

# 中国科学家创新DNA存储算法 让敦煌壁画再活两万年

## 中国科学家在新冠病毒中发现RNA加帽新机制

中青报 中青网记者 胡春艳 通讯员 赵晖

科幻大片《侏罗纪公园》里讲述了这样一个故事：科学家找到一块有史前蚊子的琥珀，从蚊子血中获得了恐龙的基因，从而让已灭绝了6000多万年的恐龙复活。

恐龙的生物信息存储在DNA中，若干年后被提取并还原出来。这听起来似乎有些道理，却也让人倒吸一口凉气。

最近，天津大学一项研究成果让人们离想象又近了一些。该校合成生物学团队将10幅精选敦煌壁画存入DNA中，并通过加速老化等实验，发现这些壁画信息在常温下可保存千年，在9.4℃下可保存两万年。

如果在合适的温度等条件下，保存千万年也是可以的。中国科学院院士、天津大学副校长元英进说。

### 小小的DNA却拥有惊人的存储容量

人类文明进化史，也是一部信息存储技术发展史。

从结绳记事、仓颉造字到磁带、硬盘等现代磁光存储技术，数据存储帮助人类延续了思想，记录下灿烂文明。造纸与印刷术的发明，让人类能够存储的数据量在几百年内获得了大约5个数量级的提升。到了计算机时代，人类产生的数据呈爆发式增长。

全世界都在建数据中心，而数据中心的能耗是惊人的。元英进说。人们一直在不断寻找更海量、更稳定、更安全的存储方式。

大自然鬼斧神工的绝妙之处就在于此。最好的存储器或许就藏身于生命体之中。

自地球上出现生命以来，大自然一直用DNA来存储信息，至今已有30多亿年。人类的五官在脸上如何摆放，体内的蛋白怎样合成，眼睛是什么颜色，诸如此类纷繁复杂的人类基因组信息，都记录在比细胞还小得多的DNA上，一代代沿用至今。

不同于各种人造存储设备，DNA极其精巧却又如此经久耐用，它存储了亿万年来无数生物的遗传信息，造就生命繁衍、进化演化及生物多样性。

那么，假如把海量的信息，像存入U盘、硬盘一样，写到小小的DNA上，岂不是一举多得？事实上，当人类发现DNA的双螺旋结构后，美俄科学家就先后提出了用DNA存储数字信息的概念。

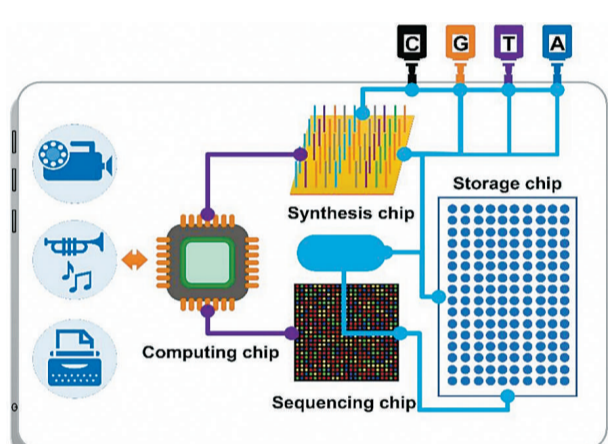
元英进解释说，DNA存储相较于磁、光、电等常规的信息存储介质有3个最显著的优势。其中最大的优势在于存储密度高。目前，天津大学研究团队将部分经典视频片段存储在DNA中，已实现了体积存储密度比普通硬盘高出6个数量级。

与此同时，存储的信息可用时间非常长。此次研究者将10幅敦煌壁画信息存储在DNA中，结合创新的算法，可以实现DNA分子在室温下保存超过千年，在9.4℃条件下保存两万年。



