DNA 数据存储中 DNA 保存技术的研究进展与挑战

熊成鹤，刘霞，高鲁娜，黄小罗*，梅辉*

（中国科学院深圳先进技术研究院，深圳合成生物学创新研究院，定量合成生物学重点实验室，广东省合成基因组学重点实验室，深圳合成基因组学重点实验室，深圳，518055）

摘 要 数据规模指数级增长与现有存储介质储存能力不足的矛盾日益凸显，亟需开发新类型介质来解决相应问题。DNA 基于其超高的数据存储密度、能耗低以及寿命长等特点，作为一种新兴的数据存储媒介备受关注，尤其针对海量“冷数据”，有望替代现有存储方式。在数据存储过程中，DNA 的有效保存是其中重要的一环，该环节直接影响了数据的存储密度、稳定性、存储时间以及数据的写入和读取。针对目前文献中 DNA 保存技术介绍较少的情况，本文综述了数据存储中 DNA 保存技术的研究进展和策略，讨论了现有的保存技术应用在 DNA 数据存储中面临的困难与挑战，对 DNA 数据存储的实现方式进行了展望。

关键词 DNA 数据存储；DNA 保存；封装技术

中图分类号 Q819 **文献标志码** **A** **doi: 10.12146/j.issn.2095-3135.20231107001**

Research Progress and Challenges of DNA Preservation Technology in DNA Data Storage

XIONG Chenghe, LIU Xia, Gao Luna, HUANG Xiaoluo*, MEI Hui*

(Shenzhen Key Laboratory of Synthetic Genomics, Guangdong Provincial Key Laboratory of Synthetic Genomics, Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

Abstract The growing contradiction between the exponential increase in data amount and the limited storage capacity of existing media is becoming increasingly evident, necessitating the development of new types of media to address this issue. Due to its ultra-high density, low energy consumption and long lifetime for data storage, DNA has attracted significant attention as an

来稿日期：2023-11-07 修回日期：2024-03-22

基金项目：国家重点研发计划资助项目（2021YFF1201700）；广东省基础与应用基础研究基金项目（2021A1515012375）；广东省合成基因组学重点实验室（2023B1212060054）；深圳合成基因组学重点实验室（ZDSYS201802061806209）

作者简介：熊成鹤，助理研究员，研究方向为核酸化学与 DNA 数据存储；刘霞，研究助理，研究方向为 DNA 数据存储；高鲁娜，研究助理，研究方向为 DNA 数据存储；黄小罗，高级工程师，博士生导师，研究方向为 DNA 数据存储，E-mail：huangxl@siat.ac.cn；梅辉，副研究员，博士生导师，研究方向为非天然核酸化学生物学，E-mail：hui.mei@siat.ac.cn。

emerging storage medium, particularly for massive “cold data”, with the potential to replace current storage methods. In the process of data storage, the effective preservation of DNA plays a crucial role, directly impacting the quality of storage density, stability, storage time, as well as data writing and reading. Due to the limited information available on DNA preservation techniques in the current literature, this paper provides an overview of current research progress and strategies in DNA preservation technology for data storage, discusses the difficulties and challenges faced when applying existing preservation techniques in DNA data storage, and presents prospects for the implementation of DNA data storage.

Keywords DNA data storage; DNA Preservation; Biomineralization

Funding This project is supported by National Key R&D Program of China (2021YFF1201700); Guangdong Basic and Applied Basic Research Foundation (2021A1515012375); Guangdong Provincial Key Laboratory of Synthetic Genomics (2023B1212060054); Shenzhen Key Laboratory of Synthetic Genomics (ZDSYS201802061806209)

1 前言

目前，数据信息的规模成指数级别的快速增长与其对应的存储介质存储能力的不足日益尖锐，这就亟需一种存储密度大、稳定性强以及寿命长的新型储存介质来解决迎面而来的问题，从分子生物学角度看待，DNA 具备相应潜力。与传统存储介质相比，DNA 具有多重优势：1、DNA 是一种稳定的存储介质，在自然界中，绝大多数生命的遗传信息都存储在 DNA 中；2、DNA 能实现低成本的快速复制，从而恢复信息；3、DNA 数据的信息密度非常大，如简单的大肠杆菌的储存密度约 10^{19}bit/cm^3；4、存储在 DNA 中的信息可以保存在多种容器和环境中；5、DNA 存储寿命更长，合适的条件下可以存在数千年乃至数万年之久；6、DNA 可生物降解，对环境无污染。相对于现有的存储介质（硬盘、磁带等）来说，它是一类绿色环保的高密度、高稳定性、长寿命的存储介质[1]。因此，选用 DNA 作为数据存储介质成为当前和未来一段时间内的热点发展方向 [0]。尤其对于政府及国际重要部门档案等大规模、不经常访问的“冷数据”的存储，人们更关注信息密度、保存时间和信息存储成本以及信息安全性等问题，而不是数据访问和读取的速度。传统的存储介质或密度较低、或寿命太短、或需要持续通电进而造成能源与成本的消耗，这就使得 DNA 数据存储较之硬盘等传统媒介有了更大的优势[4]。

DNA 数据存储[5-7]，大致分为 5 个步骤（图 1）。1、编码：采用二进制算法编码文本文档、图片或音频等信息，将得到的数据转换成 DNA 编码；2、DNA 合成：通过包括固相合成法、酶合成法、阵列芯片法等技术合成相应地 DNA 序列，即含有 4 种碱基（A=腺嘌呤、

T=胸腺嘧啶、C=胞嘧啶、G=鸟嘌呤）组成的脱氧核糖核酸序列；3、DNA 保存：将合成的 DNA 序列保存在体内或体外等不同的环境或容器中；4、DNA 测序：当需要读取或还原数据时，提取保存的 DNA，对其进行测序；5、解码：将序列信息还原成二进制数据和文件。

