

引文格式：

刘德霖, 沈玥, 平质. DNA 存储技术的数据安全研究进展 [J]. 集成技术, 2024, 13(3): 74-88.

Liu DRL, Shen Y, Ping Z. Research progress on data security of DNA storage technology [J]. Journal of Integration Technology, 2024, 13(3): 74-88.

DNA 存储技术的数据安全研究进展

刘德霖^{1,2} 沈玥^{2,3} 平质^{2,3*}

¹(中国科学院大学生命科学学院 北京 100049)

²(深圳华大生命科学研究院 深圳 518083)

³(华大工程生物学长荡湖研究所 常州 213299)

摘要 随着现代科学技术的不断迭代升级, 人类生产生活中产生的数据量呈爆炸式增长, 传统的硅基存储介质(如硬盘、闪存等)逐渐无法满足全球范围内的数据存储需求。DNA 因其存储密度大、存储时间长、能源消耗低等颠覆性的优势被认为是新一代的理想存储介质。该文围绕 DNA 存储技术, 首先阐述了其基本原理与流程; 其次, 在传统数据安全的范畴下, 结合近期 DNA 存储领域的相关研究, 归纳介绍了 DNA 存储技术的数据安全研究现状, 包括数据加密、数据弹性、恶意攻击检测等; 最后, 讨论了现阶段 DNA 存储技术的数据安全所面临的挑战及发展趋势。随着多学科交叉协同发展, 在广角度多层面的数据安全保护下, DNA 存储技术有望解决海量数据存储需求与有限存储资源间的供需不平衡问题, 实现多场景的商业化应用。

关键词 DNA 存储; 数据存储; 数据安全

中图分类号 Q 819; TP 309.2 文献标志码 A doi: 10.12146/j.issn.2095-3135.20231030002

Research Progress on Data Security of DNA Storage Technology

LIU Deruilin^{1,2} SHEN Yue^{2,3} PING Zhi^{2,3*}

¹(College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China)

²(BGI Research, Shenzhen 518083, China)

³(BGI Research, Changzhou 213299, China)

*Corresponding Author: pingzhi@genomics.cn

Abstract With the advancement of modern science and technology, the volume of data generated through human production and daily life has surged. Traditional silicon-based storage media such as hard disks and flash memory are gradually becoming inadequate to meet the growing global data storage demands. Due to

收稿日期: 2023-10-30 修回日期: 2024-03-15

基金项目: 国家重点研发计划项目(2021YFF1200100); 国家自然科学基金项目(32101182); “广东特支计划”科技创新青年拔尖人才项目(2019TQ05Y876)

作者简介: 刘德霖, 博士研究生, 研究方向为 DNA 存储; 沈玥, 研究员, 研究方向为合成生物学、合成基因组学、DNA 合成技术与工具开发; 平质(通讯作者), 副研究员, 研究方向为合成生物学、DNA 存储、生物信息分析算法, E-mail: pingzhi@genomics.cn.

its remarkable advantages, such as incredibly high storage density, extremely long-term storage capabilities, and minimal energy consumption, DNA is regarded as the ideal next-generation storage medium. This review primarily focuses on DNA storage technology. It introduces the fundamental theory and workflow, and subsequently provides an overview of the research status regarding data security in DNA storage technology, within the context of conventional data security. This includes aspects like data encryption, data resilience, the detection of bio-cyber attacks and so on. The article also discusses the challenges and emerging trends in data security of DNA storage technology. Through interdisciplinary development, DNA storage technology is expected to resolve the conflict between massive data and limited conventional storage space and can be commercialized in multiple scenarios eventually, all while ensuring broad-spectrum and multi-tier data security.

Keywords DNA storage; data storage; data security

Funding This work is supported by National Key Research and Development Program of China (2021YFF1200100), National Natural Science Foundation of China (32101182), Tip-top Scientific and Technical Innovative Youth Talents of Guangdong Special Support Program (2019TQ05Y876)

1 引言

信息存储是人类文明发展的重要标志。从古代结绳记事、仓颉造字的传说,到现代人人普及的磁盘闪存,信息通过不断迭代更新的载体,跨越时间和空间代代相传。随着全球科学技术的进步,精准医疗健康、5G 数字技术及以 ChatGPT 大语言模型为代表的人工智能技术等领域的发展带来了大数据的繁荣^[1]。国际数据公司的统计显示,2021 年,全球数据总量达到了 85 ZB(1 ZB \approx 10⁹ TB),预计到 2026 年,全球结构化与非结构化的数据总量将达到 221 ZB^[2]。面对如此海量的数据,传统硅基存储介质将无法未来存储需求,亟需变革式的新兴存储介质来应对未来数据的爆炸式增长。

DNA 存储的概念最早可追溯到 20 世纪 60 年代,控制论创始人 Norbert Wiener 和苏联科学家 Mikhail Neiman 几乎同时独立提出利用 DNA 存储数据的构想^[3]。率先将这一想法变为现实的是艺术家 Joe Davis,1996 年,他将 35 位的暗亮

像素信息编码为 DNA 后导入到大肠杆菌中,并成功读取^[4],实现了硅基数据载体与碳基数据载体之间的信息传输。进入 21 世纪以来,DNA 测序技术超“摩尔定律”的成本下降速度和高通量 DNA 合成技术的不断突破为 DNA 存储技术的发展带来了契机。一方面,转码算法的设计作为 DNA 存储技术的核心之一,在信息能否成功写入或读出,以及 DNA 所承载信息的有效恢复等方面起着至关重要的作用。从 2012 年开始,美国^[5-6]、欧洲^[7-9]及中国^[10]的研究团队先后根据合成 DNA 的生物约束条件分别开发了不同的转码算法。另一方面,在贯通整个 DNA 存储流程的基础上,对标当今计算机存储系统进行 DNA 存储系统功能模块和集成化的开发是推动 DNA 存储技术产业化应用的必经之路。从随机访问到布尔搜索的实现^[11-13],从增删查改到数据加密的完善^[14-16],从自动化端到端 DNA 存储设备到 DNA 存储微流控体系的诞生^[17-18],研究人员通过多学科交叉融合研究使得 DNA 存储技术越来越接近实际应用。

在 DNA 存储技术日趋成熟的道路上，与其相关的数据安全也需要得到相应的重视和保护。与常规存储的数据安全相比，DNA 存储技术的数据安全因其发展历程相对较短、对应生态不够完善而尚未有较全面的论述。本文围绕 DNA 存储技术简要阐述其基本原理与流程，在数据安全范畴下，结合近期相关研究着重归纳介绍 DNA 存储技术的数据安全，最后讨论当前 DNA 存储技术的数据安全所面临的挑战及发展趋势。

2 DNA 存储技术的基本原理与流程

DNA 是自然界中天然的遗传信息载体。以腺嘌呤 (Adenine, A)、胞嘧啶 (Cytosine, C)、鸟嘌呤 (Guanine, G)、胸腺嘧啶 (Thymine, T) 4 种碱基为基元，DNA 能够稳定地将亲代遗传信息传递给子代已有亿万年之久。DNA 存储技术通过预设计好的转码算法将计算机中代表二进制数据的 0 和 1 编码为由 A、C、G、T 4 种碱基组成的 DNA 序列，依靠 DNA 合成技术合成指定碱基序列的 DNA，以实现信息的存入，通过 DNA 测序读取对应的 DNA，实现信息的读出^[19]。

