

· 前沿基础研究 ·

冯森 博士，中国科学院天津工业生物技术研究所副研究员，中国科学院青年促进会会员。2011年博士毕业于南开大学，同年就职于中国科学院天津工业生物技术研究所，主要从事DNA合成技术开发与应用。主持和参与国家重点研发计划、中科院重点部署项目、中科院科研装备研制项目、国家自然科学基金、天津科委重点项目等十余项各类课题。



工业生物技术中DNA合成发展现状及展望

冯森^{1,2}，王丽娜^{1,2}，汪保卫^{1,2}，田会娟^{1,2}，白雪莲^{1,2}，刘洋^{1,2}，夏海容^{1,2}

1 中国科学院天津工业生物技术研究所，天津 300308

2 国家合成生物技术创新中心，天津 300308

冯森，王丽娜，汪保卫，田会娟，白雪莲，刘洋，夏海容. 工业生物技术中DNA合成发展现状及展望. 生物工程学报, 2022, 38(11): 4115-4131.

FENG M, WANG LN, WANG BW, TIAN HJ, BAI XL, LIU Y, XIA HR. Current status and prospect of DNA synthesis in industrial biotechnology. Chin J Biotech, 2022, 38(11): 4115-4131.

摘要：DNA合成是生命科学领域的共性支撑技术和合成生物学的关键使能技术。以合成生物学为基础的工业生物技术持续快速发展，迫切需要更加便捷、经济、安全的DNA来源以满足其日益增长的大规模DNA合成需求。工业化DNA合成在通量、成本、速度等方面的优势日益凸显，有力推动了工业生物技术研发效率的提升和研发成本的下降。但是现有技术在生产过程中还存在着使用大量有机试剂、资源浪费等问题。随着DNA合成规模的持续快速提升，有毒化学品危害、成本负担、环境负担等问题日益突出。本文结合我们的工作实践，对工业生物技术中DNA合成需求、合成策略以及可持续发展面临的问题和解决方案研究进展进行探讨。

关键词：工业生物技术；DNA合成；基因合成；可持续发展

Received: August 4, 2022; **Accepted:** October 5, 2022

Supported by: Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-PTJS-002); Youth Innovation Promotion Association Chinese Academy Sciences (2019181); Fellowship of China Postdoctoral Science Foundation (2021M693355)

Corresponding author: FENG Miao. E-mail: feng_m@tib.cas.cn

基金项目：天津市合成生物提升行动项目 (TSBICIP-PTJS-002); 中国科学院青年创新促进会 (2019181); 中国博士后科学基金 (2021M693355)

Current status and prospect of DNA synthesis in industrial biotechnology

FENG Miao^{1,2}, WANG Lina^{1,2}, WANG Baowei^{1,2}, TIAN Huijuan^{1,2}, BAI Xuelian^{1,2}, LIU Yang^{1,2}, XIA Hairong^{1,2}

1 Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China

2 National Technology Innovation Center of Synthetic Biology, Tianjin 300308, China

Abstract: DNA synthesis is one of the most basic, widely-used tools in life science as well as a key enabling technology in synthetic biology. The rapid development of industrial biotechnology promoted by synthetic biology is creating an insatiable demand for large-scale DNA synthesis from more convenient, economical and safe sources. Industrial DNA synthesis platforms have remarkable advantages in terms of throughput, cost and speed. The research and development processes of industrial biotechnology benefit from these advantages, achieving a higher efficiency and lower cost. However, challenges in DNA manufacturing process remain, such as the use of large amounts of organic reagents, waste of resources and so on. With the continuous and rapid increase of DNA synthesis scale, the hazard of toxic chemicals, cost burden and environmental burden are becoming prominent. Based on our practical work on DNA synthesis, we discuss the demand and strategies for large-scale DNA synthesis in industrial biotechnology as well as the issues and potential solutions for sustainable development.

Keywords: industrial biotechnology; DNA synthesis; gene synthesis; sustainable development

脱氧核糖核酸 (deoxyribonucleic acid, DNA) 是生命体遗传信息的主要载体, 对生命体的研究、改造乃至创造都要从 DNA 入手。DNA 合成技术是按需从头构建 DNA 分子的技术。利用 DNA 合成技术能够将遗传信息的电子数据快速转化为真实的 DNA 分子, 为生命科学各领域提供寡核苷酸、基因、基因组等不同形式的 DNA 原料。

作为生命科学领域的共性支撑技术, DNA 合成技术的进步将有力推动研发效率的提升和实验成本的下降, 使无数生命科学猜想得到验证, 帮助人类不断加深对生命体的认识和理解。DNA 合成技术更是合成生物学的关键使能技术, 为合成生物学基础研究和应用领域提供大量人工设计合

成的 DNA 分子作为改造和构建生命体的起始原料。过去十余年中, 合成生物学从实验室走向产业, 已成为生物制造领域革新的核心推动力。以合成生物学为基础的工业生物技术正在快速成为“碳中和”背景下绿色制造的重要力量。工业生物技术的持续快速发展渴望更加便捷、经济、安全的 DNA 来源, 以满足其日益增长的大规模 DNA 合成需求。这些需求驱动着合成 DNA 的供给方式从实验室关注单一实验目的的少量 DNA 合成向解决共性问题的工业化 DNA 制造拓展, 推动构建自动化、智能化的 DNA 合成工厂, 以高效、低成本方式生产各种 DNA, 覆盖从简单的 DNA 元件到基因、长片段 DNA, 再到整个基因组。

1 工业生物技术中 DNA 合成需求

寡核苷酸的应用场景除最常见的聚合酶链式反应 (polymerase chain reaction, PCR) 引物外, 还包括基因合成、基因编辑、DNA 存储、药物递送、分子诊断以及 DNA 折纸等纳米材料科学领域。快速增长的基因合成需求来自飞速发展的合成生物学等新兴领域。依赖模板的经典分子生物学手段难以满足大规模构建人工设计基因的迫切需求, 但基因合成不需要任何模板且具有高度的灵活性。当模板无法获得 (如自然界中不存在的人工设计序列) 或者获取途径风险高、伤害大、费时费力 (如天然 DNA 来自传染病、珍稀动植物或极端环境) 时, 基因合成将是获得目的 DNA 分子唯一便捷可靠的途径。这些合成基因像“砖块”一样, 被用于构建大量生物元件、基因通路、代谢网络、全基因组乃至人工合成细胞。

随着合成生物技术向工业领域的快速渗透, 生物能源、生物化工和生物医药等应用领域均提出了日益迫切的大规模 DNA 合成需求。工业化 DNA 合成在通量、成本、速度和产品质量控制等方面的优势日益凸显。工业化 DNA 合成可以使研发人员不再疲于应对繁琐和重复的基础分子克隆工作, 将更多的时间和精力投入到关键核心研究工作中, 从而缩短新技术、新产品的研发周期。同时, 工业化 DNA 合成技术发展带来的低成本优势, 也使基于大规模 DNA 合成的前沿创新研究的成本负担大幅减轻, 极大地促进生命科学和相关产业链的创新突破。

下游各应用领域对于工业化 DNA 合成能力的提升抱有持续的期待。现有技术中, 化学合成寡核苷酸的单步合成效率为 98.5%–99.5%, 长度限制在 200 个碱基以内, 错误率普