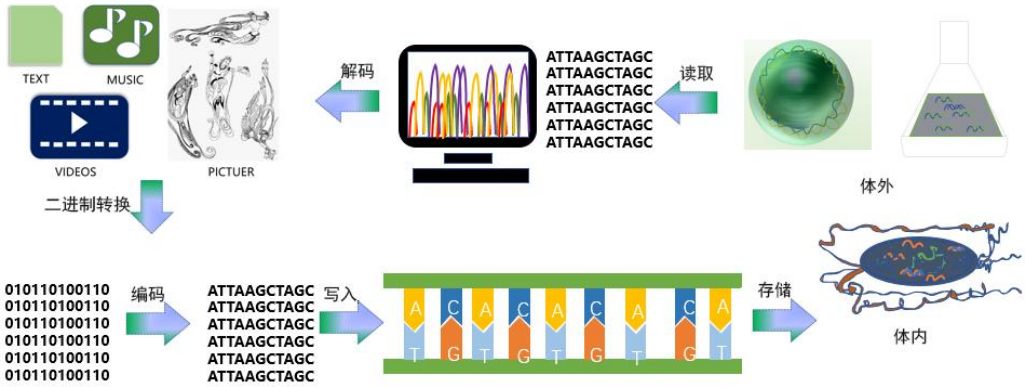

图 1：DNA 存储流程示意图。

DNA 作为新兴的数字信息存储载体，将对飞速增长的信息化时代产生重要的影响。近些年来，DNA 数据存储吸引了越来越多的科研工作者对其进行深入探索和研究。为了将 DNA 代替传统数据存储介质的想法付诸实践，研究者们正在进行多方面的研究：在 DNA 合成方面，除了现有的固相化学合成法，以 TdT 酶为基础的酶法合成 DNA 正在大力发展[8]，以实现更高效率、更长片段的 DNA 的合成；在 DNA 测序方面**错误！未找到引用源。**, DNA 第二代和第三代测序技术的发展和迭代、手持式单分子 DNA 测序仪的发明[10]等，增加测序高通量及降低成本；在算法方面，各种编译码方法也层出不穷[11]，用以增加信息密度和降低成本等。

然而，在 DNA 保存方法和载体等方面的研究进展不是很理想。原因可能为天然且未受保护的 DNA 是一种相对脆弱的生物分子，直接裸露在外界，会遭受到许多的“威胁”，比如不适宜的温度（一般是高温）[12-13]、湿度[14]、辐射[15]、氧化[16-19]等，甚至于有些环境中的化学物质和核酸酶等，都会严重加速 DNA 的分子结构破坏。而 DNA 数据存储的成本和持久性将很大程度上受到 DNA 保存技术的影响。因此，DNA 保存技术在实现 DNA 数据存储这一目标中占据着重要的位置。本文将着重从 DNA 数据存储中 DNA 保存技术的策略、现状、挑战及展望来进行概述。

2 DNA 数据保存技术的策略和现状

2.1 体外保存法

2.1.1 固态保存法

DNA 溶液在常温下保存时间通常不超过半年，在-20℃等低温环境下可保存数年之久。将 DNA 液体形式直接转变成 DNA 干粉可有效地延长其保存时间，并且具备所占体积小、能耗低、易于运输等优点。制备 DNA 干粉的方法有喷雾干燥法、喷雾冷冻干燥法、空气干燥法以及冷冻干燥法，目前冷冻干燥法是成本最低、应用最为广泛的方法[22]。但是，需要注意的是，DNA 样品在干燥过程中，其稳定性也可能会发生很大变化。例如，同样在 70℃、50%相对湿度的条件下进行 5 天的加速寿命实验，真空离心干燥 DNA 的稳定性要明显优于冷冻干燥的[21]。目前已知在 50%相对湿度、50℃条件下，没有任何其他保护措施的干燥后的固态 DNA 最多保存两周[14]。从当前角度出发，为了达到较好的保存效果，将干燥的 DNA 样品保存在较低相对湿度下已经成为了常用的 DNA 保存方法[22-23]。2015 年，Cozzolino 等人[24]利用化学物质浸渍，能够溶解细胞膜和变性蛋白的 FTA 滤纸，将 DNA 包裹、固定和稳定在其中，然后在常温下存储，以此来保护 DNA 免受核酸酶、氧化、紫外损伤及其他微生物的降解等。

封装技术保存 DNA 主要是为了隔绝外界水、氧气和光等可能影响 DNA 稳定的因素，而被封装的 DNA 样品一般是如上述经过干燥处理后的 DNA 干粉**错误！未找到引用源。**。Bonnet 等人[26]在 2009 年的实验中得出，如果不受水和氧气的影响，DNA 干粉的保存时间很长，并且其一级和二级结构仍可以保留，从而验证了长期室温保存 DNA 的可能性。DNAshell®技术是基于密封的不锈钢微型胶囊，在惰性气氛下，给予 DNA 干粉的保护。Clermont 等人[27]的实验证实基于 DNAshell®技术保存的 DNA 干粉在室温条件下保存 18 个月，未发现其有降解。在极端条件(76℃，50%相对湿度)下 30 小时，DNAshells®的保护性也不受影响，可以保证 DNA 干粉在室温条件下 100 年不发生降解。

2.1.2 添加剂法

化学添加剂二甲基亚砜(DMSO)，可以防止 DNA 单链断裂。不过由于 DMSO 具有一定毒性，且对于 DNA 数据存储中大量的 DNA 保存来说，DMSO 的使用也是大量的，因此它并不是一种理想的保护剂[28]。

海藻糖，作为一种天然的双糖，普遍认为是一种多功能的、可保护生物体免受冷冻、加热、干燥和辐射的试剂，早在 1997 年，海藻糖就在 DNA 防辐射损伤的应用中充当保护剂[31]。在 DNA 干粉中添加保护性基质—海藻糖，再利用 DNAshell®技术保存，在 76℃、50%相对湿度条件下可将保存时间从 30 小时延长至 1 周甚至 1 个月。这预示着在常温下，添加海藻糖的 DNA 干粉利用 DNAshell®技术保存的时间可远远大于 100 年[27]。

2014 年，Grass 等人[**错误！未找到引用源。**利用了与骨矿物晶体组成和结构上类似的矿物羟基磷灰石与 DNA 吸附，验证了羟基磷灰石对 DNA 的保护能力。2022 年，Sarkar 等人[33]设计出一种在水介质中长期稳定和存储 DNA 的阳离子液体材料。并且发现 DNA 在双阳离子液体中比单阳离子液体更加稳定。

某些碱式盐也可以提升 DNA 的保存时间和稳定性。2020 年，Grass 等人[21]的实验证明磷酸钙、氯化镁和氯化钙的添加，都有利于 DNA 的稳定保存，而且对于高负载量（＞30%）时，历经 70 ℃ 加速老化实验 6 天，DNA 仍然完好。2022 年，Grass 等人[34]在 pH = 6 的条件下，将碱式盐 KH_2PO_4/K_2HPO_4 和 $CaCl_2$ 的混合物共沉淀，得到了 18 wt% (wt%为质量百分比；本文指数据 DNA 重量占总重量的百分比)高负载的共沉淀结晶磷酸钙，该结构有助于保存古代 DNA 的遗传信息。在条件为 70 ℃、50%相对湿度的加速老化实验中证实，5 天后，结晶型磷酸钙中负载 DNA 的浓度高于不定性磷酸钙中 DNA 有三个数量级。

