 60–71, 1994; discussed in Ref. 42, p. 75.
45. Axe, D., Foster, N. and Ferst, A., *Proc. Nat. Acad. Sci. USA* 93:5590, 1996. Discussed in Ref. 42, p. 75.
46. Lodish et al., Ref. 4, p. 62: ‘Any polypeptide chain containing n residues could, in principle, fold into 8ⁿ conformations.’

罗尤·杜鲁门 (ROYAL TRUMAN)

纽约州立大学的化学学士和计算机科学学士学位、密西根大学 MBA、密西根州立大学有机化学博士。就职于德国巴斯夫 (BASF)。

迈克尔·海西希 (MICHAEL HEISIG)

在卡尔斯鲁厄大学通过食品科学的国家测试 (同等于科学硕士), 弗莱堡大学分子生物学博士, 就职于德国海德堡大学。

Appendix 1.

The size of the polypeptide search space.

The number of possible sequences using n amino acids is given by $20 \times 20 \times 20 \dots n$ times. For the subset of $n=110$ residues of cytochrome c this indicates

$$(20)^{110} = 1.3 \times 10^{143} \quad (11)$$

candidate amino acid sequences.

Now, the proportion of each amino acid used by proteins varies significantly. Yockey¹ showed that to a good approximation the number of synonymous codons allows a good estimate for the frequency each amino acid is used by proteins, except for arginine (Table 1). Then the probability, p_j , of finding an amino acid at a specific position, j , is affected by an existing genetic code and to a first approximation,

$$p_j = r_j p_i \quad (12)$$

where r_j represents the number of codons (between 1 and 6) coding for amino acid j and $p_i = 1/61$, the codon probability, as explained in Table 1.

Table 1. Calculated and experimental amino acid frequencies, p_i (King & Jukes analyzed 5492 amino acid residues from 53 vertebrates).^(a)

Residue	Probability, P_i	Calculated Value	King & Jukes	Goel <i>et al.</i>
leu	p_1	$6/61 = 0.0984$	0.076	0.0809
ser	p_2	$6/61 = 0.0984$	0.081	0.0750
arg	p_3	$6/61 = 0.0984$ ^(b)	0.042	0.0419
ala	p_4	$4/61 = 0.0656$	0.074	0.0845
val	p_5	$4/61 = 0.0656$	0.068	0.0688
pro	p_6	$4/61 = 0.0656$	0.050	0.0494
thr	p_7	$4/61 = 0.0656$	0.062	0.0634
gly	p_8	$4/61 = 0.0656$	0.074	0.0748
ileu	p_9	$3/61 = 0.0492$	0.038	0.0458
term	p_{10}	$0 = 0$ ^(c)	0	0
tyr	p_{11}	$2/61 = 0.0328$	0.033	0.0345
his	p_{12}	$2/61 = 0.0328$	0.029	0.0222
gln	p_{13}	$2/61 = 0.0328$	0.037	0.0413
asn	p_{14}	$2/61 = 0.0328$	0.044	0.0535
lys	p_{15}	$2/61 = 0.0328$	0.072	0.0605
asp	p_{16}	$2/61 = 0.0328$	0.059	0.0555
glu	p_{17}	$2/61 = 0.0328$	0.058	0.0538
cys	p_{18}	$2/61 = 0.0328$	0.033	0.0230
phe	p_{19}	$2/61 = 0.0328$	0.040	0.0402
trp	p_{20}	$1/61 = 0.0164$	0.013	0.0153
Met	p_{21}	$1/61 = 0.0164$	0.018	0.0155

(a) Table from.¹ 3 codons are used for the terminator, leaving $64 - 3 = 61$ codons to distribute among the other amino acids.
 (b) In nature arginine (arg) is coded almost exclusively by 2 (AGA and AGG) of the 6 codons available.
 (c) 3 codons are used in the Universal Genetic Code.