遍为 0.2%–0.5%。当用于组装基因时, 寡核苷酸中的合成错误在基因片段中积累, 使基因合成错误率普遍为 0.1%–1.0%^[1]。同时, 并非所有基因序列都能被成功合成出来, 一些序列内部因素会干扰寡核苷酸的正确组装, 导致基因合成困难。2019 年, 美国工程生物学研究联盟 (Engineering Biology Research Consortium, EBRC) 发布《工程生物学: 下一代生物经济的研究路线图》^[2], 明确提出了具有可行性的 DNA 合成技术近期和远期预期突破, 以实现每分钟内快速合成 kb 级长度的寡核苷酸链, 合成效率达到 99.99% 以上, 错误率达到 0.01% 以下和快速、高保真、低成本组装 Mb 级 DNA 片段为目标, 从而能够全面满足各种下游应用的大规模 DNA 合成需求。

2 工业化 DNA 合成策略

自 1953 年 DNA 双螺旋结构被破译以来, DNA 合成研究便开始进行, 直到 20 世纪 90 年代发展出目前 DNA 合成仪上普遍使用的亚磷酸胺三酯固相合成法并实现商业化合成应用。双链 DNA 合成 (>100 bp) 可追溯到 1970 年^[3]。2000 年前后, 市场中出现了商业化基因合成服务。当时人工合成基因的价格为每碱基对约 12 美元, 合成一条长度为 1 kb 的常规基因约需 1–2 万美元^[4]。在随后的十几年中, 基因合成的成本和价格不断下降。目前, 工业化基因合成的平均成本约为最初成本的 2%, 与采用传统分子克隆方法构建目的基因的成本相比, 已具有很高的竞争力。工业化 DNA 合成的完整流程从接收序列开始到交付产品结束, 是一个涉及生物信息学、化学、分子生物学、自动化等多学科的过程, 由各个学科力量组成的团队需要齐心协力, 将整个工作流程从小规模制备转变为工业高通量制造。

2.1 标准化与自动化

实现高通量并行和高质量规模化生产的基础是标准化和自动化。标准化涉及实验步骤、试剂耗材、反应条件等多个方面。为降低自动化整合门槛,需对生产流程进行标准化分解,分解为可定义、可管理、可独立实施和控制、可重复、小而明确的操作步骤。由此可见,工业规模和实验室规模的 DNA 合成之间的差异,不仅仅体现在高通量合成技术的迭代和高性能合成装备的开发与应用,更重要的是需要针对工艺规范和标准,对技术进行合理的分解、整合和运用,对每个操作步骤进行调整,以确保:(1) 节约试剂耗材;(2) 允许高通量并行处理样品;(3) 降低系统误差,提高可靠性和可重复性。这其中涉及到方法优化、过程控制、信息管理、仪器研发和自动化整合,以及质量控制相关的一系列具体措施。这些措施的有机结合是构建并稳定运行工业化 DNA 合成平台的必要条件。

2.2 信息管理智能化

实现大规模 DNA 合成需要依靠强大的实验室信息管理系统 (laboratory information management system, LIMS) 以实现流程管理的智能化。LIMS 贯穿从订单接收、样品处理到产品交付的全链条,包括内部工作端和外部客户端两个交互的链条,具体功能包括客户 DNA 序列处理、工作清单生成和进度跟踪、数据统计(按照时间线、用户、产品类型等不同的筛选方式进行数据查询和统计)等。在生产前,LIMS 将订单信息整理分解为流程任务,对处于同一时间节点的任务进行聚合,并创建任务列表;在生产中,LIMS 不仅管理生产流程中各个环节的执行进度,还需要管理每个环节上的大量并行操作,同时跟踪记录样品流转和实验报告等过程信息,融合条形码等编码识别技术,将信息流和样品实体相匹配,从而实现对样品的精

确定位和处理环节的精确分配;在生产后,LIMS 形成完整的生产报告和样品交付报告。

2.3 通用实施策略

以下结合常规寡核苷酸和基因合成的一种通用流程,简要介绍工业化 DNA 合成的一些具体实施策略。

寡核苷酸制造通常由 4 个关键工艺步骤组成:(1) 在固相载体上合成寡核苷酸;(2) 寡核苷酸从固相载体上分离和脱保护;(3) 纯化、鉴定、定量;(4) 浓缩干燥。目前最普遍使用的是亚磷酰胺三酯固相合成法,其他方法如氢磷酸法、可溶载体合成法等在某些应用中具有相对优势,也在不断发展和优化。通常在合成结束后采用氨水等弱碱性试剂将全长产物从固相载体上切割下来,并脱去寡核苷酸链上的保护基团(核糖核酸 (ribonucleic acid, RNA) 合成产品去除 2'-O-硅烷基保护基团通常采用氟化物处理)。针对寡核苷酸不同的下游应用,可选择脱盐、聚丙烯酰胺凝胶电泳 (polyacrylamide gel electrophoresis, PAGE) 和高效液相色谱 (high performance liquid chromatography, HPLC) 等不同的纯化工艺。除寡核苷酸合成与纯化外,还需要具备生化废物管理、废气处理与检测以及实验室环境控制等相关设备和支撑条件。

基因合成通常从序列设计开始,包括序列优化和拆分两个步骤。序列优化是可选步骤,包括针对指定宿主进行密码子优化、添加或指定特征序列或模块,以及对指定序列或特征进行规避处理等。经过确认的基因序列,根据长度和复杂性(重复序列、GC 含量等),采取相应的算法拆分成寡核苷酸库,并为其匹配首末端扩增引物(包含克隆位点、同源重组位点等序列)和测序引物。该设计过程通常由设计软件自动完成,并自动评估寡核苷酸库内成员之间是否存在错配、自身杂交等不利于组装反应的因素,

如果超过某个阈值,则修改起始参数进行重新计算设计。为提高效率,设计过程通常采用多组生产参数并行计算,从中选取最佳合成方案。对于有多个子片段组装需求或亚克隆需求的合成序列,还需要从目标产品入手,采用自上而下的策略,预先设计好所有必要的拼接和亚克隆方案。选择大规模通用生产方式的条件通常是通过单轮筛选有 95% 以上的几率获得至少一个序列完全正确的片段。由于片段越长,累积错误的几率和数量越高,从而导致筛选到正确片段所需的工作量增加,因此对于子片段的长度有一定的限制要求。设计好的寡核苷酸序列以标准化的合成格式分配给寡核苷酸合成系统进行高通量并行合成,输出的产品用于组装 1–2 kb 的片段,并采用无痕克隆技术^[5]将线性 DNA 分子重组到载体上,转化宿主并培养后,选取单克隆菌落,通过菌落 PCR 测序进行序列验证。序列完全正确的片段可进一步完成组装或亚克隆,以及交付前的样品准备等操作。

3 工业化 DNA 合成可持续发展面临的主要问题

自 20 世纪 70 年代以来, DNA 合成技术取得了显著进步,但与 DNA 测序技术相比, DNA 合成能力远远落后,生命信息的“读”和“写”能力严重不匹配^[6]。同时,现有的通用 DNA 制造工艺虽然久经考验,成熟且灵活,但生产过程中存在着使用大量有机试剂、资源浪费等问题。随着 DNA 合成规模的增加,有毒化学品危害、成本负担、环境负担等问题会更加突出。

3.1 现有生产工艺与绿色合成差距较大

如前所述,目前工业化 DNA 合成工艺通常从化学合成寡核苷酸起始,更长的 DNA 分子是以寡核苷酸为原料通过酶促反应逐步拼接和组

装得到。生产过程中,除合成与组装步骤外,还包括多步产物分离、纯化和检测步骤。绿色化学旨在从源头上减少和消除污染物,提高资源利用率和降低能耗,我们对照“绿色化学 12 项原则”^[7],对 DNA 合成工艺进行评估(表 1),采用“√”来表示符合,“×”表示偏离,数量越多表示偏离度越高。