除了上述添加剂以外，某些添加剂已成功商用，例如 Biomatrica ® DNAstable ®是一种商业上允许在无水环境中储存 DNA，保护样品免受水解、氧化和微生物降解的试剂。与之有类似功能的商业试剂，还包括 DNA SampleMatrix ®（测试了 26 个月，估测 60 ℃ 下，保存 30 年）和 GenTegra™ DNA（测试了 4 个月，估测 76 ℃ 下，保存 10 年）等**错误！未找到引用源。**。 2013 年的报道中将海藻糖、Biomatrica ® DNAstable ®、Tris 缓冲的聚乙烯醇（polyvinyl alcohol, PVA）做了 DNA 保存的对比试验，其结论为：海藻糖在 DNA 的短期保存中是足以胜任的，但建议使用 Biomatrica ® DNAstable ®和 Tris 缓冲的 PVA 进行 DNA 的长期保存，特别是在热带国家等高温环境中进行 DNA 的长期保存[36]。类似地，GenTegra-DNA 非常适合在环境温度变化较大的情况下运输，可耐受美国军方严格的环境运输规格-80 ℉ (-62 ℃)至 160 ℉ (71 ℃)，预计在运输过程中可以耐受温度范围为-80 ℃ 至 76 ℃。

2.1.3 二氧化硅封装法

最早的、非人为的 DNA 体外保存，是古代化石中保存的 DNA，受这一现象启发，可以将 DNA 保存在化石或类化石材料中。古代 DNA(ancient-DNA, a-DNA)的保存与物理、化学和生物因素的变化息息相关，比如埋藏学、化石储存、氧化作用、微生物成岩作用、pH 值和离子强度，以及阳离子和腐殖质的存在等，都会影响 DNA 的保存时间和保存质量[37]。2012 年， Allentoft 和 Bunce 等人[37]就通过建立相关模型对 DNA 的半衰期进行研究，他们的结果表明：在 5 ℃ 时，该模型预测骨骼中一个 30 bp 的线粒体 DNA（mitochondrial DNA, mtDNA）片段的半衰期为 15.8 万年。2013 年，考古专家从西班牙北部阿塔普埃卡的“白骨之坑”（La Sima de los Huesos）中发掘了一些超过 30 万年历史的德宁格尔熊的化石，并通

过测序得到了它的 mtDNA 序列，这一发现表明，DNA 可以在多年冻土之外的环境存在数十万年。同年，Grass 等人通过模拟化石气密性的封装结构来实现了用于聚合物标记的 DNA 条形码[38]在外源应激（氧化/高温/紫外）的作用下，所封装的 DNA 序列相对于未保护的 DNA 序列有着更高的稳定性[39]。同样在上述这个遗址，2014 年，德国马普进化人类学研究所的 Meyer 教授与其科研团队发现了一副约 43 万年前的人类股骨，并从中提取出 mtDNA [40]。研究发现这些化石自沉积以来一直没有受到重大干扰，较低的温度、恒定的湿度以及稳定的环境，或许是该遗址中 DNA 保存如此之久的重要因素。

模仿化石材料使用的二氧化硅在保存 DNA 方面得到了广泛的应用。Grass 最早使用二氧化硅封装 DNA[14]，并在实验室测得封装的 DNA 最高可以承受 200 ℃ 高温，且在 60 ℃ 下可保存 2 个月（相当于在室温可保存 2 年），但该方法的缺点是 DNA 负载量太小（0.7%），进而造成 DNA 存储密度太低。2019 年，Grass 等人[41]通过磁性纳米颗粒上 DNA 层和阳离子聚合物聚乙烯亚胺（PEI）层交替覆盖的设计，最后在最外层包裹二氧化硅以隔绝外部因素的影响，使 DNA 免受外部损害，实现了 DNA 存储密度的提升（155 ng/cm），并将质量负载提升至 3.4%。2020 年，Erlich 和 Grass 等人[42]设计了一种“万物皆可存储 DNA 信息（DNA-of-things, DoT）”概念的存储体系。在 DoT 框架中，合成的 DNA 用于数据的记录，然后将这些 DNA 分子封装在纳米二氧化硅颗粒中，进而将其与可降解热塑性聚酯混合，制成可用于三维打印（3D 打印）的材料， 随后打印了一个包含 45kB 信息的“斯坦福兔子”。后续，通过从兔子身上取下一小块材料，释放其包含的 DNA 并通过扩增和测序，完美地还原了原始信息并用于下一代兔子的打印。最终，科研人员通过兔子的五次复制和打印，验证了 DNA 作为数据存储介质的潜力。

数字微流控是一种可精确控制和操纵微尺度流体的技术，尤其特指亚微米结构的技术[43]。2022 年 Grass 等人[44]结合数字微流控和二氧化硅封装这两种技术，首先将 DNA 封装在二氧化硅纳米颗粒中，然后利用数字微流控处理，可以实现完全自动化的信息检索。该方法将当今档案储存系统的存储容量提高 3 个数量级。在其老化试验条件为 70 ℃，50%相对湿度下 4 天（相当于室温下 116 年），存储的数据通过 PCR 扩增被成功恢复了，这提供了一种自动化、稳定存储 DNA 编码数据的方案，有望解决未来数据存储能力不足的态势。2022，Grass 等人[45]设计了含有二硫醇的二氧化硅颗粒的核壳来存储 DNA，并且封装在此材料中的 DNA 呈现持久和高密度的数字储存格式。2022 年 Fei 和 Xiao 等人[46]提出了一种基于孔径可调节的二氧化硅水凝胶的新型 DNA 存储系统，含有 DNA 数据的二氧化硅颗粒保存在可用于 3D 打印的水凝胶材料中，其数据密度可达到 1.04×10^{10} GB/g，该数据也是迄今为止

报道的 DNA 数据存储密度最大值之一（图 2）。2022 年，Zhu 和 Wuttke 等人[47]利用全血原位冷冻硅化方法来保存 DNA，该方法廉价、可靠并且存储的 DNA 数据在长期的高温潮湿条件下（70 ℃，60%相对湿度）仍然稳定。以书本格式大小的基因组信息通过这种方法储存，推测可以超过 1000 年，每个人成本仅为 0.5 美元。最后该课题组通过 3D 打印建立了基因组库，这为未来面临灭绝的动物和植物的 DNA 信息保存留下了希望。