Polypeptides composed mostly of amino acids of low occurrence are very unlikely to exist. The odds of obtaining a polypeptide $n=110$ residues long based on only residues represented by 1 codon (such as *trp* and *met*, with chances of ca. $1/61$) compared to one based on only residues represented by 6 codons (such as *leu* and *ser*, with odds of $6/61$), is:

$$\frac{(0.0164)^{110}}{(0.0984)^{110}} = 2.5 \times 10^{-86} \quad (13)$$

For n greater than 110 residues this proportion drops rapidly.²

Treating the genetic code as given, it appears that the search space given by (11), in the absence of intelligent guidance, is exaggerated since many worthless, but highly improbable, candidates would not be tested by chance. One can define two collections of sequences, one consisting of those polypeptides which as a collection possess negligible chance of being generated compared to the second, higher probability set.

We avail ourselves of some mathematics developed by Shannon for telecommunication purposes and applied by Yockey to the analysis of gene and protein sequences.³

The entropy, H , for each residue position of a protein can be calculated by:

$$H = -\sum_{j=1}^{20} p_j \log_2 p_j \quad (14)$$

which gives $H = 4.139$ 'bits' using p_j from (12).

The number of different polypeptides using n amino acids, neglecting the set of those belonging to the very low probability class, is given by

$$a^{nH} \quad (15)$$

where $a = 2$, if we choose to work with base 2 logarithms, which is mathematically convenient. This reduces the potential search space suggested by (11) to

$$2^{(4.139 \times 110)} = 1.15 \times 10^{137} \quad (16)$$

candidate polypeptides of length 110 amino acids. Were the probability of obtaining any amino acid identical, meaning $1/20$ for every position, then equations (14) and (15) would predict the same number of candidate sequences, $(20)^{110}$, as found in (11).

The set of functional cytochrome c sequences.

We restrict ourselves now to single protein family, cytochrome c. The entropy, H_l , of the probability distribution of the synonymous residues at any site l is given by

$$H_l = \sum_j p'_{lj} \log_2 p'_{lj} \quad (17)$$

where

$$p'_{lj} = \frac{p_j}{\sum p_j} \quad (18)$$

The summation in (18) includes only the synonymous residues, based on available sequence data, at position l on the polypeptide. p_j was defined in (12). The effective number of synonymous residues at each site l is calculated as:

$$2^{H_l} = N_{\text{eff}}^l$$

where H_l is defined in (17). Finally, multiplying these values for all 110 sites provides the number of known functional cytochrome c variants:

$$\text{Functional sequences} = \prod_{l=1}^{110} N_{\text{eff}}^l \quad (19)$$

It is possible additional sequences will be discovered, may have gone extinct, or would be functional but have not been produced by mutations. This potential was estimated using a prescription developed by Borstnik and Hofacker.¹⁴ 20 amino acid physical properties were used, from which 3 orthogonal eigenvectors were sufficient to describe the data adequately. This differs from an earlier approach^{1,3,4} which was based on Grantham's⁵ prescription.

The estimated functional sequences reported⁶ are

$$2.316 \times 10^{93} \quad (20)$$

Appendix 2.

Alternative calculations of probability to obtain a functional cytochrome c protein

The amino acids presumed to be tolerated at each position on the protein is used, along with the probability of generating the acceptable amino acid (based on synonymous codons from the universal genetic code). Then it is straightforward to calculate the odds of finding an acceptable residue for each position. This is provided in the column labelled $\sum p_i$ at the bottom of Table 2. Multiplying the individual probabilities leads to an overall probability of

$$2.71 \times 10^{-44} \quad (2)$$

of obtaining a minimally functional cytochrome c protein.

This simplification could be justified by that fact that generally the proportion of a particular amino acid in proteins **does** parallel fairly well the number of codons assigned to it, see Table 1. Yockey's more rigorous mathematics, which removes from consideration polypeptide sequences of very low probability, leads to an estimated of 2.0×10^{-44} , see (1).