与绿色化学原则有偏离的方面主要涉及反应路线和所采用的试剂材料,从源头上减少乃至消除大规模工业合成产生的大量有机废气和废液需要持续努力。一些可行的解决方案包括:在保证产量的情况下,通过提高单体转化率、减少损失率来减少反应原料用量;对未反应的原料、溶剂、催化剂等进行回收、再生,通过重复利用来降低成本和减少废物排放;对无法回收、再生和重复使用且毒副作用、污染作用明显的原料,积极寻找完全替代方案,不再继续使用;开发具有高效性、高选择性、反应条件温和、环境友好的催化剂,如生物酶;进一步优化合成和纯化等相关载体,提高效能和重复利用率;利用计算机辅助设计和模拟来优化反应路线、推动生产步骤(特别是分离、提纯等后处理步骤)和实验过程的简化等等。

3.2 合成能力提升面临诸多挑战

提升 DNA 合成通量的传统解决方案通常是直接增加设备和劳动力数量,但从长远来看,这种解决方案不利于节能减排和降低成本。更具有可持续潜力的策略包括发展小型化、并行反应体系,尽可能提升合成效率以获得更长的 DNA,以及整合多种技术来提高合成平台的可扩展性等。

目前,基于亚磷酰胺三酯化学合成法的寡核苷酸单步合成效率虽然已高达 99.5%,但合成长度达到 200 个碱基时,产率即降至约 35%^[14],低于这一产率,杂质过多难以纯化得

表 1 DNA 合成工艺与绿色化学 12 项原则对照评估

Table 1 Assessment of the DNA synthesis process according to the 12 principles of green chemistry

No.	Principle of green chemistry	Current DNA synthesis process	Deviation
1	Prevention: It is better to prevent waste than to treat or clean up waste after it has been created.	Large amounts of organic and aqueous waste are generated during chemical synthesis of oligonucleotides.	×××
2	Atom economy: Synthetic methods should be designed to maximize incorporation of all materials used in the process into the final product.	Atom efficiency is 34%–41% ^a for monomers in standard phosphoramidite chemistry. The synthesis of a full-length gene requires a few to tens of picomoles of oligonucleotides which account for less than 1% of the yield of column-based oligonucleotide synthesis (usually tens of nanomoles) ^[8] .	×××
3	Less hazardous chemical syntheses: Wherever practicable, synthetic methods should be designed to use and generate substances that possess little or no toxicity to human health and the environment.	Almost all the synthetic reagents for chemical synthesis of oligonucleotide are hazardous, e.g., trichloroacetic acid, iodine, pyridine, acetic anhydride. The process of DNA assembly is catalyzed by enzymes in non-toxic aqueous solution.	×××
4	Designing safer chemicals: Chemical products should be designed to preserve efficacy of function while reducing toxicity.	Synthetic DNA generally has low chemical toxicity.	√
5	Safer solvents and auxiliaries: The use of auxiliary substances (e.g., solvents, separation agents, etc.) should be made unnecessary wherever possible and, innocuous when used.	Large amount of acetonitrile is used as solvent during oligonucleotide synthesis.	×××
6	Design for energy efficiency: Energy requirements should be recognized for their environmental and economic impacts and should be minimized. Synthetic methods should be conducted at ambient temperature and pressure.	Most reactions are carried out at ambient temperature and pressure. Heating or cooling is generally used for performing enzymatic reactions (e.g., PCR, ligation) and drying DNA samples.	×
7	Use of renewable feedstocks: A raw material or feedstock should be renewable rather than depleting whenever technically and economically practicable.	Solid supports for DNA synthesis or purification can be reused ^[9-11] . Some reagents and starting materials are from non-renewable sources ^[9-12] .	×
8	Reduce derivatives: Unnecessary derivatization (use of blocking groups, protection/deprotection, temporary modification of physical/chemical processes) should be minimized or avoided if possible, because such steps require additional reagents and can generate waste.	The protection of monomers used in phosphoramidite chemistry includes the exocyclic amino groups of the nucleobases, the hydroxyl group at the 5'-positions of the sugar moieties, and the hydroxyl group of phosphite.	××
9	Catalysis: Catalytic reagents (as selective as possible) are superior to stoichiometric reagents.	Activator is used to raise the coupling efficiency and rate in oligonucleotide synthesis. The process of DNA assembly is catalyzed by enzymes.	×
10	Design for degradation: Chemical products should be designed so that at the end of their function they break down into innocuous degradation products and do not persist in the environment.	Synthetic DNA is generally degraded into products that are naturally occurred.	√
11	Real-time analysis for pollution prevention: Analytical methodologies need to be further developed to allow for real-time, in-process monitoring and control prior to the formation of hazardous substances.	Sensors are generally used to detect the leakage of organic vapor. There is no real-time monitoring and controlling means used for pollution prevention purposes.	×××

(待续)

(续表 1)

No.	Principle of green chemistry	Current DNA synthesis process	Deviation
12	Inherently safer chemistry for accident prevention: Substances and the form of a substance used in a chemical process should be chosen to minimize the potential for chemical accidents, including releases, explosions, and fires.	Some reagents used in oligonucleotide synthesis are flammable, e.g., acetonitrile, pyridine, acetic anhydride.	××

^a: According to the data from Glen Research, the molecular weights (M.W.) of dA, dG, dC and dT phosphoramidites are 857.95, 839.92, 833.93 and 744.83, respectively. When the phosphoramidites are incorporated into an oligonucleotide with all the protecting groups removed, the formula weights (F.W.) of the corresponding monomer units are 313.21, 329.21, 289.18, 304.20, respectively (<https://www.glenresearch.com/reports/gr34-18>). Atom efficiency is measured as the ratio of the molecular weight of the desired product over the molecular weights of all reactants used in the reaction^[7,13]. The atom efficiency of dA, dG, dC and dT phosphoramidites, estimated by the ratio of F.W. to M.W., are 36.51%, 39.20%, 34.68% and 40.84%, respectively.

到目的寡核苷酸。如要合成kb级长度的寡核苷酸,单步合成效率必须达到99.9%以上,才能获得同样的产率。同时,要实现工业化合成,还需要兼顾合成速度的提升。开发基于不依赖模板的酶促DNA从头合成技术,利用酶促反应的高选择性和高催化活性,有望突破上述瓶颈。但目前酶促DNA从头合成技术在酶对非天然底物的特异性选择、催化效率以及可控精确合成的实现等方面还需要进行很多优化和改进才能满足工业化合成的需要。

基因合成方面,降低合成错误率和成本是主要挑战。合成错误来自寡核苷酸自身和组装过程。提高寡核苷酸的质量和长度是减少合成基因中累积错误的主要途径。基因序列中的高级结构、重复序列、GC含量异常等内部因素对DNA片段组装的干扰机理尚不完全清楚,因此对可能生成的副产物也无法完全准确预测。虽然高通量自动化平行组装已能够实现,但是仍有大量“困难基因”无法通过通用的自动化流程合成出来,需要依赖手工操作因“难”施策。如何以高通量平行的方式进行序列鉴定并从不完美的合成DNA混合物中经济、快速地提纯出全长、无错误的DNA片段也是当前面临的主要瓶颈之一,纳米孔测序等新技术的发展和应用有

可能为开发单分子分辨率下测序和纯化整合技术提供支持。

4 提高 DNA 合成可持续性的策略与进展

提高 DNA 合成可持续性方面的努力始于 20 世纪 90 年代后期,相关优化和创新工作涉及合成反应原理、原料、自动化实现方式等多方面,已取得很多有益的进展。

4.1 减少化学合成中有机试剂用量

在合成产量规模和 DNA 长度不变的情况下,减少化学合成中有机试剂用量的一种策略是减少合成循环数。在经典亚磷酰胺三酯化学合成法中,将常用的单碱基亚磷酰胺单体替换为双碱基亚磷酰胺(双联单体)^[15]或三碱基亚磷酰胺(三联单体)^[16],在合成同样长度的寡核苷酸序列时,所需进行的反应循环数可减少 1/2 或 2/3,从而显著减少试剂消耗,且副反应的发生几率也相应降低。受化学反应效率限制,采用单碱基单体合成的长度难以超越 200 nt,合成更长的寡核苷酸时,多联单体合成法更具优势。同时,三联单体中含有 3 个碱基,对应一种氨基酸密码子,对应 20 种氨基酸密码子的一系列三联单体常