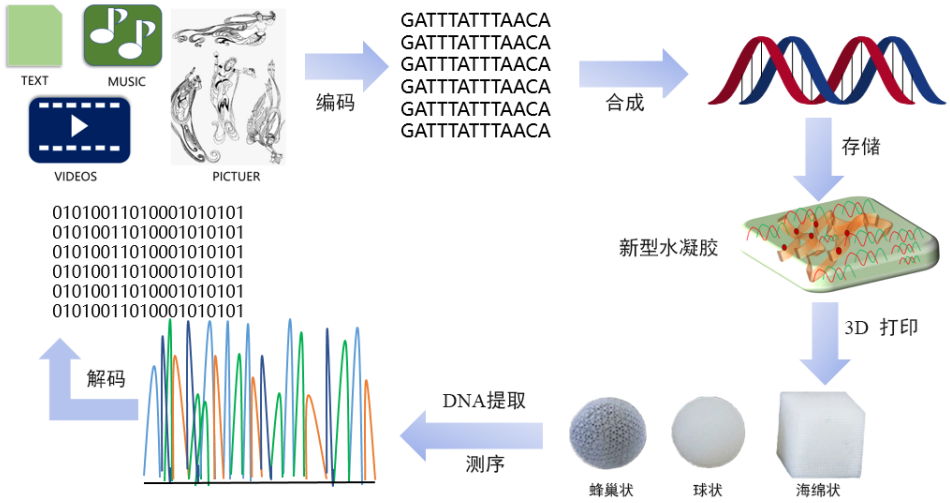

图 2：水凝胶结合 3D 打印技术储存 DNA 数据。该技术首先将数字信息转换成不同序列的 DNA，再使用二氧化硅包裹，随后将其储存在可用于 3D 打印的孔径可调节的水凝胶材料中，当需要读取时，剪下一部分材料，对其中的 DNA 进行扩增和测序，进而还原出原始数据[46]。

2.2 体内保存法

相对于 DNA 体外保存来说，在细菌或细胞等体内的保存可利用机体自身的 DNA 复制、校对以及纠错机制，实时记录生物状态以及高效且随机的访问路径[48-49]。DNA 体内保存大致分为两类；

2.2.1 质粒自主复制保存法

1996 年，Davis 团队[3]在质粒中存储了小维纳斯女神“Microvenus”图片，并将其转入大肠杆菌中保存。2010 年，Venter 团队[50]首次将人工合成的一个支原体基因组（1,077,947 bp）存入酵母细胞中并成功进行复制和传代，这在将体外信息存储在细胞中的历史上具有里程碑意义。

2.2.2 染色体整合保存法

2014 年，Yang 和 Voigt 等人[51]用不同的重组酶形成逻辑门“开关”，利用这些“开关”，在大肠杆菌中构建了一个存储阵列可以记录 1.375 字节的信息，证明了重组酶可以分层并用于永久记录转录逻辑门的瞬时状态。2020 年，Qi 等人[52]利用携带大量寡核苷酸池的混菌培

养物，成功将 445kb 的数字信息保存在大肠杆菌中。2021 年，天津大学元英进教授团队[53]在酿酒酵母中利用转化偶联重组技术，成功构建了含有两张图片和一条视频信息的人工染色体。通过将携带数据的染色体传递至 100 代，证明了该方法保存数据的稳定性。但是，对于基于重组酶的 DNA 数据存储，在特定的位置需要一个特定的重组酶，因此，数据存储的密度不高，且与宿主基因组的接口有限，最终导致数据写入效率比较低[54]。以上是利用重组酶将 DNA 数据整合进染色体而实现数据的体内保存。

另一方面，使用 CRISPR-Cas 系统也可以将 DNA 数据整合进染色体。2017 年，哈佛大学 Church 团队[55]通过 CRISPR-Cas 系统在大肠杆菌的基因组中，插入了一些黑白图像和一张飞驰骏马动图编码成的大约 2.6KB 大小的短片（GIF），并在大肠杆菌中经过多代繁殖后，最终实现了总体精度为 90%的数据恢复。2021 年，哥伦比亚大学的研究人员[56]用电化学方法调控 CRISPR 系统在生物体内实现了数据信息的自动存储。他们将“hello world”录入大肠杆菌的 DNA 中，发现它们繁衍 80 代后正确率达 90%以上。他们还将大肠杆菌混入土壤微生物中，对混合物进行测序，发现仍然可以恢复存储的信息。2024 年，Hou 等提出了一类“细胞磁盘”的存储系统，实现了可随机读取和重写的高密度体内 DNA 数据存储。该系统中，具有自我复制能力的酵母细胞被用作存储信息的“室”，类似于“磁盘块”，高达 10^5 个“块”组成了“细胞磁盘”，而酵母细胞的每个基因组都插入了一个定制的基于 CRISPR/cas9 的“锁与钥匙”模块，利用该模块，可以从“细胞磁盘”中选择性地检索、擦除或重写目标细胞“块”。配合 DNA 合成和测序，构成了一个完整的基于“细胞磁盘”的 DNA 数据存储系统[57]。

综上所述，目前适配 DNA 数据存储系统中 DNA 保存的策略大致分为体外存储（干燥封存、糖类包裹、碱式盐添加、纳米颗粒包裹存储等）和体内储存等。其中，纳米颗粒存储主要研究思路是带阳离子聚合物上的正电荷与 DNA 上的负电荷相结合，然后包裹二氧化硅；或二氧化硅与存储体系结构且具有不变内存材料相融合，然后再采用 3D 打印技术将 DNA 数据进行保存；或同样采用阴阳离子相结合原理，将二硫醇环绕在 DNA 上，然后再将二氧化硅掺杂进去，由于二硫化合物会生物降解，最后形成了二氧化硅保存 DNA 信息的短暂数据存储系统。另外，除了利用有机或无机液体/固体介质，DNA 还能存储在生物体内，充分证明了 DNA 可以储存在任何均一介质中，而对于 DNA 数据所储存的密度、寿命、成本、稳定性等问题则由相对应的 DNA 存储技术来决定。这为科研工作者提供了一种新的思路，在体内或体外均可以开展研究，对于如何改进和创新 DNA 存储技术需要不断的开拓思维和