A simpler alternative ignores the number of synonymous codons used by the universal genetic code and considers only the number of acceptable amino acids at each protein site. This leads to an estimate (last column, bottom of Table 2) of

$$6.88 \times 10^{-45} \quad (3)$$

As a back-of-the-envelope estimate this later approach is useful as a rough orientation. It has relevance for abiogenesis scenarios when: only the 20 amino acids are present; in relative proportions reflecting usage in proteins; in the absence of interfering reactants, including water; only L form amino acids are present; no chemical side-reactions occurs (such as intramolecular rings, condensation of side chain carboxylic acids, oxidation reactions, etc.).

Appendix 3

Trials available to chance.

How many opportunities might chance have to stumble on a functional cytochrome c sequence? We permit all random mutational processes able to generate a new sequence on a suitable portion of DNA not needed for other purposes. Let us assume some generous settings from an evolutionary perspective to avoid argument:

Table 2. Effective number of amino acids for iso-1-cytochrome c^(g). Explanation of the data from Yockey¹⁷ based on data from Hampsey, Das & Sherman^{18,19} derived from 92 eukaryotic cytochromes c.²⁰

Site Nr ^(b)	p _i ^(a)	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Tyr	Trp	Thr	Val
		4/61	6/61	2/61	2/61	2/61	2/61	2/61	4/61	2/61	3/61	6/61	2/61	1/61	2/61	4/61	6/61	2/61	1/61	4/61	4/61
23		1		1	<i>I</i>	<i>I</i>	1	1	1		<i>I</i>	1	<i>I</i>			<i>I</i>	1		<i>I</i>	<i>I</i>	<i>I</i>
36		1		<i>I</i>	<i>I</i>	<i>I</i>	1	1	<i>I</i>		1	<i>I</i>	<i>I</i>			<i>I</i>	<i>I</i>		1	1	1
...																					
34			<i>I</i>	1	<i>I</i>		1	<i>I</i>		1			<i>I</i>				1				
38																1					
40												1									
...																					
17		1	<i>I</i>	1	<i>I</i>	<i>I</i>	<i>I</i>	<i>I</i>	<i>I</i>	1	1	1	<i>I</i>		1		<i>I</i>		<i>I</i>	1	1
...																					
35				<i>I</i>		<i>I</i>	1	<i>I</i>	1			1	1				1	1	1	1	

Site Nr	∑p _j ^(e)	H ₁ ^(d)	2 ^{H₁} = N ₁ ^{eff}	AAs ^(e)	Prob. ^(f)
23	0.787	3.7563	13.513	15	0.75
36	0.787	3.7563	13.513	15	0.75
...					
34	0.328	2.8464	7.192	8	0.4
38	0.066	0	1	1	0.05
40	0.098	0	1	1	0.05
...					
17	0.885	3.9163	15.098	17	0.85
...					
35	0.541	3.2559	9.553	11	0.55
...					
Prob.:	2.71×10⁻⁴⁴	310.1508 bits		Prob.:	6.88×10⁻⁴⁵

(a) p_j = r_j/61 where r_j is the number of codons (1 to 6) coding for amino acid j
 (b) 1 means residue is known at that position; 1 (in italics) means residue is predicted to be allowed at that position
 (c) ∑ of probabilities, using known and postulated amino acids at that position
 (d) H₁ = - ∑ (p_j) log₂ (p_j), where p_j = p_j/∑ (p_j)
 (e) Number of amino acids (AAs) assumed to be tolerated at that position of the protein
 (f) Probability of getting a tolerated AA by chance
 (g) The calculations²¹ were checked with an Excel spreadsheet. Some typographical errors appear in Dr Yockey's book. For example, for N₁^{eff} residue position #17: the reported was 15.908, the correct value is 15.098; residue position #80: reported was 6.420, correct is 6.240. In private correspondence, Yockey confirmed that in the text several residues were accidentally left out, but had been taken into account for the final calculations. To complete his table 9.1, transfer the following residues from his Table 6.2: 5, 43, 60; 67, 73, 75, 82, 91, 102, 103 and 105. Our careful calculations, using an Excel spreadsheet, confirm almost exactly the reported value of H₁ = 310 bits (our value is slightly higher).