用于构建突变基因文库, 满足蛋白质定向进化筛选需求^[17]。不过, 目前市售多联单体的价格相当昂贵, 并且溶解性低于单碱基单体, 限制了它们的大规模应用。

减少化学合成中有机试剂用量的另一种策略是简化合成步骤。将单体 5'羟基的保护基团二甲氧基三苯基 (dimethoxytrityl, DMT) 替换为分子量更小的芳氧基羰基 (aryloxycarbonyl), 采用兼具强亲核性与温和氧化性的过氧阴离子溶液在一步反应中完成寡核苷酸链脱保护和氧化两个步骤^[18]。合成较短寡核苷酸链时, 封闭 (capping) 步骤可省略, 从而实现将经典的“脱保护-偶联-封闭-氧化”四步合成循环简化为“偶联-脱保护并氧化”两步合成循环。该方法的优势还在于能够完全消除经典亚磷酸酰胺三酯合成法中采用酸脱保护导致的脱嘌呤突变, 这种突变是 DNA 合成中最主要的副反应。

4.2 发展微缩集成的高通量合成体系

以固相合成柱为载体的柱式合成产量通常高于 20 nmol, 对于许多下游应用来说, 是浪费合成试剂和原料的过度生产。例如, 用于基因合成的寡核苷酸用量只需几十 pmol。多余的产品要么被丢弃, 要么保存在深孔板 (体积为 1.2 mL/孔) 中长期储存, 占用宝贵的低温保藏空间。减少过度生产的直接途径是缩小固相合成柱的负载量以降低寡核苷酸产量, 但合成柱负载量缩小到 3.5–5.0 nmol 水平 (对应的通量提升到 1 536 条寡核苷酸)^[19–20]后, 受限于固相合成柱的设计和试剂投放模式, 很难进一步缩减合成产量和提升通量。

开发芯片式微缩集成的寡核苷酸合成体系是以指数级水平削减大规模寡核苷酸过剩产量、减少合成试剂浪费的一种有效途径。在一个微表面上以高密度方式设置大量 DNA 化学

合成位点, 通过设计合适的开关机制来自动控制表面上成千上万个合成位点的开启和关闭, 从而在特定区域按需引发合成反应, 实现不同序列的寡核苷酸链的高通量并行合成。目前已发展出多种“开关机制”, 如光化学法^[21–26]、喷墨打印法^[27–28]、电化学法^[29]、微流控法^[30–31]、集成电路控温法^[32]、分选^[33]等。随着底层技术的优化, 基于这些方法的自动化合成平台也在不断地改进和发展, 多篇综述文章进行了详细介绍和总结^[34–36]。采用高密度芯片合成能够将合成寡核苷酸所需的化学品消耗减少 99%以上, 合成成本比柱式合成降低几个数量级, 每个碱基的合成成本从 0.000 01 到 0.001 美元不等^[37]。

芯片合成提供了低成本的寡核苷酸池 (oligonucleotide pool), 商品化用途主要作为检测探针, 当用于基因合成时, 单个寡核苷酸池中包含的寡核苷酸过多, 会因交叉杂交干扰过多而导致很难有效地进行基因组组装。上述问题的解决, 目前主要有两种策略。第一种策略是利用条形码序列, 通过选择性 PCR 扩增出用于组装同一条基因的所有寡核苷酸, 然后通过酶切去除条形码序列, 进而组装成基因。基于这一策略的优化方法包括分层法基因合成^[38]、“散弹枪”法大片段 DNA 合成^[39]、集成高通量纠错的高保真基因合成^[40]等。第二种策略是物理隔离, 通过在芯片表面设置多个独立微池, 每个微池内的寡核苷酸子池用于合成一条基因, 使芯片表面可平行合成多条基因, 输出基因文库^[41–42]; 物理分割也可以通过液滴来实现, 修饰了条形码序列的微球将不同的寡核苷酸子池从芯片合成的寡核苷酸池中提取出来, 在皮升 (picoliter) 级“油包水”液滴内, 微球上的寡核苷酸子池被酶切割下来并组装成基因片段, 破乳后, 大量液滴释放出大规模的基因文库^[43–44]。基因文库构建是酶改造和定向进化

中的一个关键环节,获得库容足够大的候选基因库可显著提高筛选到目标酶突变体的概率。基于高通量合成技术所构建的基因库几乎没有氨基酸偏好,能够在给定基因序列的基础上,对指定位点、突变数量和密码子种类等实现完全或部分随机化,克服基于易错 PCR、DNA 混编 (DNA shuffling) 等方法构建基因文库时,难以完全控制突变方向的缺点,为高通量代谢工程、菌株进化和蛋白质工程提供高质量基因文库^[45-48]。

4.3 预防和消除合成错误

从合成寡核苷酸到采用不同组装方式合成基因,以及进一步组装成长片段的过程中,由于化学反应效率、酶的偏好性和扩增过程保真度的局限,都会不同程度地导致合成错误的发生。这些错误的存在常会对后续应用产生重要的影响。如何高效、低成本地对人工合成 DNA 进行纠错和质量验证已成为各种后续应用要解决的重要共性问题。

用作 PCR 引物、杂交探针、治疗药物的寡核苷酸,通常会根据含合成错误的寡核苷酸链与正确寡核苷酸链在长度和分子量方面的差异,采用聚丙烯酰胺凝胶电泳、高效液相等分离纯化手段去除大部分截短的错误寡核苷酸链,但对于碱基替换或部分单碱基插入/缺失错误的去除率很低。基因合成常用的错误修复方法是实施基因克隆后进行测序分析和定点突变,经常是基因合成中成本最高和最耗时的步骤。

为应对工业化高通量 DNA 合成的错误修复问题,已发展出一些可行的策略。杂交筛选策略采用完全正确的“参考寡核苷酸库”作为互补链,在严谨的杂交条件下(如高精度控温)进行杂交,可去除人工合成寡核苷酸库中的部分错误序列^[49],但由于获得完全正确的参考

寡核苷酸库本身非常困难,该策略只适用于固定序列的寡核苷酸探针文库的批量化生产纠错。荧光选择策略以表达荧光蛋白的序列为报告基因,与目的基因在载体上融合表达,只有当目的基因不含缺失或插入错误时,报告基因才能正确表达,使转化后的克隆菌落呈现荧光^[50]。长度选择策略采用通用引物与芯片上的寡核苷酸退火,以带有可切割阻断基团的脱氧核苷三磷酸 (deoxyribonucleoside triphosphates, dNTPs) 为单体,以可控加成的方式将寡核苷酸延长 n bp,然后添加生物素-dATP,只有不含缺失或插入错误的寡核苷酸互补链才能偶联上生物素-dATP,最后利用生物素-链霉亲和素相互作用富集回收筛选出的寡核苷酸^[51]。芯片合成与二代测序相结合的方法可实现对测序正确的寡核苷酸链进行分拣回收^[52],但该技术存在效率低、设备及运行费用昂贵等问题。