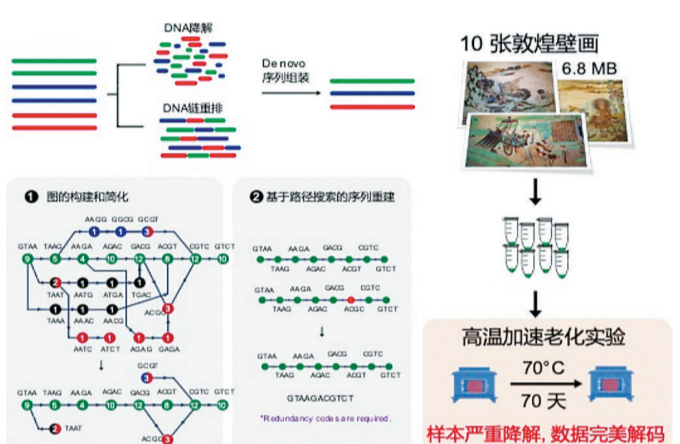
DNA存储的敦煌壁画。

天津大学供图



DNA存储技术概念图。

天津大学供图



基于德布莱英图理论设计的序列重建算法高效解决DNA断裂、降解问题。

天津大学供图

这样的长期保存需要的能耗却很低。元英进认为，DNA存储被视为一种极具潜力的存储技术，已经成为应对数据存储增长挑战的新机遇。

### 壁画变身DNA需要几步

DNA信息存储的原理共分两步：信息写入和信息读取。

这个过程实际上跨越了极难逾越的鸿沟：它打破了有机与无机的界限，连起生命和信息两大系统。

DNA是脱氧核糖核酸的缩写，含有A T C G四种碱基。如果用数字中的0、1、2、3分别代表一个碱基，就组成了一个四进制的存储方式，类似于

计算机采用的0和1二进制代码。

通过编码转化，碱基四进制和计算机二进制就可以实现对话。天津大学合成生物学前沿科学中心博士生韩明哲解释说，壁画的数字图像本质上就是二进制的比特串，我们通过编码将这些二进制的比特串，转化为四进制的ATGC碱基序列，再通过DNA合成技术将碱基序列写入DNA中，壁画的数据图像就变为DNA了。

此前，该团队成功在酿酒酵母中合成了一条额外的人工染色体，并在上面存储了两张图片及一段视频信息，将其称之为酵母CD。随着酵母的不断繁殖扩增，数字信息也随之廉价且稳定地复制。

我们代代培养酵母到100代，依然

可以完美地恢复出原始数据。元英进说，假如脑洞更大一点，将信息存储到一棵树中，随着树生长千百年，人类的子孙后代都可以随时从这棵树中读取到千百年前存储的信息。

这一次，这支年轻团队的创新之处在于，能实现更恶劣条件下可靠读取信息。韩明哲说，存了壁画信息的DNA，本质上其实跟天然的DNA没有什么不同，同样也存在在长时间存放而产生的断裂和降解等问题，影响信息存储的长期可靠性，这也成为亟待解决的关键科学问题。

于是，他们设计了基于德布莱英图理论的序列重建算法来解决DNA断裂等问题，可以从严重降解的DNA样本中，恢

复原始的信息。

为了验证数据的长期可靠性，团队制备了一个没有任何特殊保护的DNA水溶液样本，随后在70℃的温度下加速样本断裂、降解长达十周。韩明哲说：这个过程使得DNA片段80%以上都发生了断裂错误，模拟了DNA在自然环境下千上万年的降解情形。

随后，团队依靠设计的序列重建算法，依然可以准确组装并解码96.4%以上的片段，再通过一种编码方式解决了少量片段丢失的问题，使原始的敦煌壁画图片能够完美恢复。

### DNA存储走向实用化还有多远

尽管DNA存储还不被大众所熟知，但它正在努力走出实验室，距离实用化并不遥远。元英进说，惊人的数据存储需求是新技术走向市场的最大推动力。

据国际数据公司估计，到2025年全球数据总量将达到175ZB（1ZB为十万亿字节）。到2024年，全球将有30%的数字业务进行DNA存储试验。然而从目前来看，DNA存储想要大规模应用，尤其是在中国实用化还需要突破几个关键瓶颈。

团队分析了当前DNA信息存储面临的主要挑战。信息存储成本高、信息读写速度慢，以及无法高效对接现有信息系统是三大主要限制因素。

根据测算，目前DNA存储写入成本相当于20世纪80年代内存的存储成本，而要达到当前数据存储成本还需要降低7-8个数量级。

DNA信息存储成本在未来有很大下降的潜力。韩明哲认为，今后可以从优化合成反应、改良芯片结构、替换廉价耗材、优化试剂配方量等方面着手，大幅降低合成成本。

与此同时，由于信息存储领域市场规模巨大，随着半导体器件、微纳加工在DNA信息存储领域的应用，该领域的巨大投入将对DNA合成技术产生重大影响，DNA合成技术与装备快速迭代升级，也有望使成本快速下降。

DNA信息存储的读取依赖测序技术，与磁、光、电等存储相比，读取速度较慢。目前DNA测序仪的读取速度与硬盘相比，还存在3-4个数量级的差距。现有电、磁存储技术通常每秒可读取几十到几百兆字节数据。此外，DNA存储的标准尚待建立，面临与现有数字存储系统兼容的问题。

DNA信息存储是一个新兴的、多学科深度交叉融合的研究方向。元英进认为，DNA存储在将来极有可能成为庞大冷数据存储的主要存储介质。所谓冷数据，就如同档案馆的历史资料，需要把海量信息保存好，但平时又很少去使用。因为这些数据需要长期存储、耗能又大，而电子存储设备的寿命往往只有十年到几十年，并需要不断更新迭代，难以满足冷数据存储的需要。