$$5 \times 10^{13} \text{ generations} \tag{21}$$

based on 1 generation per 10 minutes on average for 1 billion years. For comparison, 'In ideal growth conditions, the bacterial cell cycle is repeated every 30 minutes.'⁷ Our proposed value is surely about a factor of 10 too generous on average.

$$4 \times 10^{20} \text{ liters living space} \tag{22}$$

based on an assumed primitive ocean of volume 20,000 × 20,000 × 1 km. The current oceans are believed to contain about **1.4 × 10²⁰** liters of water:⁸

Since the ancient earth framework assumes water accumulated from comets and water vapour from volcano eruptions over billions of years, the putative aqueous living space would actually have been on the order of only 1/10th the volume we are using.

One evolutionist would like to enlighten us⁹ that such a primitive ocean would have contained **1 × 10²⁴ liters** of water, conveniently stocked full of just the right 20 amino acids. Which is more probable, his claim or (22)? Post French Revolution the circumference of the earth was measured at 40,000 km. 1 thousandth of a km became the definition of a metre and a cube of 1/10th of meter full of water became the definition of a kilogramm. A liter is a cube having length 1/10th of meter.

If the earth were a perfect sphere the measured circumference would indicate a radius of about 6,366.2 km. Using the accepted radius⁵⁴ of **6,378.15 km** and assuming a perfect sphere provides an estimate of the earth's **total** volume:

(4/3) πr³ = (4/3) π (6378.15 × 10⁴ dcm)³ = 1 × 10²⁴ liters. For the evolutionist's statement⁵⁵ to be true the **whole** earth would have to consist of water, clearly absurd.

As an alternative calculation to see whether (22) is reasonable let us assume there were no continents and the ocean had an average of 1 km depth. The outer radius r₁ is **6378.15 km** and the inner radius r₂ is (**6378.15 - 1**) km This would provided a volume of water:

$$(4/3) \pi r_1^3 - (4/3) \pi r_2^3 = 1.2 \times 10^{20} \text{ litres.}$$

This confirms that (22) has been deliberately exaggerated to favour the evolutionary scenario.

1×10¹¹ members per liter (23)

based on 10% of the levels available for concentrated E. coli under optimal laboratory conditions.¹⁰ We assume sufficient nutrients are available in nature during the billion years and that this high concentration was maintained from water surface to a depth of 1 km. On average over a billion years this is probably **at least** 100 times too generous.

Note that the maximum number of organisms thus estimated agrees almost precisely with other work performed independently.¹¹

1 novel mutation per cytochrome C protein per 1000 generations (24)

Estimates of error rates during DNA duplication vary. Yockey¹² suggested between 10⁻⁷ and 10⁻¹² per nucleotide. Other literature indicate between 10⁻⁷ and 10⁻¹⁰.^{13,14,15,16} Using the fastest mutation rate proposed in the literature above indicates about 3.3 × 10⁻⁵ base pair changes per generation, based on:

(330 bases per cytochrome c gene) (10⁻⁷ mutations per base per generation) .

Let us by generous and use a mutation rate of 1 per 1000 generations, which is about 30 times greater than the fastest estimate proposed to avoid argument. Furthermore we will neglect the effect such random mutations would have on the rest of the genome.

These assumptions offer a generous maximum number of attempts possible:

$$(5 \times 10^{13}) \times (4 \times 10^{20}) \times (1 \times 10^{11}) \times (1 \times 10^{-3}) = 2 \times 10^{42} \quad (25)$$

mutational opportunities.