基于生物酶处理的纠错技术利用错配特异性结合蛋白 (mismatch specific binding protein, MSB) 或错配特异性内切酶 (mismatch specific endonuclease, MSE) 对 DNA 双链中错配位点的特异性结合或切割来实现清除或修复合成错误的目的(图 1),具有高效快捷、可自动化、成本相对较低等优势。目前,已有研究报道的错配特异性结合蛋白主要包括来源于大肠杆菌 DNA 修复系统的 MutS 蛋白及其同源蛋白^[53-55]。有关错配特异性内切酶的研究报道较多,包括大肠杆菌核酸内切酶 V (*Escherichia coli* endonuclease V)、T4 噬菌体核酸内切酶 VII (phage T4 endonuclease VII)、T7 噬菌体核酸内切酶 I (T7 endonuclease I)、来源于芹菜的核酸内切酶 CEL I/II、米曲霉核酸酶 S1 (nuclease S1 from *Aspergillus oryzae*)、绿豆核酸酶 (mung bean nuclease)等^[1,56-57]。

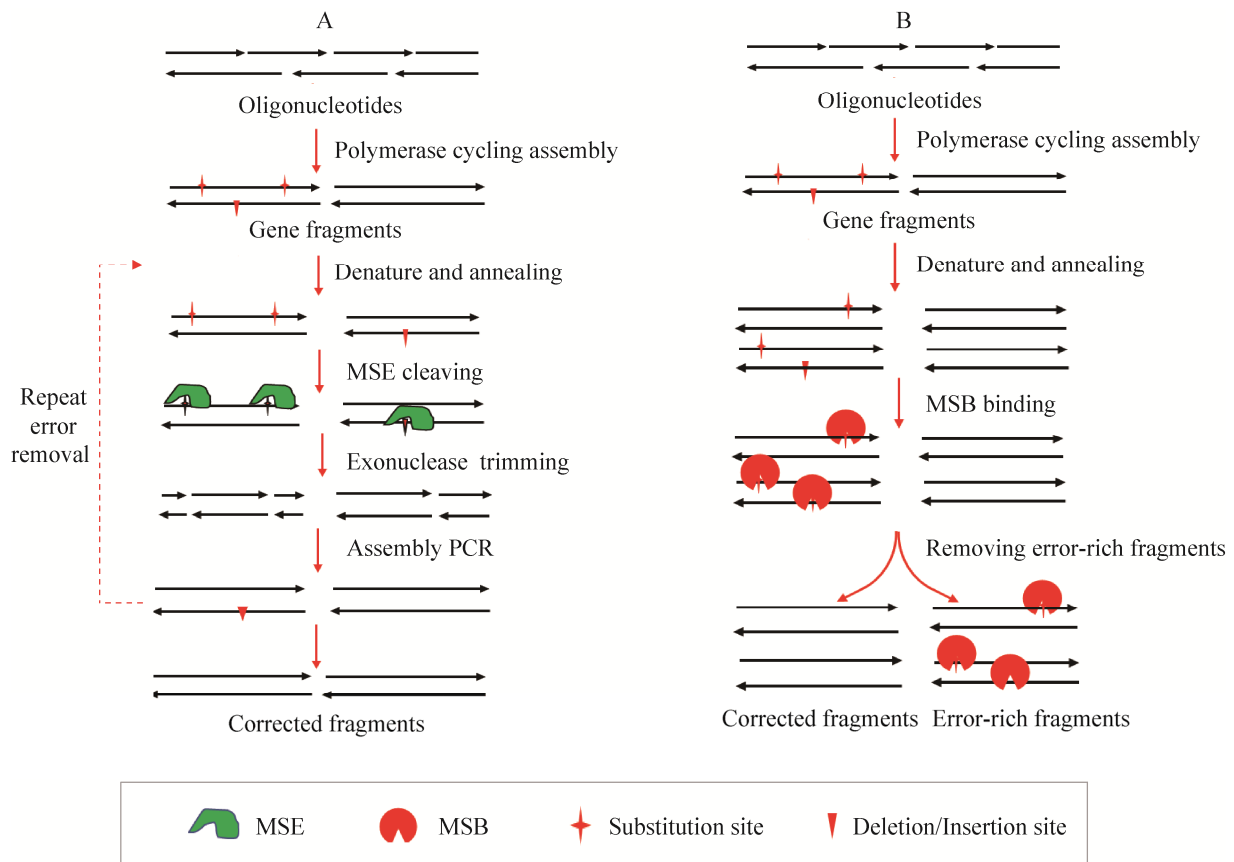


图1 错配特异性内切酶 (A) 与错配特异性结合蛋白 (B) 纠错反应流程

Figure 1 Working scheme of mismatch correction with mismatch-specific endonuclease (MSE) (A) and mismatch-specific binding protein (MSB) (B).

人工合成基因组，特别是较大型的真核生物基因组合成过程相当复杂，其合成错误的纠正需要采取新的思路和策略，包括采用多层次协同作用的纠错策略。这种多层次协同作用的纠错策略包括在寡核苷酸合成阶段采用严谨的合成工艺参数条件，在基因组组装阶段加入基于生物酶处理的纠错过程以提高用于人工基因组合成的基础元件的质量，在大片段拼接阶段采用高保真的重组体系，在基因组拼接和完工阶段采用快速准确的错误定位方法实现对错误或漏洞 (bug) 的修复^[58]。

为解决 DNA 存储过程中信息解码错误的问题，从信息科学的角度开发了多种编码和解

码策略来进行纠错以降低人工合成 DNA 的错误对 DNA 存储质量的影响，属于一种广义的合成 DNA 纠错方式^[59-61]，与其他基础的纠错技术一起共同服务于 DNA 存储技术开发。

4.4 生物酶催化合成

化学法合成 DNA 的长度难以超过 200 nt，且需要使用无水溶剂，并产生有毒废物。此外，使用有机酸脱保护易导致脱嘌呤突变^[37]。随着合成生物学的发展，利用合成生物学策略和技术设计构建环境友好的高效生物酶催化路线来替代化学合成路线，有望从源头上解决 DNA 合成瓶颈问题。

酶促 DNA 合成的概念源于 DNA 聚合酶的

发现^[62]。为了像化学合成法一样从头合成所需的 DNA 序列,需要不依赖模板的聚合酶,如多核苷酸磷酸化酶 (polynucleotide phosphorylase, PNPase)、T4 RNA 连接酶和末端脱氧核苷酸转移酶 (terminal deoxynucleotidyl transferase, TdT)。受亚磷酸胺三酯化学合成方法的启发,采用带有阻断基团的核苷酸衍生物作为单体,通过外加化学试剂按需切去阻断基团,从而提高酶催化合成 DNA 过程的可控性^[63]。在选择合适的阻断基团时,需要满足以下要求:(1) 在酶促反应中具有化学稳定性;(2) 可被化学试剂去除,但不影响寡核苷酸链结构;(3) 体积足够小,不影响酶与底物结合;(4) 能够有效阻止酶继续加成新的核苷酸。

PNPase 利用带有阻断基团的核苷酸单体衍生物 2' (3')-O-(α -甲氧基乙基) 核苷 5'-二磷酸或 2' (3')-O-异戊酰逐步合成单链 DNA^[64-68],优化 Mn^{2+} 浓度等条件,使预先由化学合成得到的 4 mer 寡核苷酸延伸了 9 mer^[69]。但 PNPase 的一个关键缺点是随着寡核苷酸链的延长,产物中磷酸盐积累增多,会催化发生逆反应,使寡核苷酸链降解^[70]。

T4 RNA 连接酶能够将 2'-脱氧核糖核苷 3',5'-二磷酸连接到寡核苷酸的 3'羟基上,寡核苷酸链的 3'-磷酸能够阻断链的进一步延伸^[71]。但是优化后合成反应时间仍需要数天,并且合成效率因底物类型而不同,也会有自连等副反应发生^[72-74]。