理论与实际相结合（图 3）。

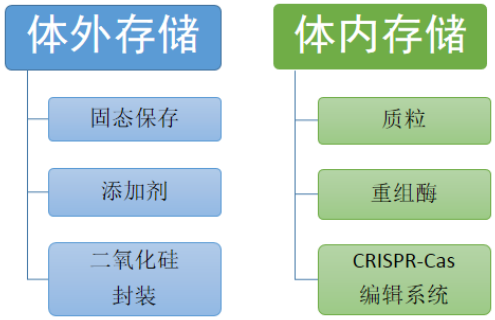

图 3：不同类型的 DNA 数据存储技术。

3 DNA 数据存储系统中 DNA 保存所面临的挑战

虽然传统的磁性/光学/固态存储介质的写入和读取成本低且易于使用，但其寿命较短和密度低，储存能力会受到密度的限制，相应的吞吐量也比较低。传统磁性介质有磁带和硬盘，光学介质有 CD 光盘，固态介质有快闪储存器等。一般磁盘和磁带寿命＜10 年，光盘相对于它们而言，寿命较长，可达 50 年。此外，这些信息存储系统还存在消磁、粘连和磨损等问题，对于长期存储的数据则需要不断地更新载体，反复拷贝，这在一定程度上增加了维护费用，同时也反应了传统介质存储数据的耐久性较差。而 DNA 作为数据存储载体，其在寿命、密度、耐久性、吞吐量和储存能力等诸多方面都比传统数据存储介质要优越，尤其是寿命可达数十万年，预估存储系统可保存达数千年之久[9, 26]。（图 4）

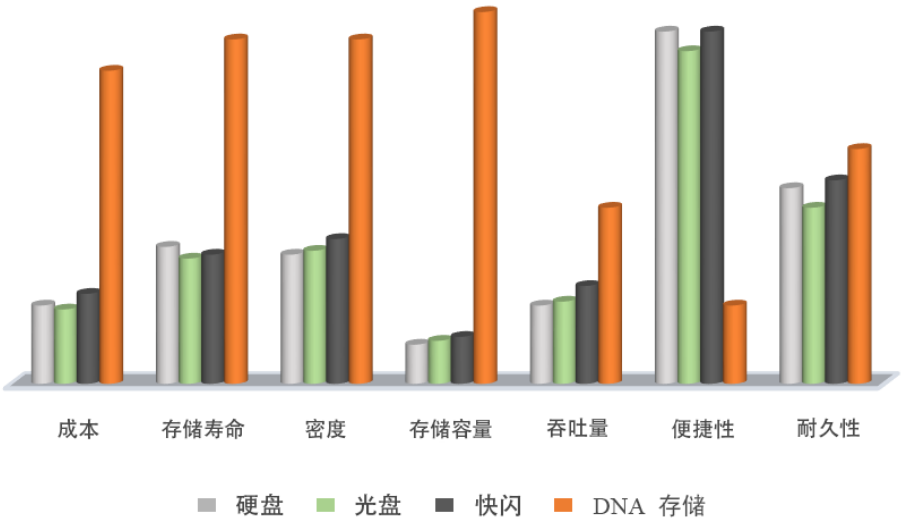

图 4：磁带、光盘、快闪和 DNA 存储介质保存 DNA 情况。

总的来说，与传统存储系统相比，DNA 数据存储系统具有许多优势，但仍存在许多技

术局限，如 DNA 存储系统中写入和读取成本仍然很高，高效快速对数据进行随机访问仍存在困难，兼顾存储持久性与简易读取数据的存储技术有待完善。尤其是后者，DNA 保存的策略和形式将直接影响数据存储的密度、稳定性、存储时间以及数据的写入和读取等，需要克服以下多方面的困难和挑战。

1、DNA 测序系统与 DNA 保存形态的矛盾

首先是测序与 DNA 保存之间物态要求的矛盾。目前主流的三代测序手段都要求待测样品呈液态，而经过反复的冻融循环，冷冻保存的 DNA 会受到损伤。在此方面，人们已经开始尝试改变 DNA 的空间结构，新的 DNA 结构更稳定，用于探索减小冻融循环对 DNA 的损伤，并且已经有所成效[58] 。2020 年，Keller 和 Smith 等人[62]将有效冷冻保存 2D 和 3D DNA 折纸纳米结构的方法用在了冷冻保存 DNA 上，该研究还报道了反复冷冻和解冻循环的效果。然而，二维 DNA 折纸纳米结构在至少 32 个冻融循环中保持其结构完整性，但冰晶的形成使 DNA 折纸对苛刻的样品处理条件逐渐变得敏感。在经历 32 次冻融循环的 3D DNA 折纸纳米结构中，没有发现冻损，然而 1000 次冻融循环会导致明显的碎片。与此同时，该篇报道还发现添加冷冻保护剂甘油和海藻糖可以有效保护 DNA 折纸纳米结构。

考虑到未来的使用环境，人们始终期待开发出常温即可简单操作的 DNA 保存方法，然而到目前还没有找到合适的。DNA 数据存储系统是离不开测序技术的[63]，除了样品的反复冻融，PCR 作为测序的必要步骤，必须要有加热过程，这与利于 DNA 保存的条件是矛盾的，至于影响的大小，目前还没有可靠的实验数据来证实。除此之外，反复读取信息，即反复测序，是否会对 DNA 造成不可逆的伤害，现在也未可知。

从测序角度而言，添加物的增加甚至会干扰测序的结果。除此之外，各添加物之间的反应也是需要考虑的重要因素。相对来讲，不含有添加物的、直接干燥后的 DNA 虽然在保存时间上还有待加强，但其作为“器件”存在具有天然优势：固体、能分离能连接、无杂质干扰且具备良好的光电特性等，这都为未来 DNA 数据存储系统形成提供了可能。

目前来说，DNA 数据存储各步骤之间仍处于相对独立地发展阶段，而形成数据存储系统时，就需要各部分之间产生联系。以二氧化硅封装法为例，此方法在保存效果上很有优势，然而封装技术现在是个难点，Grass 等人的实验中仅封装就用了 4 天[14]，而密闭状态使得它的信息读取有着很大的阻碍。

2、信息存储效率和稳定性的问题

提及信息存储效率问题，科研工作者们最初都习惯在编、译码环节上进行探究[64]。诚

然，完善的编译码方法会大大提高存储效率，合理的纠错方法还能大幅度减少冗余，增加数据存储的鲁棒性[66]。而对于 DNA 存储整个流程的存储效率问题，尤其是保存环节，需要考虑 DNA 负载量，即 DNA 重量占总重量的百分比。较高的 DNA 负载会使得相同的信息量在存储系统中所占据空间更小，换句话说，更有利于高密度数据存储。Grass 等人[14]开发的 DNA 储存程序中 DNA 数据被二氧化硅包裹，实现了与天然化石相当的 DNA 稳定性。但是，以现有的实验技术手段，即使在最佳条件下，二氧化硅包裹的 DNA 负载量也仅为 0.7 wt％。该课题组的另一种封装方法[41]，磁性纳米颗粒——二氧化硅逐层设计可将 DNA 负载量提高到 7.8 wt％。Grass 等人[21]通过向 DNA 中添加碱式盐，将 DNA 负载量提高到 30 wt%。由此可见，在负载量方面，还有很大的进步空间。