与传统的硅基存储介质相比，DNA 在存储密度、存储时间、能源消耗等方面具有无可比拟的优势。研究表明，DNA 的存储密度高达 455 EB/g (1 EB \approx 10⁶ TB)^[20-21]，据此推测，存储 2025 年的全球数据大约需要 10 kg DNA 分子^[22]。与闪存、光盘等传统存储介质相比，DNA 分子的存储密度要高出 8 个数量级^[23]。除了高存储密度外，在适宜的条件下，DNA 还具有比较强的稳定性。在化石中，DNA 的半衰期约为 521 年^[24]。在格陵兰岛上，研究人员甚至能够对 200 万年前的 DNA 进行恢复和测序，以追踪古代生物群落的生态和进化^[25]。此外，在能源消耗方面，用 DNA 存储 1 EB 数据的功耗约为 10⁻¹ W，而用硬盘存储 1 EB 数据的功耗约为 4.0 \times 10⁷ W^[26]。尽管有研究表明，过去 10 年中，在数据需求迅速增加的情况下，数据中心的大量能效提升使得能源使用保持了大致平稳的水平^[27]，但是当前对于需要长期归档存储的冷数据来说，DNA 是更为合适的存储载体。

如图 1 所示，DNA 存储技术的基本流程可以分为两个阶段。(1)信息存入阶段，包括信息编码、DNA 合成和 DNA 保存；(2)信息读出阶段

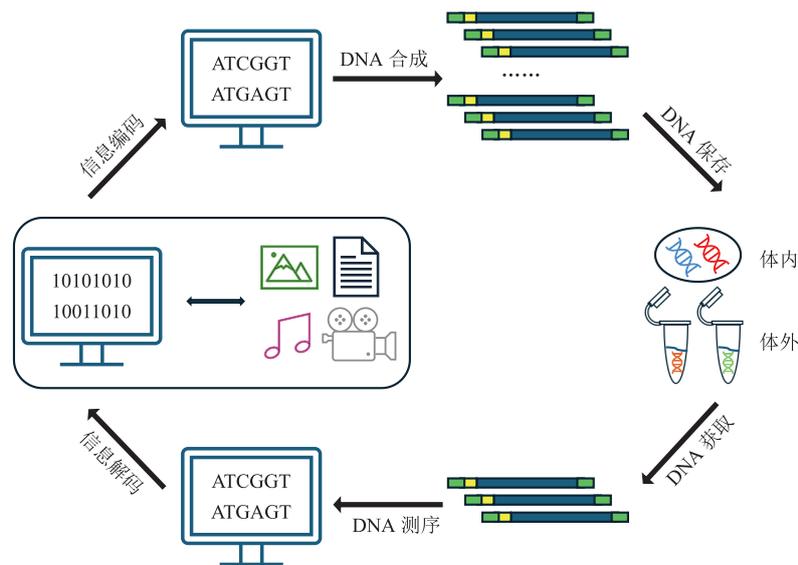


图 1 DNA 存储的基本流程

Fig. 1 Basic steps of DNA storage

段, 包括 DNA 获取(随机访问)、DNA 测序和信息解码。

2.1 信息存入

(1) 信息编码: 通常情况下, 计算机中待存储的文本、图像、音视频等数字文件需要先处理为以 0、1 为基元的二进制序列, 再通过相应的转码算法实现 0、1 和 A、C、G、T 之间的映射转换, 完成数字信息的编码。不同的转码算法之间存在差别, 但是基本上都包含文件数据压缩、纠错编码及地址索引添加、二进制-碱基映射与编码等过程^[28]。

(2) DNA 合成: 根据信息编码过程中产生的 DNA 序列进行人工合成。当前, 人工合成 DNA 的方法主要包括柱式合成、高通量芯片合成在内的化学合成法和基于酶的生物合成法^[29]。

(3) DNA 保存: 合成好的 DNA 需要选择适宜的条件进行保存。现阶段, 人工合成 DNA 的保存方式主要分为体外保存和体内保存。体外保存指将人工合成的 DNA 通过冻干、稳定剂添加^[30-31]或保护材料包被^[32-33]等方法保存在生物体外。体内保存的方式主要包括质粒重组^[34]、人工染色体封装^[35]及 CRISPR-Cas 体系插入拼接^[36]。

2.2 信息读出

(1) DNA 获取(随机访问): 捕获全部或部分已保存的 DNA 片段, 准备进行下游的测序读取。捕获 DNA 片段的方式主要分为两种: 一种是基于多重聚合酶链式反应 (polymerase chain reaction, PCR) 的目标引物扩增^[12]; 另一种是以生物素-链霉亲和素反应为代表的纳米磁珠分离捕获^[14]。

(2) DNA 测序: 将捕获的 DNA 分子片段进行序列测定。DNA 测序技术包括第一代 Sanger 测序^[37]、第二代高通量测序^[38-39]和第三代单分子测序^[40-41]。后两者为现阶段科学研究的主流测序技术。

(3) 信息解码: 根据信息编码阶段的转码算

法, 将通过 DNA 测序获取的 A、C、G、T 序列进行信息解码、纠错, 并还原为二进制序列, 进而恢复为文件数据, 实现信息读取。

3 DNA 存储技术的数据安全

数据安全与数据存储密不可分。随着数字化时代的到来, 大体量的数据存储不仅意味着需要满足海量数据的存储需求, 更重要的是在此基础上保证数据的安全归档。围绕数据写入到数据读出的完整闭环, DNA 存储技术的数据安全由内而外表现在 3 个层面。

第一层面是数据的自身安全, 即对 DNA 分子承载的序列信息和结构信息的安全保护, 通常利用生物技术中的复杂生物分子反应过程对 DNA 分子进行信息加密和隐藏^[42-43]。现阶段, DNA 链置换反应^[44-45]和 DNA 折纸术^[15]是保护序列信息和结构信息比较常用的生物技术方法。CRISPR-Cas 系统能够替代 DNA 片段对体内数据进行重写^[46], 也具备潜在的数据安全保护能力。第二层面是数据的存储安全, 即尽可能地避免携带信息的 DNA 分子因受外界影响而发生突变、断裂或者损伤, 主要通过数据弹性来衡量数据的存储安全程度。在 DNA 存储技术中, 面对偶发的随机性错误或者系统性错误, 为生成的信息序列添加纠错码^[47]是在有效纠正序列突变, 甚至断裂重排的同时保持极低成本的数据修复方式。对于不确定性更强的外界威胁因素来说, 采用无机或有机材料外壳保护 DNA 分子是当下被认为最有前途的体外保存方式之一^[48]。另外, 在体内存储中, 宿主细胞的生物安全也关乎着携带数据的 DNA 分子能否稳定长久地保存下去。第三层面是数据的传输安全, 即在数据写入 (DNA 合成) 或数据读取 (DNA 测序) 环节进行的生物网络安全检测, 以及对废弃陈旧的 DNA 存储系统设备进行的彻底性数据清除。其中, 前者是为了防止

硅基存储系统中的恶意序列在编码与数据写入环节混入 DNA 存储系统中，造成大规模的数据泄露与数据破坏；后者是保证数据迁移机密性的必要流程。

DNA 存储技术的数据安全总体图景及其子领域如图 2 所示。近年来，得益于各学科之间的交叉融合发展，DNA 存储技术在数据安全层面的相关研究取得了一定进展。

3.1 数据加密

数据加密是通过加密算法和加密密钥将明文转换为密文的过程，是对信息进行保护的可靠方法。DNA 存储技术中的数据加密基于分子生物技术及生物学中的困难问题，利用 DNA 分子自身的物理性质进行明文序列转换。

最早的 DNA 加密方法诞生于 1999 年，Gehani 等^[49]根据传统密码学中的“一次一密”机制，利用 DNA 分子高存储密度的特性和 PCR 扩增技术设计出了 DNA 层面的“一次一密”加