TdT 属于 DNA 聚合酶 X 家族^[75]。目前,利用 TdT 合成 DNA 的方法主要有两种策略。一种是竞争合成策略,即采用酶或离子笼分子在一定程度上阻断天然 TdT 加成天然 dNTPs 的活性,从而控制加成核苷酸单体的数量。其中,一种方法是 TdT 与三磷酸腺苷二磷酸水解酶 (apyrase) 竞争结合核苷酸单

体。Apyrase 可将 dNTPs 降解为脱氧核苷二磷酸 (deoxyribonucleoside diphosphates, dNDPs) 或脱氧核苷单磷酸 (deoxyribonucleoside monophosphates, dNMPs)。由于 TdT 对 dNTPs 的活性高于 dNDPs,这种方法控制 TdT 加成核苷酸的数量平均为 3-5 个碱基^[76-77]。与上述方法类似,采用螯合剂 1-(4,5-二甲氧基-2-硝基苯基)-1,2-二氨基乙烷-N,N',N',N'-四乙酸 (1-(4,5-dimethoxy-2-nitrophenyl) ethylene diamine tetraacetic Acid, DMNP-EDTA) 与 TdT 竞争结合 Co^{2+} 离子^[61]。在反应开始之前, Co^{2+} 离子被 DMNP-EDTA 笼住,不能被 TdT 结合。紫外光照后,笼状分子结构被破坏,释放出的 Co^{2+} 离子与 TdT 结合,启动 dNTPs 加成反应。该调控机制是不可逆的,通过添加过量的离子笼分子来终止反应,形成均聚物的长度取决于紫外光照强度和照射时间。

上述竞争方法不能确保按需加成核苷酸,不可避免产生均聚物。采用带有阻断基团的核苷酸作为单体,能够进一步提高 DNA 合成的可控性。除加入化学试剂切割^[78]外,还可以采用光照裂解阻断基团^[79]。一种解决方案是 TdT 与 dNTPs 通过共价键形成交联体, TdT 酶本身充当“阻断基团”。当 TdT 酶将所连的 dNTPs 连到寡核苷酸链上后,自身能够阻止其他 dNTPs 继续加成。利用特定波长光照切去 TdT 酶后,寡核苷酸链 3'端除去阻碍,得以继续进行延伸。该方法单步合成产率为 93.4%-99.5%,平均单步产率为 97.7%^[80]。由于需要预先制备 TdT 与 dNTPs 交联体,该方法对酶的消耗较大。

使用经过修饰的非天然核苷酸作为底物,需要对 TdT 酶进行针对性的改造,不仅涉及底物结合口袋的形状、催化位点的改变,甚至会改变酶的动力学性质。尽管 TdT 对多种核苷酸都具有兼容性,但与 3'端羟基游离的天然

核苷酸相比,对 3'端羟基修饰的核苷酸的活性较差^[81]。因此,尽管该策略看似简单,但开发具有快速、精确合成能力的 TdT 和与之匹配的 3'端可逆阻断的核苷酸反应体系并不容易。国外很多团队提出了不同的解决方案^[81-85],天津工业生物所在国内率先实现技术突破,经工程化改造后的 TdT 对 3'-ONH₂-dNTPs 单体的催化活性比常用的哺乳动物 TdT 高 3 个数量级,可有效延伸寡核苷酸链,平均单步产率为 98.7%,与 DNA 化学合成法相当(图 2)^[86]。

4.5 防止 DNA 合成的滥用

DNA 合成价格的快速下降使其更易获得,在助力科技发展的同时,也增加了被滥用的可能。例如,脊髓灰质炎病毒和“西班牙流感”病毒的 DNA 分别于 2002 年和 2005 年被成功合成^[87-88],这些病原体如被用作生物武器,后果不堪设想。为防止滥用 DNA 合成技术,美国卫生与公众服务部于 2010 年发布了《合成双

链 DNA 供应商筛查框架指南》^[89],呼吁双链 DNA 的商业供应商自愿筛查所有订单。国际基因合成联盟(International Gene Synthesis Consortium, IGSC) 内的 DNA 合成公司已经根据上述指南实施了协调筛选协议(harmonized screening protocol, HSP)^[90]。虽然并非所有 IGSC 成员都使用单一的 DNA 筛选算法,但 DNA 筛选软件通常遵循 HSP 指导。生物安全评估过程包括两项任务,即评估客户的身份和合法性,以及对所有订购的基因产品的序列进行鉴定,通过在特定数据库中进行比对,确定是否与危险或受管制药物、毒物相关的序列相匹配。尽管 DNA 提供方自愿实施序列筛查程序,但随着 DNA 合成仪器的小型化和易用性提高,全球范围内获得合成 DNA 的门槛可能会进一步降低。2019 年,核威胁倡议协会(Nuclear Threat Initiative, NTI) 和世界经济论坛(World Economic Forum, WEF) 在全球组织了

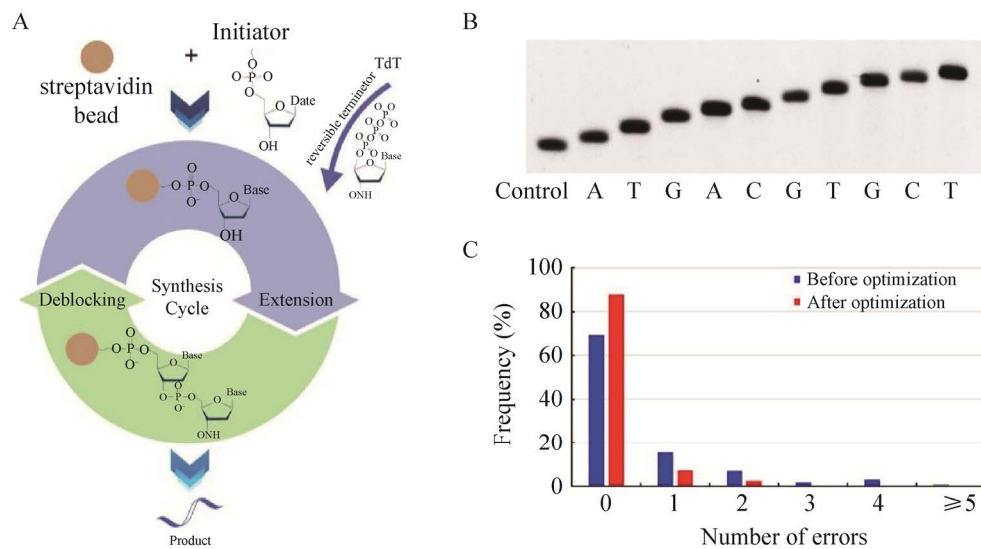


图 2 工程化改造的 TdT 在从头 DNA 合成中的应用^[86] A: 寡核苷酸二步合成循环; B: 进行 10 个合成循环后所得寡核苷酸产物的 PAGE 结果; C: 过程优化前后寡核苷酸合成策略性能评估

Figure 2 Utilization of engineered TdT in *de novo* DNA synthesis^[86]. (A) Two-step oligonucleotide synthesis cycle. (B) PAGE analysis of the 10 steps of extension in oligonucleotide synthesis. (C) Evaluation of the performance of our oligonucleotide synthesis strategy before and after process optimization.