DNA存储走向实用化仍面临很多挑战。元英进认为，眼下的突破可能还只是冰山一角，技术进步需要十年磨一剑的耐心，还需要一点运气。

# 2022年诺贝尔自然科学奖解读

编者按

7位科学家，领走了今年诺贝尔奖的3个自然科学类奖项。从人类起源，到量子奥秘，再到分子结构。社交平台上的关注度说明，许多人哪怕从事的工作与科学研究毫无关联，也仍然对这些难以理解的前沿领域充满兴趣。

这可能就是人类与生俱来的、对未知事物的好奇心。我们思考自身命运，好奇人类来自何方，好奇是什么让人类与众不同；我们将视线投向微观世界，好奇粒子之间奇妙的超距纠缠；我们好奇分子构建块是如何有效地结合在一起。科学研究让人类掌握规律，让未知成为已知。

诺贝尔奖的意义之一，在于它让更多好奇的人知道，这个时代有哪些杰出发现和杰出人才。在这个庞大而未知的世界面前，人类的探索永远不会止步。

## 探索人类起源：揭示演化之谜

中青报 中青网记者 叶雨婷

北京时间10月3日17时30分许，2022年诺贝尔生理学或医学奖揭晓。瑞典科学家斯万特·帕博(Svante Pääbo)获奖，表彰他在已灭绝古人类基因组和人类进化方面的发现。

帕博发现，在大约7万年前人类离开非洲后，基因从这些现已灭绝的古人类转移到了智人身上。这种古老的基因流向现代人类，在今天具有生理相关性，比如影响人们免疫系统对感染的反应。

不仅如此，帕博建立了一门全新的科学学科——古基因组学。他通过揭示现存人类与灭绝古人类之间的基因差异，为探索是什么使我们成为独一无二的人类提供了基础。

来自诺贝尔奖官方网站的相关资料指出，现代智人大约在30万年前起源于非洲，人类的近亲尼安德特人约40万年前在非洲以外的欧洲和西亚居住，智人和尼安德特人在欧亚大陆共

存了数万年。

但对于人类与已灭绝尼安德特人的关系，人们长期以来了解甚少。关键的线索来自基因组信息。到上世纪90年代末，几乎整个人类基因组都已得到测序。然而，研究现代人与尼安德特人之间的关系，需要对从远古样本中收集的基因组DNA进行测序。

帕博意识到了技术上的难度。随着时间推移，DNA会被化学修饰、降解成片段。数千年后只剩下微量的DNA，并且，残留的DNA会被细菌和当代人类的DNA污染。帕博开始开发研究尼安德特人DNA的方法，这项工作持续了几十年。

2010年，帕博发表了首个尼安德特人基因组序列：尼安德特人和智人最近的共同祖先生活在大约80万年前。帕博和团队分析尼安德特人和来自世界不同地区的现代人类之间的关系。对比分析表明，尼安德特人的DNA序列与起源于欧洲或亚洲的当代人类的DNA序列更相似，而不是非洲。

这意味着，尼安德特人和智人在几千年的共存中进行了杂交。在现代具有欧洲或亚洲血统的人类中，大约1%-4%的基因组来自尼安德特人。

2008年，帕博博士的团队对西伯利亚南部

一块距今4万年前的指骨碎片进行DNA测序。结果引起了轰动，帕博发现了一种以前不为人知的古人类，命名为丹尼索瓦人。与来自世界不同地区的现代人人类的序列比较表明，丹尼索瓦人和智人之间也发生过基因流动。这种关系首先在美拉尼西亚和东南亚其他地区的人群中被发现，那里的个体携带高达6%的丹尼索瓦人DNA。

如今，因为帕博的发现，人们终于可以了解到，这些已经灭绝的亲属们的古老基因序列，仍影响着当今人类。

## 探索微观世界：纠缠也很美丽

中青报 中青网记者 张渺

著名物理学家费曼曾说，世界上没有人真的懂量子力学。量子力学描述的一些现象，看上去甚至有些反常识，不仅让普通人头疼，也让很多物理学家头疼。

北京时间10月4日17时45分，诺贝尔奖官方网站公布，2022年诺贝尔物理学奖得主为法国物理学家阿兰·阿斯佩(Alain Aspect)、美国科学家约翰·克劳泽(John F. Clauser)和奥地利科学家安东·塞林格(Anton Zeilinger)。他们用光子纠缠实验确认贝尔不等式在量子世界不成立，并开创了量子信息科学。