密体系。随后，Chen^[50]提出了基于 DNA 并行计算的“一次一密”加密方案。近年来，基于 DNA 链置换反应的数据加密策略发展迅速。Zhu 等^[45]提出了一种基于 DNA 链置换的新型加密框架，该框架包含 3 种分别用来混淆明文、生成密钥及混淆扩展密钥的加密策略，同时将参数敏感的生化反应与密钥获取紧密关联在一起，提高了密钥传输的安全性。此外，生物学中的困难问题也展现出其在加密领域的潜在优势。2020 年，Grass 等^[51]证明遗传短串联重复序列可以用来生成强加密密钥，并从人类基因组中生成了 80 bit 的强密钥，成功实现了 17 kB 的数字信息在 DNA 存储中的加密和解密。

伴随着 DNA 加密思维范式从序列到结构的转变，Zhang 等^[15]开发了 DNA 折纸密码学系统，如图 3 所示。在该系统中，发送方 Alice 会将骨架链的折叠作为加密密钥，将文本信息先转换为类似盲文图案的点阵列，再通过密钥将点阵

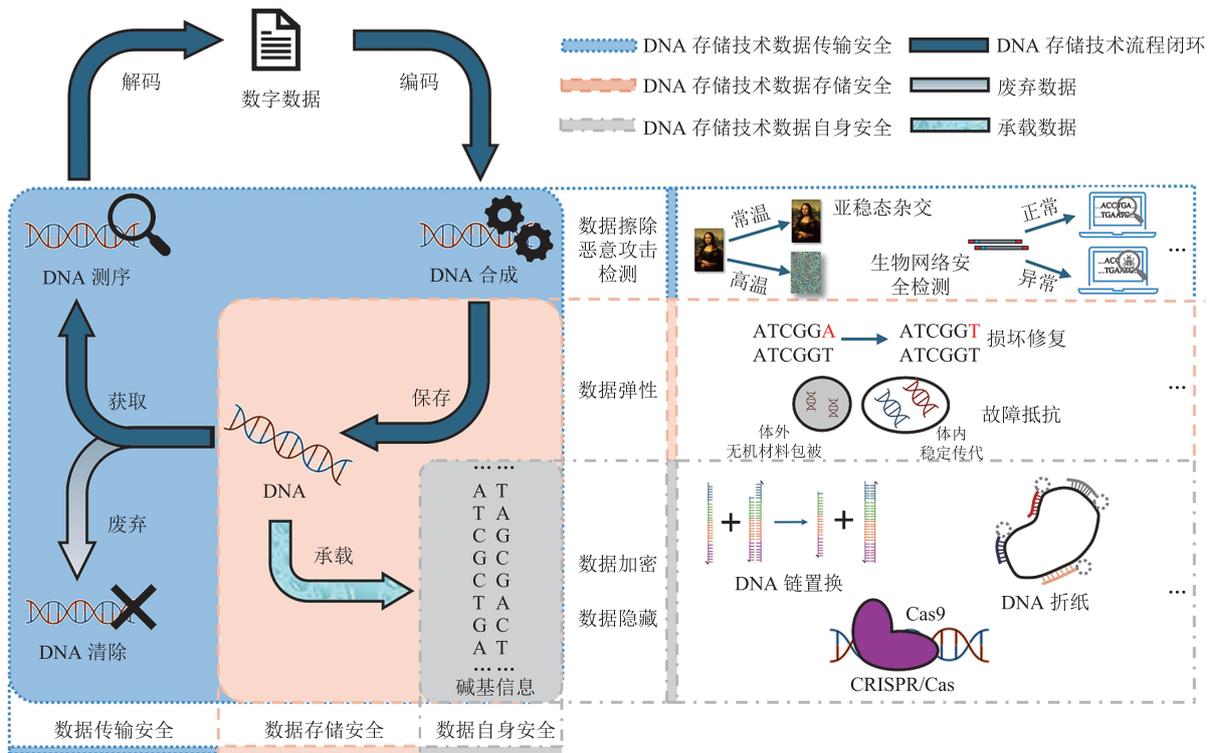


图 2 DNA 存储技术的数据安全总体图景及其子领域

Fig. 2 Overview and sub-areas of data security in DNA storage technology

列加密为杂交若干生物素化短链的骨架链。接收方 Bob 可以从获取的密钥上得到订书链, 并将其与主链发生物理折叠, 形成点阵列的形状, 在加入链霉亲和素后便可以在原子力显微镜下识别, 最终获取原始信息。整个 DNA 折纸密码学系统是数据加密和数据隐写的结合, 能够根据不同的 DNA 折纸结构实现对不同类型信息的加密。

3.2 数据隐藏

数据隐藏(即隐写术)的原理是将敏感数据混合在非敏感的信息载体中, 不让除预期的接收者之外的任何人知晓信息的内容。1999 年, Clelland 等^[52]利用微点技术在 DNA 层面上实现了数据隐写, 开创了 DNA 分子隐写术的先河。DNA 隐写术主要包含 3 类对象: 携带敏感数据的 DNA、大量的非敏感载体、提取敏感数据的特定引物密钥。因此, DNA 隐写术的安全程度主要取决于引物密钥的机密性及非敏感载体的混淆性。针对引物密钥的安全问题, Li 等^[16]开发了 Cas12a 辅助 DNA 隐写术, 在 DNA 隐写的过程中对特异性引物密钥也进行了隐写操作。在 Cas12a 辅助 DNA 隐写术中, 真密钥(特异性引物)首先与假密钥(非特异性引物)混合或者与 3'-端的冗余序列连接从而形成预密钥, 完成真密钥的隐写, 然后利用 CRISPR/Cas12a 技术切割假密钥或者 3'-端的冗余序列, 以获取真密钥, 从

而确保 DNA 编码信息的存储和传输。Cas12a 辅助 DNA 隐写术为经典 DNA 隐写术中的密钥添加了保护层, 进一步提升了其安全性能。更重要的是, Koch 等^[53]提出的万物 DNA (DNA-of-things) 概念极大地扩展了 DNA 隐写术的应用场景。将携带敏感数据的 DNA 分子封装在二氧化硅小球内, 再将这些小球嵌入各种日常物品中, 即可为敏感数据提供由外向内的多重安全保障: 二氧化硅小球的寻找困难、DNA 分子的检测困难以及敏感数据的提取困难。不过, 基于万物 DNA 的隐写术也具有高成本、长周期的局限, 需要在未来的研究中予以克服。

3.3 数据弹性

数据弹性用于衡量存储系统对各种故障的恢复力或者承受力。在转码算法遵循 DNA 序列生物约束的前提下, DNA 存储系统中的数据恢复力主要依靠编码环节引入的纠错码; 由于信息存储载体介质的稳定性会影响数据安全, 因此, DNA 存储系统对故障的承受力主要包括体外保存阶段的 DNA 分子环境耐受性及体内保存阶段的活细胞存储稳定性。

3.3.1 纠错恢复

现阶段, 引入纠错码是 DNA 存储技术中能够同时保证成本和准确率的纠错机制。目前使用较多的纠错码是里德-所罗门(RS)码, 少数

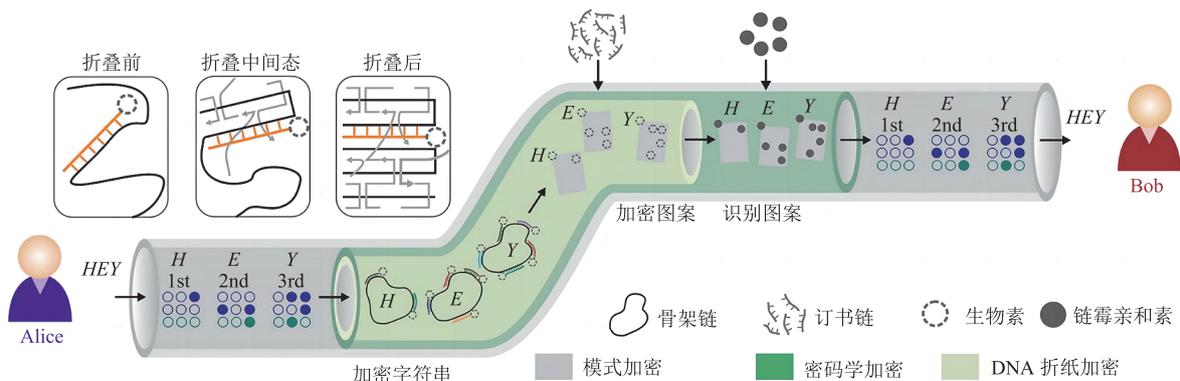


图 3 DNA 折纸加密系统的工作流程^[15]