一个防止 DNA 非法合成的国际专家工作组，并于 2020 年 1 月 9 日发布了《生物安全创新和减少风险：可获取、安全和可靠的 DNA 合成全球框架》^[91]报告，倡议用标准化筛查的方法减少相关风险。报告建议同时开展两项活动：一是由多方利益相关者组成技术联盟，开发 DNA 序列筛查的通用机制；二是成立一个新的全球实体来监督联盟，促进 DNA 序列筛查形成全球规范。

为应对涉及基因操作的前沿生物技术风险，多年来，我国也先后发布和实施了一系列有关风险防控的法律法规，包括《生物技术研究开发安全管理办法》（2017 年 7 月）、《中华人民共和国人类遗传资源管理条例》（2019 年 3 月）等^[92]。2021 年我国发布实施《中华人民共和国生物安全法》^[93]，规定了维护生物安全应当遵循的原则、领导体制、生物安全风险的防控机制和措施，并明确了相关法律责任。该法适用范围涉及重大新发突发传染病、动植物疫情、生物技术研发与应用、生物安全实验室、人类遗传资源与生物资源、外来物种入侵与生物多样性、生物恐怖袭击与生物武器威胁和微生物耐药等，为相关行政法规的制定和修订提供了上位法基础^[92]。同时，我国科技工作者在国际生物安全新规则的制定中提出“中国方案”、发出中国声音，积极融入国际立法和对话。2021 年，中国倡导推动的《科学家生物安全行为准则天津指南》（简称《天津指南》）受到《禁止生物武器公约》缔约国的广泛关注和支持。《天津指南》包含 10 项指导原则和行为标准，从科研责任、成果传播、科技普及、国际交流等多个环节倡议提高科研人员生物安全意识，倡导负责任的生物科研，鼓励各国政府及科研机构加强监管和自律，制定必要的自愿性行为准则，以促进生命科学造福人类，同时避

免误用滥用、防范安全风险^[94-95]。

5 总结与展望

为应对指数级增长的 DNA 合成需求，以碱基水平的精度进行大规模 DNA 制造的工业化合成平台正在将基因工程从劳动密集型工作转变为以工业、信息和技术驱动领域。同时，通过提高反应载体数量、反应位点密度、反应效率等综合措施，积极探索突破现有 DNA 合成技术瓶颈的有效途径。随着 DNA 合成技术的不断发展和迭代，DNA 合成流程中越来越多的手工操作被自动化系统取代。如何以高通量和高精度的方式，以更低的成本、在更短的时间内获得更长的高质量 DNA 产品，是当前工业化 DNA 合成领域的竞争焦点。自动化、微型化、生物酶法合成、组装与纠错等新技术和新工具的应用使 DNA 合成能力不断提升和拓展的同时，也使工业 DNA 生产系统更加绿色、低碳、环保。规范地运用 DNA 合成技术也是保障 DNA 合成技术可持续发展的重要因素，对于工业生物技术稳定、健康发展至关重要，也为保障生物技术和生物经济繁荣发挥积极的作用。

REFERENCES

- [1] Ma SY, Saaem I, Tian JD. Error correction in gene synthesis technology. *Trends Biotechnol*, 2012, 30(3): 147-154.
- [2] Consortium EBR. Engineering biology: a research roadmap for the next-generation bioeconomy[EB/OL]. [2019-06]. <https://roadmap.ebrc.org>.
- [3] Agarwal KL, Buchi H, Caruthers MH, et al. Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast. *Nature*, 1970, 227(5253): 27-34.
- [4] Notka F, Liss M, Wagner R. Industrial scale gene synthesis. *Methods Enzymol*, 2011, 498: 247-275.
- [5] Chao R, Yuan YB, Zhao HM. Recent advances in DNA assembly technologies. *FEMS Yeast Res*, 2015, 15(1): 1-9.

- [6] Potomac Institute for Policy Studies. The Future of DNA Data Storage[EB/OL]. [2018-09] https://potomacinstitute.org/images/studies/Future_of_DNA_Data_Storage.pdf.
- [7] Anastas P, Eghbali N. Green chemistry: principles and practice. *Chem Soc Rev*, 2010, 39: 301-312.
- [8] Yantsevich AV, Shchur VV, Usanov SA. Oligonucleotide preparation approach for assembly of DNA synthons. *SLAS Technol*, 2019, 24(6): 556-568.
- [9] Pon RT, Yu SY, Guo ZQ, et al. Reusable solid-phase supports for oligonucleotides and antisense therapeutics. *J Chem Soc Perkin Trans 1*, 2001(20): 2638-2643.
- [10] Hyodo M, Morimura M, Hayakawa Y. A solid support with a hydroxyallyl linker, full parts of which are potentially reusable for the synthesis of oligonucleotides. *Nucleosides Nucleotides Nucleic Acids*, 2005, 24(5/6/7): 585-588.
- [11] Kumar P, Mahajan S, Gupta KC. Universal reusable polymer support for oligonucleotide synthesis. *J Org Chem*, 2004, 69(19): 6482-6485.
- [12] Sanghvi YS, Ravikumar VT, Scozzari AN, et al. Applications of green chemistry in the manufacture of oligonucleotide drugs. *Pure Appl Chem*, 2001, 73(1): 175-180.
- [13] Trost BM. The atom economy--a search for synthetic efficiency. *Science*, 1991, 254(5037): 1471-7.
- [14] Hughes RA, Ellington AD. Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb Perspect Biol*, 2017, 9(1): a023812.
- [15] Kumar G, Poonian M. Improvements in oligodeoxyribonucleotide synthesis: methyl N,N-dialkylphosphoramidite dimer units for solid support phosphite methodology. *J Org Chem*, 1984, 49: 4905-4912.
- [16] Kayushin AL, Korosteleva MD, Miroshnikov AI, et al. A convenient approach to the synthesis of trinucleotide phosphoramidites--synthons for the generation of oligonucleotide/peptide libraries. *Nucleic Acids Res*, 1996, 24(19): 3748-3755.
- [17] Suchsland R, Appel B, Janczyk M, et al. Solid phase assembly of fully protected trinucleotide building blocks for codon-based gene synthesis. *Applied Sciences-Basel*, 2019, 9(11), 2199-2209.
- [18] Sierzchala AB, Dellinger DJ, Betley JR, et al. Solid-phase oligodeoxynucleotide synthesis: a two-step cycle using peroxy anion deprotection. *J Am Chem Soc*, 2003, 125(44): 13427-13441.
- [19] Jensen M, Roberts L, Johnson A, et al. Next generation 1 536-well oligonucleotide synthesizer with on-the-fly dispense. *J Biotechnol*, 2014, 171: 76-81.
- [20] Cheng JY, Chen HH, Kao YS, et al. High throughput parallel synthesis of oligonucleotides with 1 536 channel synthesizer. *Nucleic Acids Res*, 2002, 30(18): e93.
- [21] Lietard J, Schaudy E, Holz K, et al. High-density DNA and RNA microarrays-photolithographic synthesis, hybridization and preparation of large nucleic acid libraries. *J Vis Exp*, 2019(150): e59936.
- [22] Lietard J, Ameer D, Damha MJ, et al. High-density RNA microarrays synthesized *in situ* by photolithography. *Angewandte Chemie Int Ed*, 2018, 57(46): 15257-15261.
- [23] Hoelz K, Hoi JK, Schaudy E, et al. High-efficiency reverse (5'→3') synthesis of complex DNA microarrays. *Sci Reports*, 2018, 8: 15099-15110.
- [24] Sinyakov AN, Nikolaenkova EB, Os'kina IA, et al. Photogenerator of trichloroacetic acid as a promising detritylation agent for oligonucleotide microarray synthesis. *Russ J Bioorg Chem*, 2014, 40(5): 586-588.
- [25] Agbavwe C, Kim C, Hong D, et al. Efficiency, error and yield in light-directed maskless synthesis of DNA microarrays. *J Nanobiotechnol*, 2011, 9: 57-73.
- [26] Chow BY, Emig CJ, Jacobson JM. Photoelectrochemical synthesis of DNA microarrays. *PNAS*, 2009, 106(36): 15219-15224.
- [27] Saaem I, Ma KS, Marchi AN, et al. *In situ* synthesis of DNA microarray on functionalized cyclic olefin copolymer substrate. *ACS Appl Mater Interfaces*, 2010, 2(2): 491-497.
- [28] Lausted C, Dahl T, Warren C, et al. POSaM: a fast, flexible, open-source, inkjet oligonucleotide synthesizer and microarrayer. *Genome Biol*, 2004, 5(8): R58.
- [29] Egeland RD, Southern EM. Electrochemically directed synthesis of oligonucleotides for DNA microarray fabrication. *Nucleic Acids Res*, 2005, 33(14): e125.
- [30] Srivannavit O, Gulari M, Hua ZS, et al. Microfluidic reactor array device for massively parallel *in situ* synthesis of oligonucleotides. *Sens Actuators B Chem*, 2009, 140(2): 473-481.
- [31] Hua ZS, Xia YM, Srivannavit O, et al. A versatile microreactor platform featuring a chemical-resistant microvalve array for addressable multiplex syntheses and assays. *J Micromech Microeng*, 2006, 16(8):