另一方面，DNA 本身的稳定性不仅是关乎 DNA 存储系统能否投入使用的重要指标，更是这类新型数据存储的最具开发潜力和最有优势的一个特点。2013 年，Ivanova 等人**错误！未找到引用源。**进行了干燥 DNA 保存方法不同的对比实验，1 年后，除了未保护 DNA(33%)，剩下的每个样本的平均 PCR 成功率在不同方法处理之间并没有显著差异(97-100%)。在室温下放置 4 年后，最好的结果是 Biomatrica(98%),其次是 PVA(96%)和冰箱存储(- 20 ℃; 88%)。在未保护的 DNA(2%)、海藻糖(29%)和冰箱冷藏(4 ℃; 49%)中观察到严重的 DNA 降解。

类比传统数据存储，DNA 数据存储系统的可用时间也是一个十分重要的参数。所谓可用时间，也称持久性，实际上是区别于 DNA 的生存时间或者说寿命所提出，就时间长度而言，可用时间一般小于生存时间。DNA 数据存储的可用时间究竟能有多少，现在还没有统一的判断标准和可靠的实验数据。

3、信息安全及成本问题

许多技术的发展在应用时都会成为一把“双刃剑”，DNA 数据存储也不例外。未来 DNA 存储如若应用于军事，可以通过人体携带 DNA 数据，有了 DNA 存储技术，我们人体就是“云硬盘”，非常方便。但也正是因为 DNA 的载体十分丰富，不像磁盘、磁带等容易检测，DNA 数据存储系统的信息安全问题变得令人担忧，未来每个人都可以是一个超大容量的 DNA 库数据盘，如何监管就将成为新的挑战。

DNA 合成和测序的成本已经引起了人们的广泛关注，降低合成成本的方法主要依赖于合成技术的进步和编码方法的改进，其合成和测序的成本以每年 5 倍和 12 倍的指数迅速下降，这相比于电子媒体每年 1.6 倍的速度要快得多[67]。而关于 DNA 保存的成本研究还比较少。比如常用的冷冻保存法，其优点是保存时操作简单、处理材料用时短，且已经经过了实

际应用的验证，但是缺点是占用空间大，能源使用多，长期保存成本高，且长期保存时间不是很理想。如果此方法试图降低成本，可以寻找合适的自然条件，例如模仿化石藏于地下或者低温地区，未来可以用于存储一些大量且不常用的信息。但在任意地点的保存成本仍是个难题。

4 DNA 数据存储的展望

到目前为止，主流数据存储手段仍然是传统的磁盘、闪存、DVD 等，而让数据存储在 DNA 中是一个非常热点的研究方向。相比于其他介质，DNA 具备存储密度大、出错率低、能耗低等优势，尤其对于海量冷数据储存，有望替代或补充现有存储形式比如磁盘、光盘等。目前相关研究大多集中在编解码、合成、测序、纠错等环节，而对 DNA 保存这一环节关注较少。事实上，与 DNA 数据存储系统相匹配的 DNA 保存技术能直接影响数据存储的存储密度、稳定性、存储时间以及数据的写入和读取。

目前，不管 DNA 数据储存在冻土、低温环境、冰川有机质、酵母细胞、大肠杆菌及细菌/细胞里，这都是为了隔绝外界环境，提供一种类似真空或封闭的体系来保存 DNA，使 DNA 在外界应激源的刺激下缓慢降解，实现长久保存。其中，将 DNA 保存在体外，在一定程度上摆脱了活体复杂体系的不确定性，但是仍然面临着诸多问题，比如温度、湿度、紫外、光照、氧化等外源刺激条件造成的 DNA 降解，进而造成信息的丢失。对于此，科学家通过制作各种添加剂或材料组成“保护膜”，用于减少温度、湿度、紫外、氧化等对 DNA 数据的伤害。目前亟需寻找能够更适用于现有测序技术的新型 DNA 物态和保存技术，在提高保存稳定性和持久性的同时，可适配现有测序手段，从而匹配数据存储。

借鉴于自然生命，利用活细胞进行 DNA 信息编码、合成、存储、提取、检索和复制，进而构建数据信息存储系统，目前也取得了不少进展。国内外都已经在大肠杆菌/细胞等中进行了相关研究，证实了将编码合成的 DNA 数据存储在活体中相对于传统存储能力，可以高几个数量级。但是细胞/细菌体内是一个非常复杂的体系，修饰或者添加新的 DNA 数据到细胞/细菌体内的操作具有挑战性，且细胞/细菌尺寸较大、存储信息总量较低，还存在着不稳定性遗传，造成数据丢失、插入、删除和替换等风险[52]。未来工作可以在提高单个细胞存储数据容量，建立 DNA 信息存储覆盖读、写、增、删改查的一体化体内 DNA 信息存储系统。近些年开始的酵母基因组合成计划已经可以实现 Kbp 到 Mbp 级染色体的从头合成和精简，未来的技术可以考虑如何将数据信息存入这些合成染色体中，以提高容量。同时，通过 Crispr-Cas 系统的进一步优化，实现大片段基因组在细胞内的任意增删和修改，以及整个细胞层面存储信息的增删和修改。在此基础上，通过优化染色体基因组合成、基因组转化、

Crispr 编辑、细胞培养、DNA 提取、DNA 测序的一体化流程，实现覆盖读、写、增、删改查的高容量 DNA 存储系统。另外，通过三代测序进行实时的长度长测序读取能够大大提升体内基因组存储的分析效率，对于提升 DNA 信息存储数据的读取速度具有极大帮助。 与此同时，在体内 DNA 的编码设计上，通过考虑加入纠错码， 来降低 DNA 信息存储在长期传代中的错误，进一步提高系统的鲁棒性。为了保存数据的稳定性，还可以考虑利用微生物孢子的形态存储数据，以保证体内存储信息能够应对多种复杂的环境。

随着 DNA 数据存储系统的建立与完善，对于 DNA 保存技术的要求也在逐步提高，这就迫切地要求我们寻找新的方法来实现 DNA 的长久稳定保存，适应 DNA 数据存储系统的发展。我们相信，随着各方面技术的发展， DNA 数据存储必将获得突破，从而弥补现有存储方式的不足，并逐步走向商业化、产业化，走进每个人的生活。

参考文献：

[1] Meiser LC, Nguyen BH, Chen YJ, et al. Synthetic DNA applications in information technology [J]. Nature Communications, 2022, 13(1): 352.