Fig. 3 Workflow of DNA origami cryptography system^[15]

研究采用低密度奇偶校验码、汉明码等。虽然纠错码的种类并不稀少,但是其中大部分并不能完全解决 DNA 存储技术中独特的生物约束挑战(碱基替换、删除和插入)。幸运的是,近年来,国内外研究团队针对这一问题开发出了多种适配于 DNA 存储技术的纠错码。在算法原理方面,Press 等^[54]设计了一种名为基于哈希编码的穷举搜索解码的纠错码,该纠错码能够修复插入、删除和替换 3 种基本类型的 DNA 错误,但是其解码的复杂度及引入的冗余度较高。此外,图论算法在 DNA 存储技术中的纠错方面也展现出了非凡的潜力。Song 等^[55]基于德布鲁因图(de Bruijn Graph)开发出的 DNA 序列重建算法 DBGPS 能够从大量的替换、插入和删除片段中恢复原始无误的 DNA 序列。Zhang 等^[56]开发出了基于图的编解码一体化方案 SPIDER-WEB。SPIDER-WEB 不仅具有完善的纠错功能,还能够实现实时信息检索,其检索速度约为当前单分子测序的 300 倍。除了算法原理层面上的深耕外,纠错码的译码方式也得到了改进。Ding 等^[57]将软判决引入 DNA 存储系统,开发出了软判决译码软件 Derrick。与传统的硬判决策略相比,Derrick 使得 RS 码的纠错能力提高了 2 倍,并且将不可纠正的错误概率降低了几个数量级。这些先进的纠错码的开发提高了 DNA 存储系统的纠错性能,为 DNA 存储系统的内部数据恢复提供了强大的助力。

3.3.2 体外 DNA 分子的环境耐受性

众所周知,未受到保护的 DNA 分子是比较脆弱的。近年来,随着 DNA 寡核苷酸合成成本的降低,关于 DNA 存储技术的科学研究更倾向于使用体外保存^[58]。因此,设计出能够抵抗由外界理化因素改变导致 DNA 降解的风险,从而使 DNA 分子能够长期稳定存在的保存策略至关重要。受到化石中 DNA 因骨骼残骸的保护可以长期稳定存在的启发^[59],研究人员先后使用碱

金属盐^[60]、二氧化硅与磁性纳米颗粒^[61-63]、磷酸钙^[33]等材料对 DNA 进行封装保护,大大增强了体外保存中的 DNA 分子稳定性。虽然这些无机包被材料能够对抗多种恶劣环境对 DNA 分子的侵蚀,但是它们形成的保护结构的装载密度较低,部分回收技术也比较复杂。考虑到有机高分子材料的强附着性,Fei 等^[64]设计了一种具有长期稳定性且较易回收的热响应功能梯度(thermally responsive functionally graded, TRFG)水凝胶,其理论信息负载密度接近 7.0×10^9 GB/g。TRFG 水凝胶保存 DNA 分子的过程如图 4 所示。室温条件下,该水凝胶在 $\text{pH} < 5$ 的 DNA 溶液中发生溶胀,并吸附 DNA 溶液,水凝胶基质中超支化阳离子聚合物能够与溶液中的 DNA 分子发生相互作用,形成纳米颗粒聚合物,并被包封到水凝胶基质中;在温度达到 60°C 后,会发生退溶胀,并释放吸附的溶液,而 DNA 则被保留在 TRFG 水凝胶中。这种温度敏感性允许 TRFG 水凝胶进行多个溶胀和退溶胀循环,以富集 DNA 分子,提高其负载量。除了合成有机高分子保护材料外,天然有机高分子保护材料因造价成本极低而在 DNA 存储领域也展现出了巨大的应用前景。例如,Liu 等^[65]基于纤维素开发了一种可持续低成本(9.27×10^{-18} 美元/碱基)的 DNA 存储系统,理论最大物理密度约为 10^6 GB/g。该系统通过静电吸附的方式将 DNA 分子保存在纤维素上,并且可以很容易地使用碱性盐溶液进行回收,即使暴露在空气中,也能够保存多年。

3.3.3 体内活细胞的存储稳定性

在数据安全层面上,与体外保存相比,基于活细胞的 DNA 体内存储不仅要注重宿主细胞外的环境因素,还需要考虑携带数据的 DNA 作为生物大分子在宿主细胞内的潜在效应。现阶段,在拥有专业设备的实验室条件下,体内保存宿主的选择主要是大肠杆菌^[36]和酵母菌^[66],由于日常或极端环境中宿主细胞需要具有一定

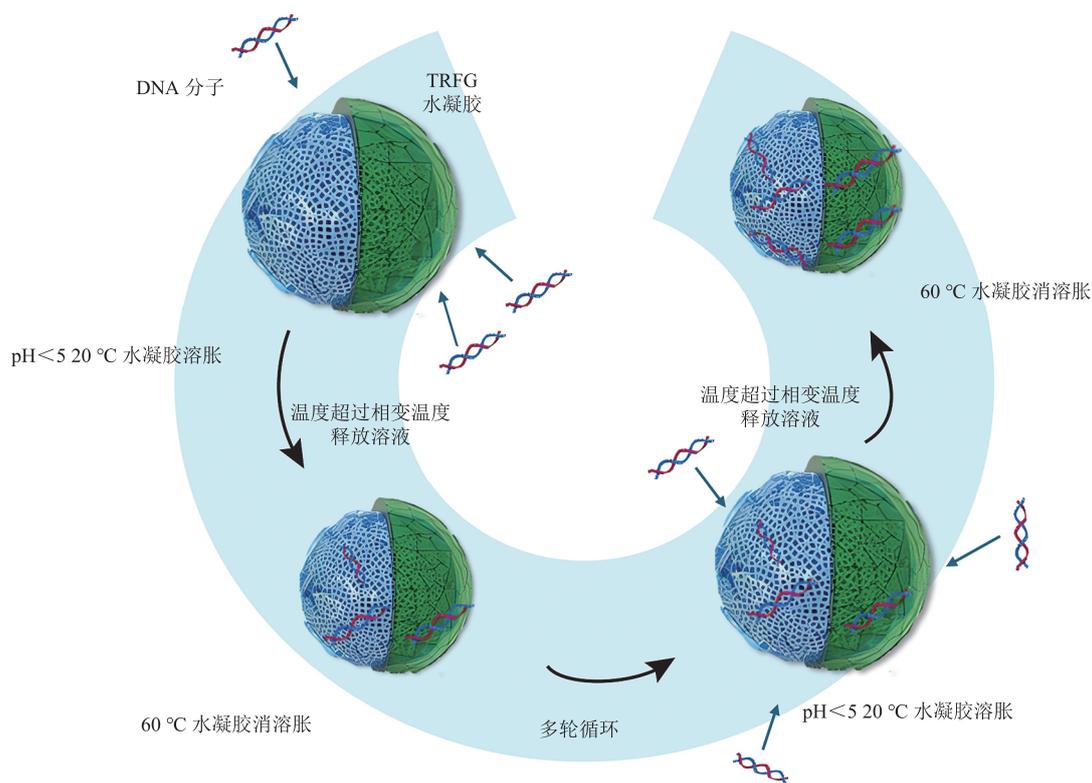


图 4 TRFG 水凝胶通过多轮循环保存 DNA 的过程^[64]