三位科学家都使用纠缠量子态进行了突破性的实验。在量子纠缠中，两个粒子即使在分离时，也像整体一样。他们的成果为基于量子信息的新技术扫清了道路。这是诺贝尔奖官方网站对三位科学家科研成果的描述。

事情可以从1935年说起，当时，爱因斯坦、波多尔斯基和罗森联合发表了一篇文章，提出量子力学还不完善，应该还有尚未发现的理论，即某种隐变量，可以完整解释物理系统所有可

观测量的演化行为，以避免任何不确定性或随机性。三位物理学家设计了一个有关量子纠缠的思想实验来论证这一观点，这篇论文的论述后来被称为EPR悖论(三人姓名首字母缩写)。

20世纪60年代，物理学家约翰·贝尔基于EPR悖论，提出了贝尔不等式并设计了一个思想实验，使EPR悖论具备了被实验验证的可能性。论文发表之后的几十年，物理学者想出很多种实验来检验贝尔不等式，其中就包括如今的三位获奖者。

想要验证贝尔不等式需要很高的实验条件，因此直到20世纪70年代，验证工作才陆续开始。美国理论和实验物理学家约翰·克劳泽首先进行了尝试，然而他的实验仍然存在一些漏洞。1982年，阿兰·阿斯佩第一次真正意义上验证了贝尔不等式。实验结果违背贝尔不等式，这说明隐变量不存在，量子力学是非确定性的。

1997年，由安东·塞林格所领导的因斯布鲁克大学研究组，首次完成了量子隐形传态的原理性实验验证。次年，他们让光子飞出相距400米，这一实验使得将量子态从一个粒子转移到远处的一个粒子成为可能。

人类对未知的探索永无止境，量子力学充满了难以言喻的美丽。如今，许多基于量子力学的研究成果，都已经应用领域发挥一定作用，比如正在快速发展的量子计算机和量子加密通信技术。

越来越清楚的是，一种新的量子技术正在兴起。我们可以看到，获奖者在纠缠态方面的工作极其重要，甚至涵盖了对于量子力学基本问题的解释。诺贝尔物理学委员会主席安德斯·伊尔贝克说。

## 探索化学王国：让复杂变简单

中青报 中青网记者 张茜

有时候简单的答案才是最好的。

2022年诺贝尔化学奖授予卡罗琳·露丝·贝尔托西(Carolyn R. Bertozzi)、摩根·梅尔达尔(Morten Meldal)、卡尔·巴里·夏普莱斯(K. Barry Sharpless)以表彰开发点击化学和生物正交化学。

诺贝尔奖官方网站介绍，点击化学和生物正交反应将化学带入了功能主义时代。

可以说，2022年诺贝尔化学奖是关于简化困难的过程。卡尔·巴里·夏普莱斯和摩根·梅尔达尔为化学的一种功能形式——点击化学奠定了基础。在这种化学中，分子构建块快速有效地结合在一起。卡罗琳·露丝·贝尔托西将点击化学带入了一个新的维度，并开始将其应用于生物体。

长期以来，化学家们一直渴望构建越来越复杂的分子。在药理学中，这通常涉及人工再制造具有药用性质的天然分子。研究过程中确实产生了许多令人叹为观止的分子结构，但通常都很耗时，而且生产成本很高。

诺贝尔化学委员会主席约翰·克维斯特说：今年的化学奖处理的不是过于复杂的问题，而是简单易行的问题。即使采取简单的方法也可以构建功能分子。

事实上，卡尔·巴里·夏普莱斯是第二次获得诺贝尔奖。2000年前后，他提出了点击化学的概念，这是一种简单可靠的化学形式。

不久之后，摩根·梅尔达尔和卡尔·巴里·夏普莱斯相互独立地发现了点击化学王冠上的明珠——铜催化叠氮化物炔烃环加成反应。这是一种优雅而高效的化学反应，目前已被广泛使用。比如用于药物开发、绘制DNA图谱和创建更能满足需求的材料。

卡罗琳·露丝·贝尔托西则将点击化学提升到了一个新的水平。为了在细胞表面绘制重要但难以捉摸的生物分子——聚糖，她开发了在生物体内起作用的点击反应。她的生物正交反应不会破坏细胞的正常化学反应。

贝尔托西关注的一个领域是肿瘤细胞表面的聚糖。她的研究揭示了一些聚糖似乎可以保护肿瘤免受人体免疫系统的影响，因为它们会使免疫细胞无法发挥功能。为了阻断这种保护机制，她和同事们发明了一种新型生物药物。这种药物目前正在对晚期癌症患者进行临床试验。

虽然我们目前还不能确定这些新疗法是否有效，但有一点是明确的，点击化学和生物正交化学的巨大潜力才刚刚开始显现。