[2] Gantz J, Reinsel, D. Extracting value from chaos [R]. International Data Corporation (IDC), Framingham, MA (2011).

[3] Davis J. Microvenus [J]. Art Journal, 1996, 55(1): 70-74.

[4] 董一名，孙法家，武瑞君，等. DNA 数字信息存储的研究进展 [J]. 合成生物学，2021, 2(3): 323-334.

Dong YM, Sun FJ, Wu RJ, et al. Research progress on DNA molecules for digital information storage [J]. Synthetic Biology Journal, 2021, 2(3): 323-334.

[5] Gervasio J, Oliveira H, Martins A, et al. How close are we to storing data in DNA? [J]. Trends in Biotechnology, 2023.

[6] Ping Z, Chen SH, Zhou GY, et al. Towards practical and robust DNA-based data archiving using the yin-yang codec system [J]. Nature Computational Science, 2022, 2, 234-242.

[7] Lu MY, Wang Y, Qiang W, et al. Towards high-density storage of text and images into DNA by the “Xiao-Pang” codec system [J]. Science China Life Sciences, 2023, 66, 1447-1450.

[8] Lee HH, Kalhor R, Goela N, et al. Enzymatic DNA synthesis for digital information storage [J]. bioRxiv, 2018. doi: 10.1101/348987.

[9] Pääbo S, Poinar H, Serre D, et al. Genetic Analyses from Ancient DNA [J]. Annual Review of Genetics, 2004, 38(1): 645-679.

[10] Pennisi E. Genome Sequencing. Search for pore-fection [J]. Science, 2012, 336(6081): 534-537.

[11] Church GM, Gao Y, Kosuri S. Next-generation digital information storage in DNA [J]. Science, 2012, 337(6102): 1628-1628.

[12] Bruskov VI, Malakhova LV, Masalimov ZK, et al. (2002). Heat-induced formation of reactive oxygen species and 8-oxoguanine, a biomarker of damage to DNA [J]. Nucleic Acids Research, 2002, 30(6): 1354-1363.

[13] Molina MDC, Anchordoquy TJ. Degradation of lyophilized lipid/DNA complexes during storage: the role of lipid and reactive oxygen species [J]. Biochimica et Biophysica Acta (BBA)-Biomembranes, 2008, 1778(10): 2119-2126.

[14] Grass RN, Heckel R, Puddu M, et al. Robust chemical preservation of digital information on DNA in silica with error−correcting codes [J]. Angewandte Chemie International Edition, 2015, 54(8): 2552-2555.

[15] Gut IG. Depurination of DNA and matrix-assisted laser desorption/ionization mass spectrometry [J]. International Journal of Mass Spectrometry and Ion Processes, 1997, 169: 313-322.

[16] Evans RK, Xu Z, Bohannon KE, et al. Evaluation of degradation pathways for plasmid DNA in pharmaceutical formulations via accelerated stability studies [J]. Journal of Pharmaceutical Sciences, 2000, 89(1): 76-87.

[17] Lindahl T, Andersson A. Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid [J]. Biochemistry, 1972, 11(19): 3618-3623.

[18] Lindahl T, Nyberg B. Rate of depurination of native deoxyribonucleic acid [J]. Biochemistry, 1972, 11(19): 3610-3618.

[19] Suzuki T, Ohsumi S, Makino K. Mechanistic studies on depurination and apurinic site chain breakage in oligodeoxyribonucleotides [J]. Nucleic Acids Research, 1994, 22(23): 4997-5003.

[20] Wan E, Akana M, Pons J, et al. Green technologies for room temperature nucleic acid storage [J]. Current Issues in Molecular Biology, 2010, 12(3): 135.

[21] Kohll AX, Antkowiak PL, Chen WD, et al. Stabilizing synthetic DNA for long-term data storage with earth alkaline salts [J]. Chemical Communications, 2020, 56(25): 3613-3616.

[22] Anchordoquy TJ, Molina MC. Preservation of DNA [J]. Cell Preservation Technology, 2007, 5(4): 180-188.

[23] Alkhamis KA. Influence of solid-state acidity on the decomposition of sucrose in amorphous systems (I) [J]. International Journal of Pharmaceutics, 2008, 362(1-2): 74-80.

[24] Peluso AL, Cascone AM, Lucchese L, et al. Use of FTA Cards for the storage of breast carcinoma nucleic acid on fine-needle aspiration samples [J]. Cancer Cytopathology, 2015, 123(10): 582-592.

[25] Karni M, Zidon D, Polak P, et al. Thermal degradation of DNA [J]. DNA and Cell Biology, 2013, 32(6): 298-301.

[26] Bonnet J, Colotte M, Coudy D, et al. (2010). Chain and conformation stability of solid-state

DNA: implications for room temperature storage [J]. Nucleic Acids Research, 2010, 38(5): 1531-1546.

[27] Clermont D, Santoni S, Saker S, et al. Assessment of DNA encapsulation, a new room-temperature DNA storage method [J]. Biopreservation and Biobanking, 2014, 12(3): 176-183.

[28] Repine JE, Pfenninger OW, Talmage DW, et al. Dimethyl sulfoxide prevents DNA nicking mediated by ionizing radiation or iron/hydrogen peroxide-generated hydroxyl radical [J]. Proceedings of the National Academy of Sciences, 1981, 78(2): 1001-1003.

[29] Yoshinaga K, Yoshioka H, Kurosaki H, et al. Protection by trehalose of DNA from radiation damage [J]. Bioscience, Biotechnology, and Biochemistry, 1997, 61(1): 160-161.

[30] Crowe JH, Hoekstra FA, Crowe LM. Anhydrobiosis [J]. Annual Review of Physiology, 1992, 54(1): 579-599.

[31] Librizzi F, Vitrano E, Cordone L. Dehydration and crystallization of trehalose and sucrose glasses containing carbonmonoxy-myoglobin [J]. Biophysical Journal, 1999, 76(5): 2727-2734.

[32] Kohll AX, Antkowiak PL, Chen WD, et al. Stabilizing synthetic DNA for long-term data storage with earth alkaline salts [J]. Chemical Communications, 2020, 56(25): 3613-3616.