Fig. 4 Process of TRFG hydrogel storing DNA in multiple cycles^[64]

的抗逆性, 因此, 芽孢杆菌^[67]和具有天然抗污染特性的极端微生物^[68]也成了体内保存的宿主。例如, Sun 等^[68]在蓝晶盐单胞菌 (*Halomonas bluephagenesis*) 中构建了一套便携式、可自我维持的 DNA 存储系统。该系统不依赖于任何专业人员或复杂设备, 可以在家庭或野外环境使用。在携带数据的 DNA 分子对宿主细胞的稳定性与活性影响方面, Zhou 等^[69]对导入酵母菌体内的数据携带染色体进行了一系列(包括遗传、表观遗传、转录组和蛋白质组学)生化分析。结果表明, 数据携带染色体能够形成高染色质可及性, 具有高水平的 H3K4me3 甲基化修饰, 其存在对野生型酵母的染色质状态没有显著影响; 虽然数据携带染色体具有高水平的转录活性, 但其转录本均为非编码 RNA。此外, Hao 等^[34]在构建的大肠杆菌分布式混菌存储系统中对每个细胞携带

的 DNA 片段都进行了生物信息学分析, 以此避免潜在的启动子序列。细胞稳定的传代表明, 包含数据的 DNA 序列对宿主细胞的影响不大, 即使是携带有毒序列的寡核苷酸, 在大肠杆菌分布式混菌存储系统中也不会干扰其他序列。

3.4 数据擦除

在进行数据迁移或者更换存储设备时, 往往需要清除陈旧设备上的数据, 以确保敏感信息的安全性。常规硅基存储系统更多的是在信息层面对原有数据进行基于特殊算法的数据覆盖, 使其永久无法恢复; 而在 DNA 存储系统中, 用紫外线照射、高温、氧化剂等物理化学方法对 DNA 分子进行摧毁是最直接的数据擦除方式, 但这些方式需要专业的设备, 并且往往会有清除残留。2020 年, Kim 等^[70]提出了基于杂交亚稳态策略的 DNA 存储系统, 该系统能够通过加热实现快

速永久的信息擦除。在该策略中,研究人员将真假信息编码混合在 DNA 溶液中,其中,携带真实信息的序列能够 and 预先设计好的真值标记链进行杂交,而携带无关信息的 DNA 分子则会因为结合假标记链而无法延伸扩增。携带真实信息的杂交分子具有温敏性,在 25 °C 的条件下能够存储 65 天,并稳定读取;而在 95 °C 的高温下,只需 5 min,真实信息就会因杂交分子的热解离而被抹去。杂交亚稳态策略的设计展示出特异的生物反应过程在数据擦除方面的独特优势。除此之外,Wang 等^[71]开发出基于 DNA 链置换反应的单指令多数据 DNA 计算范式,在增强 DNA 存储系统数据修改能力的同时,也实现了数据擦除。值得一提的是,并非所有基于生物反应过程的数据擦除方式都能够彻底清除 DNA 分子。Lin 等^[72]构建的动态操作与可重用信息存储系统虽然能够在实现 DNA 数据库可重复信息访问的同时完成数据的擦除,但是该系统的擦除只是给目标序列打

上了擦除的标记,序列包含的信息依然完整。

3.5 恶意攻击检测

作为 DNA 存储系统的数据入口和出口,DNA 合成和 DNA 测序是外来恶意攻击的主要入侵目标。早在 2017 年就有研究人员对当时的 DNA 测序流程和生物信息学分析管道进行了信息安全评估,并通过实验发现,可以将计算机木马程序编码到合成的 DNA 链中^[73]。2020 年,Puzis 等^[74]报告了一种黑客对生物网络安全的攻击方式,如图 5 所示:研究人员 Alice 从 Bob 的 DNA 合成公司订购序列,而 John 是一名网络黑客,他通过木马软件将 Alice 的全部或者部分订单序列替换成了经过 DNA 混淆技术处理的包含所有必要成分的恶意序列(能够编码病原微生物或者有害毒蛋白),因此,在合成前,Bob 检测不到恶意序列。序列合成后,木马软件会对 Alice 的测序比对结果进行干扰,使得恶意序列能够顺利进入研究下游阶段实施破坏。

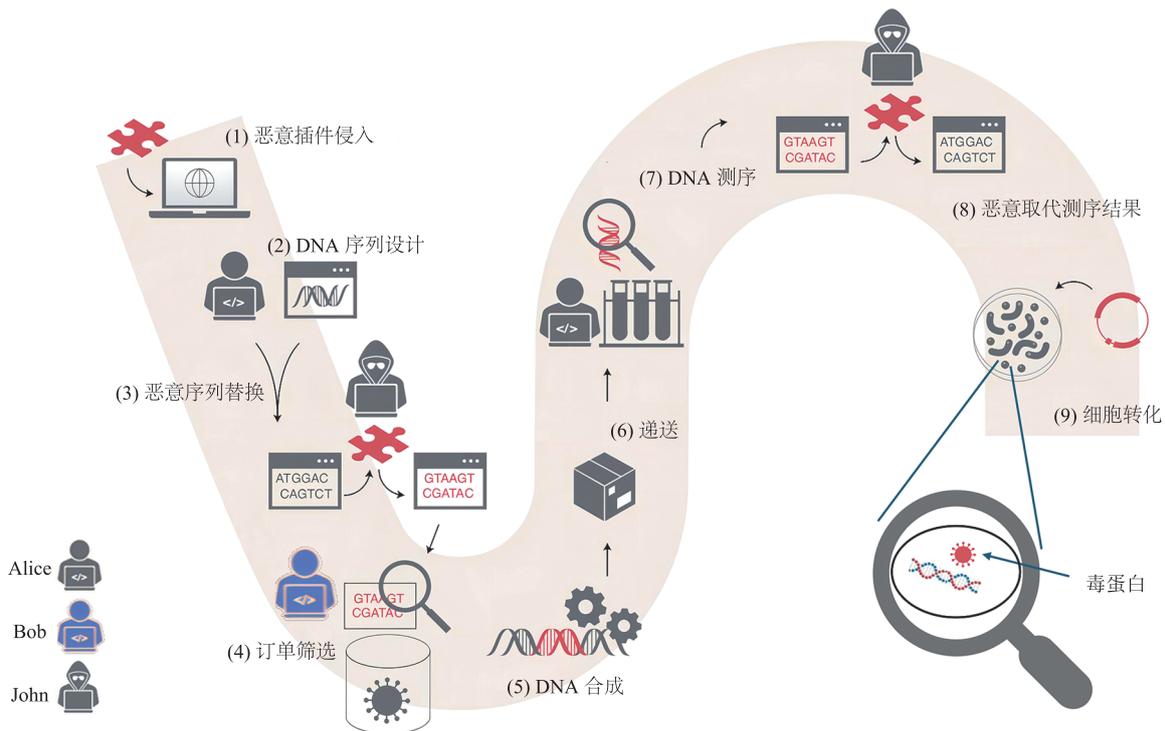


图 5 一种生物网络安全攻击^[74]

Fig. 5 A bio-cyber security attack^[74]