References and Notes

1. Yockey, H.P., *J. Theor. Biol.* **46**: 369, 1974; p. 381.
2. One sees how quickly the proportion drops with increasing polypeptide length, n, by taking the log of the ratio: proportion = 10^{n[log0.0164 - log 0.0984]} = 10^{-0.778n}
3. Yockey, H.P., *J. Theor. Biol.* **67**:377, 1977.
4. Yockey, H.P., *J. Theor. Biol.*, **67**:345, 1977; see p. 361.
5. Grantham, R., *Science*, **185**:862, 1974.
6. Pennisi, Ref. 9, p.254
7. Lodish *et al.*, Ref. 5, p. 9.
8. 'Earth', <http://seds.lpl.arizona.edu/nineplanets/nineplanets/earth.html>.
9. Musgrave, I., Lies, damned lies, statistics, and probability of abiogenesis calculations, <http://www.talkorigins.org/faqs/abioprob.html>.
10. Personal communication from Professor Scott Minnich, email from 17 March 2001.
11. Scherer, S. and Loewe, L., Evolution als Schöpfung? in: Weingartner, P. (Ed.) , *Ein Streitgespräch zwischen Philosophen, Theologen und Naturwissenschaftlern*, Verlag W. Kohlhammer, Stuttgart; Berlin; Köln: Köhlhammer, pp. 160–186, 2001.
The authors were made aware of the above essay just prior to the submission of this paper. It contains many valuable probability calculations performed independently and unknown to us. It is reassuring that although they worked with different assumptions, virtually identical numbers resulted. For example, they calculated the maximum number of cells which could have lived in **4 billion years** as **10⁴⁶** whereas we estimated for **1 billion years, 2 × 10⁴⁵** :
From Appendix 3: (5 × 10¹³) (4 × 10²⁰) (1 × 10¹¹) = 2 × 10⁴⁵.
12. Yockey, Ref. 9, p. 301.
13. In bacteria the mutation rate per nucleotide has been estimated to be between 0.1 and 10 per billion transcriptions.^{14,15} For organisms other than bacteria, the mutation rate is between 0.01 and 1 per billion.¹⁶ References found in: Spetner, L., *Not by Chance! Shattering the Modern Theory of Evolution*, The Judaica Press, Inc., Chapter 4, 1998.
14. Fersht, A.R., *Proceedings of the Royal Society (London)* , **B 212**, 351–379, 1981.
15. Drake, J.W., *Annual Reviews of Genetics*, **25**, 125–146, 1991.
16. Grosse, F., Krauss, G., Knill-Jones, J.W. and Fersht, A.R., *Advances in Experimental Medicine and Biology*, **179**:535–540, 1984.
17. Yockey, Ref. 9, p. 250.
18. Hampsey, D.M., Das, G. and Sherman, F., *J. Biological Chemistry* **261**:3259, 1986.
19. Hampsey, D.M., Das, G. and Sherman, F., *FEBS Letters* **231**:275, 1988.
20. Yockey, Ref. 9, p. 162.
21. Yockey, Ref. 9, p. 250

《创造研究》现已面世，欢迎订阅!

现今主流的学术界已经将未经证实的进化论假说视为事实，并且掩盖很多反对进化假说的事实和发现。因而很多人只能接触到经过“筛选”的信息，以为进化论大部分的问题已经解决了，只是个别少量的问题还没有完满的解释，就将进化看作已被证实的真理，而将创世记当为神话。

但真实的情况是随着现代科学研究越来越深入，覆盖面越来越广泛，进化论在自然科学各个领域面临的挑战是越来越严重，而且发现越来越多证据支持创世记是真实记述。

为了让知识背景较强的读者能更深入、更全面和更详细地明白进化论的问题和创造论最新的科学研究数据和成果，我们精选了创造论科学家近期发表的论文，现已出版《创造研究》（**免费电子版期刊**），为创造科学领域的科研人士提供展现和交流的平台。有意免费订阅《创造研究》请发邮件至以下邮箱：

chuangzaoyanjiu2@protonmail.com