[33] Mahapatra A, Barik S, Satish L. Assessing the Suitability of a Dicationic Ionic Liquid as a Stabilizing Material for the Storage of DNA in Aqueous Medium [J]. Langmuir, 2022, 38(48): 14857-14868.

[34] Antkowiak PL, Koch J, Rzepka P, et al. Anhydrous calcium phosphate crystals stabilize DNA for dry storage [J]. Chemical Communications, 2022, 58(19): 3174-3177.

[35] Wan E, Akana M, Pons J, et al. Green technologies for room temperature nucleic acid storage [J]. Current Issues in Molecular Biology, 2010, 12(3): 135-142.

[36] Ivanova NV, Kuzmina ML. Protocols for dry DNA storage and shipment at room temperature [J]. Molecular Ecology Resources, 2013, 13(5): 890-898.

[37] Allentoft ME, Collins M, Harker D, et al. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils [J]. Proceedings of the Royal Society B: Biological Sciences, 2012, 279(1748): 4724-4733.

[38] Paunescu D, Fuhrer R, Grass RN. Protection and deprotection of DNA—high-temperature

stability of nucleic acid barcodes for polymer labeling [J]. Angewandte Chemie International Edition, 2013, 52: 4269-4272.

[39] Dabney J, Knapp M, Glocke I, et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments [J]. Proceedings of the National Academy of Sciences, 2013, 110(39): 15758-15763.

[40] Meyer M, Fu Q, Aximu-Petri A, et al. A mitochondrial genome sequence of a hominin from Sima de los Huesos [J]. Nature, 2014, 505(7483): 403-406.

[41] Chen WD, Kohll AX, Nguyen BH, et al. Combining Data Longevity with High Storage Capacity—Layer-by-Layer DNA Encapsulated in Magnetic Nanoparticles [J]. Advanced Functional Materials, 2019, 29(28): 1901672.

[42] Koch J, Gantenbein S, Masania K, et al. A DNA-of-things storage architecture to create materials with embedded memory [J]. Nature Biotechnology, 2020, 38(1): 39-43.

[43] Choi K, Ng AH, Fobel R, et al. Digital microfluidics [J]. Annual Review of Analytical Chemistry, 2012, 5: 413-440.

[44] Antkowiak PL, Koch J, Nguyen BH, et al. Integrating DNA Encapsulates and Digital Microfluidics for Automated Data Storage in DNA [J]. Small, 2022, 18(15): 2107381-2107390.

[45] Koch J, Kerl AC, Schawalder N, et al. Preserving DNA in Biodegradable Organosilica Encapsulates [J]. Langmuir, 2022, 38(37): 11191-11198.

[46] Fei ZJ, Li MJ, Cheng C, et al. Bionic-structure thermo-responsive (best) hydrogels with controllable layer for high-capacity DNA data storage [J]. Nano Select, 2022, 1-16.

[47] Zhou L, Lei Q, Guo JM, et al. Long-term whole blood DNA preservation by cost-efficient cryosilicification [J]. Nature Communications, 2022, 13(1): 6265-6279.

[48] Organick L, Ang SD, Chen YJ, et al. Random access in large-scale DNA data storage [J]. Nature Biotechnology, 2018, 36(3): 242-248.

[49] Sheth RU, Wang HH. DNA-based memory devices for recording cellular events [J]. Nature Reviews Genetics, 2018, 19(11): 718-732.

[50] Gibson DG, Glass JI, Lartigue C, et al. Creation of a bacterial cell controlled by a chemically synthesized genome [J]. Science, 2010, 329(5987): 52-56.

[51] Yang L, Nielsen AK, Fernandez-Rodriguez J, et al. Permanent genetic memory with> 1-byte

capacity [J]. Nature Methods, 2014, 11(12): 1261-1266.

[52] Hao M, Qiao H, Gao Y, et al. A mixed culture of bacterial cells enables an economic DNA storage on a large scale [J]. Communications Biology, 2020, 3(1): 1-9.

[53] Chen WG, Han MZ, Zhou JT, et al. An artificial chromosome for data storage [J]. National Science Review, 2021, 8(5):nwab028.

[54]周廷尧, 罗源, 蒋兴宇. DAN数据存储：保存策略与数据加密 [J]. 合成生物学, 2021, 2(3): 371-383.

Zhou TY, Luo Y, Jiang XY. DNA data storage: preservation approach and data encryption [J]. Synthetic Biology Journal, 2021, 2(3): 371-383.

[55] Shipman SL, Nivala J, Macklis JD, et al. CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria [J]. Nature, 2017, 547(7663): 345-349.

[56] Yim SS, McBee RM, Song AM, et al. Robust direct digital-to-biological data storage in living cells [J]. Nature Chemical Biology, 2021, 17(3): 246-253.

[57] Hou ZH, Qiang W, Wang XX, et al. “Cell Disk” DNA Storage System Capable of Ranrom Reading and Rewriting [J]. Advanced Science, 2024, 2305921.

[58] Rothemund PW. Folding DNA to create nanoscale shapes and patterns [J]. Nature, 2006, 440(7082): 297-302.

[59] Kielar C, Xin Y, Xu X, et al. Effect of staple age on DNA origami nanostructure assembly and stability [J]. Molecules, 2019, 24(14): 2577.

[60] Zhu B, Zhao Y, Dai J, et al. Preservation of DNA Nanostructure Carriers: Effects of Freeze-Thawing and Ionic Strength during Lyophilization and Storage [J]. ACS Applied Materials & Interfaces, 2017, 9(22): 18434-18439.

[61] Chung WJ, Cui Y, Chen CS, et al. Freezing shortens the lifetime of DNA molecules under tension [J]. Journal of Biological Physics, 2017, 43(4): 511-524.

[62] Xin Y, Kielar C, Zhu S, et al. Cryopreservation of DNA Origami Nanostructures [J]. Small, 2020, 16(13): 1905959.

[63] Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA [J]. Genomics, 2016, 107(1): 1-8.

[64] Erlich Y, Zielinski D. DNA Fountain enables a robust and efficient storage architecture [J]. Science, 2017, 355(6328): 950-954.

[65] Yazdi SHT, Yuan YB, Ma J, et al. A Rewritable, Random-Access DNA-Based Storage System [J]. Scientific Reports, 2015, 5(1): 1-10.

[66] Press WH, Hawkins JA, Jones SK, et al. HEDGES error-correcting code for DNA storage corrects indels and allows sequence constraints [J]. Proceedings of the National Academy of Sciences, 2020, 117(31): 18489-18496.

[67] Carr PA, Church GM. Genome engineering [J]. Nature Biotechnology, 2009, 27(12): 1151-1162.