这些利用生物安全漏洞的攻击对 DNA 存储系统是致命的, 计算机木马程序会在测序阶段带来极大的数据安全风险, 而能够编码有害毒蛋白的恶意序列甚至可能会导致体内保存的活细胞数据库崩溃。因此, 对计算机存储系统和 DNA 存储系统之间的数据传输管道进行安全性检测或者设立防火墙是十分必要的。近年来, 随着人工智能领域的快速发展, 神经网络和机器学习算法在检测恶意攻击方面展现出了非常大的潜力^[75-77]。例如, Islam 等^[75]使用基于深度学习的训练模型对经过测序仪读取的 DNA 序列进行评估, 成功地检测和识别出其中的恶意编码序列, 并且检测的准确率接近 100%。

4 DNA 存储技术的数据安全前沿性挑战

当前, 作为 DNA 存储技术持续而稳定地发挥能效的重要保障, 数据安全的相关研究在取得一定成果的同时也面临着挑战。应对这些挑战, DNA 存储技术的数据安全保护程度将进一步被提升。

4.1 数据自身安全挑战

现阶段, 数据加密和数据隐藏作为 DNA 存储技术中保护数据信息不被泄露的两种重要策略, 主要依赖于生物技术中复杂的生物反应过程。这些生物反应过程往往涉及更为复杂的分子动力学行为, 与分子的性质及实验条件等因素关系密切, 因此, 实现对生物反应过程快速而准确地破解非常困难。在 DNA 存储中, 基于信息技术的数据加密和数据隐藏的相关研究虽然较多^[78], 但是当前的基于生物技术的 DNA 数据加密和数据隐藏研究仍处在探索阶段, 存在着生物反应过程较为固定单薄的挑战, 缺乏多类型适配的加密隐藏模式。一旦对手熟悉或者精通该生物反应体系, 那么, 数据自身的安全性将会受到严重的威胁。另外, 考虑到未来 DNA 存储技术的广泛应用,

类比计算机存储系统开发出属于 DNA 存储系统的身份认证环节能够极大地提升存储数据的自身安全程度。

4.2 数据存储安全挑战

在当前的数据存储安全范畴中, 多场景通用特性的实现是挑战, 也是追求的目标。在数据错误恢复方面, 主要使用的纠错码种类多样, 但是不同类型的纠错码各有优劣; 此外, 在体外存储保护材料的选择上, 常规的无机保护材料(如二氧化硅和氧化铁)虽然容易制造, 能够保护 DNA 分子免受高温、氧化等环境影响, 显著提高 DNA 分子的稳定性, 但是其装载密度较低, 目前为止, 二氧化硅的最高装载质量分数为 3.4%^[61], 并且回收嵌入无机保护材料的 DNA 分子的过程较为复杂。新兴的有机高分子保护材料(如 TRFG 水凝胶及天然纤维素)虽然在装载密度和回收技术层面优于无机材料, 但在面对剧烈的环境侵蚀时, 其性能却远不如前者, 部分合成材料的制作周期也比较长; 同样地, 在体内存储中, 不同类型的宿主细胞在面对不同温度、酸碱度的环境时, 无法实现稳定的生长繁殖。尽管有研究表明, 在死去的细胞中, 数据依旧能够被保存^[79], 但是不具备低成本、高保真度的数据备份方式的体内存储丧失了竞争优势。总体来说, 现阶段的数据弹性保护方式都具有自身的局限性。面对未来多场景转换的数据存储安全要求, 从数据错误恢复、体外存储的保护材料及体内存储的宿主细胞 3 个层面设计通用化的 DNA 存储保护体系是非常有必要的: (1) 开发文件通用的、能够平衡冗余度和数据准确度的纠错码; (2) 可以着力探索适配于多场景 DNA 存储的无机有机结合保护材料; (3) 在寻找新型宿主细胞的同时, 可以利用现有的生物技术对现有细胞进行改造, 提升宿主细胞在不同环境下的稳定性。

4.3 数据传输安全挑战

近年来, 作为 DNA 存储系统与外界交互传

输信息的两个接口, DNA 合成与 DNA 测序在以深度学习为代表的人工智能算法的加持下实现了对恶意序列的精准检测, 计算机存储系统与 DNA 存储系统间的数据传输安全得到了有效保障^[75-77]。然而, 在 DNA 存储系统之间, 数据传输安全的完善仍面临一些限制。首先, 现阶段适用于 DNA 存储系统的数据擦除工具较少, 亟需开发新型的或者基于生物反应过程的数据擦除工具, 以实现数据传输过程的无痕性。其次, 在达到无痕性要求的基础上, 如何在以 DNA 分子载体迁移的方式进行数据传输的过程中保证数据的完整性。以体外存储为例, 一般来说, 可以通过 PCR 扩增进行数据备份, 以减少寡核苷酸库在转移过程中的损耗。但是, 对于特殊情况来说, 如前文阐述的杂交亚稳态策略^[70], PCR 扩增因过程中产生的高温会使得数据全部损毁而不再适用。最后, 在将计算机作为数据中转站进行多个 DNA 存储系统间的数据传输的情况下, 基于深度学习算法的数据传输安全保护方法可能并不完全适用, 相关训练模型需要进一步调整优化。总之, DNA 存储系统间的数据传输安全仍需要科研人员进行较为深入的研究探索。

5 展望

互联网和全球化的快速发展改变了人们的生产生活方式。在 2023 全球移动宽带论坛上, 华为发布了全球首个全系列 5.5G 产品解决方案, 这标志着更强大、更高效的信息技术已经问世。不断迭代更新的科学技术在给人们带来便利的同时, 也加快了全球数据量的指数级增长。作为具有存储密度大、存储时间长、能源消耗低等优势的一代存储介质, DNA 有望替代传统硅基存储介质, 并化解数据量与存储空间之间的矛盾, 成为改变数据存储现状的先行利器。

作为一个新兴的多学科深度融合交叉的研究

领域, DNA 存储技术具有非常强的发展潜力, 其长期持久地发挥存储能效离不开完善的数据安全保障。近年来的研究虽然取得了一些进展, 但 DNA 存储技术的数据安全仍面临诸多考验。幸运的是, 近期报道的一些先进技术为 DNA 存储技术的数据安全研究拓宽了途径。其中, 人工合成的镜像 DNA 分子^[80]和非天然的 DNA 分子^[81]不但和天然 DNA 分子拥有同样的高存储密度, 而且具有独特的生物正交性。这不仅从本质上直接提高了信息载体分子的稳定性, 更重要的是开创了天然 DNA 分子与镜像 DNA 分子或者非天然 DNA 分子混合的加密隐藏模式。另外, 基于切口酶和 DNA 序列可编程性的双因子分子访问控制装置的开发^[82]使得 DNA 存储系统实现身份认证这一重要的数据安全环节成为可能。除了阶段性的技术成果外, 从存储系统的底层逻辑出发, 不难看出, 当前的计算机存储系统在数据安全领域的庞大而完善的数据安全生态离不开标准化、规范化的文件系统支撑。因此, 发展标准化的 DNA 数据存储体系是迈向未来安全可靠的 DNA 存储时代的必由之路。

DNA 存储技术距离真正能够服务于大众的目标尚有一定的距离, 相关的数据安全研究也任重道远, 但其潜在的研究价值和发展前景已经受到诸多研究人员的青睐。相信随着各学科的协调发展和各路人才的集思广益, 在广角度多层次数据安全的保驾护航中, DNA 存储, 这个诞生于 20 世纪科学家脑海中的灵感终将会在 21 世纪科研人员的努力下快速形成多方面的技术突破, 并在多场景的应用中展现出颠覆性的存储模式改变。

参考文献

- [1] 沈玥, 陈芳, 汪亮, 等. DNA 存储蓝皮书 [Z/OL]. <https://www.dydata.io/datastore/detail/2168976911718551552/>.

- Shen Y, Chen F, Wang L, et al. The blue book of DNA storage [Z/OL]. <https://www.dydata.io/datastore/detail/2168976911718551552/>.
- [2] IDC. High data growth and modern applications drive new storage requirements in digitally transformed enterprises [EB/OL]. (2022-07-01) [2023-10-30]. <https://uat-prev.delltechnologies.com/asset/pt-pt/products/storage/industry-market/h19267-wp-idc-storage-reqs-digital-enterprise.pdf>.
- [3] Ceze L, Nivala J, Strauss K. Molecular digital data storage using DNA [J]. *Nature Reviews Genetics*, 2019, 20(8): 456-466.
- [4] Davis J. Microvenus [J]. *Art Journal*, 1996, 55(1): 70-74.
- [5] Church GM, Gao Y, Kosuri S. Next-generation digital information storage in DNA [J]. *Science*, 2012, 337(6102): 1628.
- [6] Erlich Y, Zielinski D. DNA Fountain enables a robust and efficient storage architecture [J]. *Science*, 2017, 355(6328): 950-954.
- [7] Goldman N, Bertone P, Chen SY, et al. Towards practical, high-capacity, low-maintenance information storage in synthesized DNA [J]. *Nature*, 2013, 494(7435): 77-80.
- [8] 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.
- [9] Blawat M, Gaedke K, Hütter I, et al. Forward error correction for DNA data storage [J]. *Procedia Computer Science*, 2016, 80: 1011-1022.
- [10] 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(4): 234-242.
- [11] Bornholt J, Lopez R, Carmean DM, et al. Toward a DNA-based archival storage system [J]. *IEEE Micro*, 2017, 37(3): 98-104.
- [12] Organick L, Ang SD, Chen YJ, et al. Random access in large-scale DNA data storage [J]. *Nature Biotechnology*, 2018, 36(3): 242-248.
- [13] Banal JL, Shepherd TR, Berleant J, et al. Random access DNA memory using Boolean search in an archival file storage system [J]. *Nature Materials*, 2021, 20(9): 1272-1280.
- [14] Bögels BWA, Nguyen BH, Ward D, et al. DNA storage in thermoresponsive microcapsules for repeated random multiplexed data access [J]. *Nature Nanotechnology*, 2023, 18(8): 912-921.
- [15] Zhang YN, Wang F, Chao J, et al. DNA origami cryptography for secure communication [J]. *Nature Communications*, 2019, 10(1): 5469.
- [16] Li SY, Liu JK, Zhao GP, et al. CADS: CRISPR/Cas12a-assisted DNA steganography for securing the storage and transfer of DNA-encoded information [J]. *ACS Synthetic Biology*, 2018, 7(4): 1174-1178.
- [17] Takahashi CN, Nguyen BH, Strauss K, et al. Demonstration of end-to-end automation of DNA data storage [J]. *Scientific Reports*, 2019, 9(1): 4998.
- [18] Xu CT, Ma B, Gao ZL, et al. Electrochemical DNA synthesis and sequencing on a single electrode with scalability for integrated data storage [J]. *Science Advances*, 2021, 7(46): eabk0100.
- [19] 强薇, 沈玥, 戴俊彪. DNA 信息存储的机遇与挑战 [J]. *生命科学*, 2021, 33(12): 1469-1475. Qiang W, Shen Y, Dai JB. The opportunities and challenges in DNA digital data storage [J]. *Chinese Bulletin of Life Sciences*, 2021, 33(12): 1469-1475.
- [20] Jacques B, Marthe C, Delphine C, et al. Chain and conformation stability of solid-state DNA: implications for room temperature storage [J]. *Nucleic Acids Research*, 2010, 38(5): 1531-1546.
- [21] Shrivastava S, Badlani R. Data storage in DNA [J]. *International Journal of Electrical Energy*, 2014, 2(2): 119-124.
- [22] Organick L, Chen YJ, Ang SD, et al. Probing the physical limits of reliable DNA data retrieval [J]. *Nature Communications*, 2020, 11(1): 616.
- [23] Dong YM, Sun FJ, Ping Z, et al. DNA storage: research landscape and future prospects [J]. *National Science Review*, 2020, 7(6): 1092-1107.
- [24] Allentoft ME, Collins M, Harker D, et al. The half-life of DNA in bone: measuring decay kinetics

- in 158 dated fossils [J]. *Proceedings Biological Sciences*, 2012, 279(1748): 4724-4733.
- [25] Kjaer KH, Pedersen MW, Sanctis BD, et al. A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA [J]. *Nature*, 2022, 612(7939): 283-291.
- [26] Extance A. How DNA could store all the world's data [J]. *Nature*, 2016, 537(7618): 22-24.
- [27] Masanet E, Shehabi A, Lei N, et al. Recalibrating global data center energy-use estimates [J]. *Science*, 2020, 367(6481): 984-986.
- [28] 毕昆, 顾万君, 陆祖宏. DNA 存储中的编码技术 [J]. *生物信息学*, 2020, 18(2): 76-85.
Bi K, Gu WJ, Lu ZH. Coding algorithms in DNA storage [J]. *Chinese Journal of Bioinformatics*, 2020, 18(2): 76-85.
- [29] 江湘儿, 王勇, 沈玥. DNA 合成技术与仪器研发进展概述 [J]. *集成技术*, 2021, 10(5): 80-95.
Jiang XE, Wang Y, Shen Y. The review of DNA synthesis technologies and instruments development [J]. *Journal of Integration Technology*, 2021, 10(5): 80-95.
- [30] Zhu B, Furuki T, Okuda T, et al. Natural DNA mixed with trehalose persists in B-form double-stranding even in the dry state [J]. *The Journal of Physical Chemistry B*, 2007, 111(20): 5542-5544.
- [31] Howlett SE, Castillo HS, Gioeni LJ, et al. Evaluation of DNASTable™ for DNA storage at ambient temperature [J]. *Forensic Science International: Genetics*, 2014, 8(1): 170-178.
- [32] Organick L, Nguyen BH, McAmis R, et al. An empirical comparison of preservation methods for synthetic DNA data storage [J]. *Small Methods*, 2021, 5(5): 2001094.
- [33] Liu XG, Jing XX, Liu P, et al. DNA framework-encoded mineralization of calcium phosphate [J]. *Chem*, 2020, 6(2): 472-485.
- [34] Hao M, Qiao HY, Gao YM, et al. A mixed culture of bacterial cells enables an economic DNA storage on a large scale [J]. *Communications Biology*, 2020, 3(1): 416.
- [35] Chen WG, Han MZ, Zhou JT, et al. An artificial chromosome for data storage [J]. *National Science Review*, 2021, 8(5): nwab028.
- [36] 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.
- [37] Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase [J]. *Journal of Molecular Biology*, 1975, 94(3): 441-448.
- [38] Shendure J, Ji H. Next-generation DNA sequencing [J]. *Nature Biotechnology*, 2008, 26: 1135-1145.
- [39] Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies [J]. *Nature Reviews Genetics*, 2016, 17(6): 333-351.
- [40] Ardui S, Ameer A, Vermeesch JR, et al. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics [J]. *Nucleic Acids Research*, 2018, 46(5): 2159-2168.
- [41] Deamer D, Akeson M, Branton D. Three decades of nanopore sequencing [J]. *Nature Biotechnology*, 2016, 34(5): 518-524.
- [42] Leier A, Richter C, Banzhaf W, et al. Cryptography with DNA binary strands [J]. *Biosystems*, 2000, 57(1): 13-22.
- [43] Xiao GZ, Lu MX, Qin L, et al. New field of cryptography: DNA cryptography [J]. *Chinese Science Bulletin*, 2006, 51(12): 1413-1420.
- [44] Yang J, Ma JJ, Liu S, et al. A molecular cryptography model based on structures of DNA self-assembly [J]. *Chinese Science Bulletin*, 2014, 59(11): 1192-1198.
- [45] Zhu EQ, Luo XH, Liu CJ, et al. An operational DNA strand displacement encryption approach [J]. *Nanomaterials*, 2022, 12(5): 877.
- [46] Liu YY, Ren YB, Li JJ, et al. *In vivo* processing of digital information molecularly with targeted specificity and robust reliability [J]. *Science Advances*, 2022, 8(31): eabo7415.
- [47] 咎乡镇, 姚翔宇, 许鹏, 等. DNA 存储中的纠错方法综述 [J]. *广州大学学报(自然科学版)*, 2021, 20(2): 13-22.

- Zan XZ, Yao XY, Xu P, et al. A survey on error correcting algorithms in DNA storage [J]. *Journal of Guangzhou University (Natural Science Edition)*, 2021, 20(2): 13-22.
- [48] Wang SP, Mao XH, Wang F, et al. Data storage using DNA [J]. *Advanced Materials*, 2024, 36(6): 2307499.
- [49] Gehani A, Labean T, Reif J. DNA-based cryptography [M] // Winfree E, Gifford DK. *DNA Based Computers V*. American Rhode Island: AMS DIMACS, 1999: 233. <http://archive.dimacs.rutgers.edu/Volumes/Vol54.html>.
- [50] Chen J. A DNA-based, biomolecular cryptography design [C] // *Proceedings of the 2003 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2003: 822-825.
- [51] Grass RN, Heckel R, Dessimoz C, et al. Genomic encryption of digital data stored in synthetic DNA [J]. *Angewandte Chemie International Edition*, 2020, 59(22): 8476-8480.
- [52] Clelland CT, Risca V, Bancroft C. Hiding messages in DNA microdots [J]. *Nature*, 1999, 399(6736): 533-534.
- [53] 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.
- [54] 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.
- [55] Song LF, Geng F, Gong ZY, et al. Robust data storage in DNA by de Bruijn graph-based de novo strand assembly [J]. *Nature Communications*, 2022, 13: 5361.
- [56] Zhang HL, Lan ZJ, Zhang WW, et al. SPIDER-WEB generates coding algorithms with superior error tolerance and real-time information retrieval capacity [Z/OL]. *arXiv Preprint*, arXiv: 2204.02855, 2023.
- [57] Ding LL, Wu SG, Hou ZH, et al. Improving error-correcting capability in DNA digital storage via soft-decision decoding [J]. *National Science Review*, 2023, 11(2): nwad229.
- [58] Ping Z, Ma DZ, Huang XL, et al. Carbon-based archiving: current progress and future prospects of DNA-based data storage [J]. *GigaScience*, 2019, 8(6): giz075.
- [59] 周廷尧, 罗源, 蒋兴宇. DNA 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.
- [60] Kohll AX, Antkowiak PL, Chen WD, et al. Stabilizing synthetic DNA for long-term data storage with earth alkaline salts [J]. *Chemical Communications(Cambridge, England)*, 2020, 56(25): 3613-3616.
- [61] 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.
- [62] Koch J, Kerl AC, Schawalder N, et al. Preserving DNA in biodegradable organosilica encapsulates [J]. *Langmuir*, 2022, 38(37): 11191-11198.
- [63] Ma JY, Wang KX, Ma M, et al. Magnetic microsphere/silica nanoparticle composite structures for switchable DNA storage [J]. *ACS Applied Nano Materials*, 2022, 5(10): 15619-15628.
- [64] Fei ZJ, Gupta N, Li MJ, et al. Toward highly effective loading of DNA in hydrogels for high-density and long-term information storage [J]. *Science Advances*, 2023, 9(19): eadg9933.
- [65] Liu Q, Wei YN, Wang ZG, et al. Sustainable DNA data storage on cellulose paper [J]. *Small Methods*, 2023, 7(9): 2201610.
- [66] Sun J, Wang QM, Diao WY, et al. Digital information storage on DNA in living organisms [J]. *Medical Research Archives*, 2019, 7(6): 6.
- [67] Liu F, Li JS, Zhang TZ, et al. Engineered spore-forming bacillus as a microbial vessel for long-term DNA data storage [J]. *ACS Synthetic Biology*, 2022, 11(11): 3583-3591.

- [68] Sun FJ, Dong YM, Ni M, et al. Mobile and self-sustained data storage in an extremophile genomic DNA [J]. *Advanced Science*, 2023, 10(10): 2206201.
- [69] Zhou JT, Zhang C, Wei R, et al. Exogenous artificial DNA forms chromatin structure with active transcription in yeast [J]. *Science China Life Sciences*, 2021, 65: 851-860.
- [70] Kim J, Bae JH, Baym M, et al. Metastable hybridization-based DNA information storage to allow rapid and permanent erasure [J]. *Nature Communications*, 2020, 11: 5008.
- [71] Wang B, Wang SS, Chalk C, et al. Parallel molecular computation on digital data stored in DNA [J]. *Proceedings of the National Academy of Sciences*, 2023, 120(37): e2217330120.
- [72] Lin KN, Volkel K, Tuck JM, et al. Dynamic and scalable DNA-based information storage [J]. *Nature Communications*, 2020, 11(1): 2981.
- [73] Ney P, Koscher K, Organick L, et al. Computer security, privacy, and DNA sequencing: compromising computers with synthesized DNA, privacy leaks, and more [C] // *Proceedings of the 26th USENIX Security Symposium*, 2017: 765-779.
- [74] Puzis R, Farbiash D, Brodt O, et al. Increased cyber-biosecurity for DNA synthesis [J]. *Nature Biotechnology*, 2020, 38(12): 1379-1381.
- [75] Islam MS, Ivanov S, Awan H, et al. Using deep learning to detect digitally encoded DNA trigger for Trojan malware in bio-cyber attacks [J]. *Scientific Reports*, 2022, 12(1): 9631.
- [76] Bae H, Min S, Choi HS, et al. DNA privacy: analyzing malicious DNA sequences using deep neural networks [J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2022, 19(2): 888-898.
- [77] Tavella F, Giaretta A, Conti M, et al. A machine learning-based approach to detect threats in bio-cyber DNA storage systems [J]. *Computer Communications*, 2022, 187: 59-70.
- [78] 董校生, 王雷, 王燕, 等. DNA 密码学研究进展与应用 [J]. *广州大学学报(自然科学版)*, 2020, 19(2): 48-56.
- [78] Dong XS, Wang L, Wang Y, et al. Research progress and application of DNA cryptography [J]. *Journal of Guangzhou University (Natural Science Edition)*, 2020, 19(2): 48-56.
- [79] Ham TS, Lee SK, Keasling JD, et al. Design and construction of a double inversion recombination switch for heritable sequential genetic memory [J]. *PLoS One*, 2008, 3(7): e2815.
- [80] Fan CY, Deng Q, Zhu TF. Bioorthogonal information storage in L-DNA with a high-fidelity mirror-image *Pfu* DNA polymerase [J]. *Nature Biotechnology*, 2021, 39(12): 1548-1555.
- [81] Hoshika S, Leal NA, Kim MJ, et al. Hachimoji DNA and RNA: a genetic system with eight building blocks [J]. *Science*, 2019, 363(6429): 884-887.
- [82] Zhang XK, Liu Y, Wang B, et al. Biomolecule-driven two-factor authentication strategy for access control of molecular devices [J]. *ACS Nano*, 2023, 17(18): 18178-18189.