- 1433-1443.
- [32] Crosby SR, Jennison M, Brennan J. Thermally-cleavable protecting and linker groups: US, 16604329. 2018-04-12.
- [33] 江湘儿, 王勇, 沈玥. DNA 合成技术与仪器研发进展概述. 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 Integr Technol, 2021, 10(5): 80-95 (in Chinese).
- [34] Tang N, Ma SY, Tian JD. New tools for cost-effective DNA synthesis. Synthetic Biology. Amsterdam: Elsevier, 2013: 3-21.
- [35] 闫汉, 肖鹏峰, 刘全俊, 陆祖宏. DNA 微阵列原位化学合成. 合成生物学, 2021, 2(03): 354-370.
Yan H, Xiao PF, Liu QJ, et al. *In situ* chemical synthesis of DNA microarrays. Synthetic Biology Journal, 2021, 2(03): 354-370 (in Chinese).
- [36] Tian JD, Ma KS, Saaem I. Advancing high-throughput gene synthesis technology. Mol BioSyst, 2009, 5(7): 714-722.
- [37] Kosuri S, Church GM. Large-scale *de novo* DNA synthesis: technologies and applications. Nat Methods, 2014, 11(5): 499-507.
- [38] Kim H, Jeong J, Bang D. Hierarchical gene synthesis using DNA microchip oligonucleotides. J Biotechnol, 2011, 151(4): 319-324.
- [39] Kim H, Han H, Ahn J, et al. 'Shotgun DNA synthesis' for the high-throughput construction of large DNA molecules. Nucleic Acids Res, 2012, 40(18): e140.
- [40] Wan W, Lu M, Wang D, et al. High-fidelity *de novo* synthesis of pathways using microchip-synthesized oligonucleotides and general molecular biology equipment. Sci Reports, 2017, 7: 6119-6130.
- [41] Quan J, Saaem I, Tang N, et al. Parallel on-chip gene synthesis and application to optimization of protein expression. Nat Biotechnol, 2011, 29(5): 449-452.
- [42] Banyai W, Peck BJ, Fernandez A, et al. *De novo* synthesized gene libraries: US, 15233835. 2016-08-10.
- [43] Plesa C, Sidore AM, Lubock NB, et al. Multiplexed gene synthesis in emulsions for exploring protein functional landscapes. Science, 2018, 359(6373): 343-347.
- [44] Sidore AM, Plesa C, Samson JA, et al. DropSynth 2.0: high-fidelity multiplexed gene synthesis in emulsions. Nucleic Acids Res, 2020, 48(16): e95.
- [45] Kuiper BP, Prins RC, Billerbeck S. Oligo pools as an affordable source of synthetic DNA for cost-effective library construction in protein- and metabolic pathway engineering. ChemBioChem, 2022, 23(7): e202100507.
- [46] Sinyakov AN, Ryabinin VA, Kostina EV. Application of array-based oligonucleotides for synthesis of genetic designs. Mol Biol (Mosk), 2021, 55(4): 562-577.
- [47] Li AT, Sun ZT, Reetz MT. Solid-phase gene synthesis for mutant library construction: the future of directed evolution? ChemBioChem, 2018, 19(19): 2023-2032.
- [48] Li AT, Acevedo-Rocha CG, Sun ZT, et al. Beating bias in the directed evolution of proteins: combining high-fidelity on-chip solid-phase gene synthesis with efficient gene assembly for combinatorial library construction. ChemBioChem, 2018, 19(3): 221-228.
- [49] Tian JD, Gong H, Sheng NJ, et al. Accurate multiplex gene synthesis from programmable DNA microchips. Nature, 2004, 432(7020): 1050-1054.
- [50] Kim H, Han H, Shin D, et al. A fluorescence selection method for accurate large-gene synthesis. ChemBioChem, 2010, 11(17): 2448-2452.
- [51] Choi H, Choi Y, Choi J, et al. Purification of multiplex oligonucleotide libraries by synthesis and selection. Nat Biotechnol, 2022, 40(1): 47-53.
- [52] Matzas M, Stähler PF, Kefer N, et al. High-fidelity gene synthesis by retrieval of sequence-verified DNA identified using high-throughput pyrosequencing. Nat Biotechnol, 2010, 28(12): 1291-1294.
- [53] Zhang J, Wang YF, Chai BH, et al. Efficient and low-cost error removal in DNA synthesis by a high-durability MutS. ACS Synth Biol, 2020, 9(4): 940-952.
- [54] Wan W, Li LL, Xu QQ, et al. Error removal in microchip-synthesized DNA using immobilized MutS. Nucleic Acids Res, 2014, 42(12): e102.
- [55] Wan W, Wang DM, Gao XL, et al. Immobilized MutS-mediated error removal of microchip-synthesized DNA. Methods Mol Biol, 2017, 1472: 217-235.
- [56] Sequeira AF, Guerreiro CIPD, Vincentelli R, et al. T7 endonuclease I mediates error correction in artificial gene synthesis. Mol Biotechnol, 2016, 58(8/9): 573-584.
- [57] Lubock NB, Zhang D, Sidore AM, et al. A systematic comparison of error correction enzymes by next-generation sequencing. Nucleic Acids Res, 2017, 45(15): 9206-9217.
- [58] Wu Y, Li BZ, Zhao M, et al. Bug mapping and fitness testing of chemically synthesized chromosome X. Science, 2017, 355(6329): eaaf4706.
- [59] Zan XZ, Yao XY, Xu P, et al. A hierarchical error

- correction strategy for text DNA storage. *Interdiscip Sci*, 2022, 14(1): 141-150.
- [60] Antkowiak PL, Lietard J, Darestani MZ, et al. Low cost DNA data storage using photolithographic synthesis and advanced information reconstruction and error correction. *Nat Commun*, 2020, 11(1): 5345.
- [61] Lee H, Wiegand DJ, Griswold K, et al. Photon-directed multiplexed enzymatic DNA synthesis for molecular digital data storage. *Nat Commun*, 2020, 11(1): 5246.
- [62] Jensen MA, Davis RW. Template-independent enzymatic oligonucleotide synthesis (TiEOS): its history, prospects, and challenges. *Biochemistry*, 2018, 57(12): 1821-1832.
- [63] Mackey JK, Gilham PT. New approach to the synthesis of polyribonucleotides of defined sequence. *Nature*, 1971, 233(5321): 551-553.
- [64] Gillam S, Waterman K, Smith M. Enzymatic synthesis of oligonucleotides of defined sequence. Addition of short blocks of nucleotide residues to oligonucleotide primers. *Nucleic Acids Res*, 1975, 2(5): 613-624.
- [65] Kaufmann G, Fridkin M, Zutra A, et al. Monofunctional substrates of polynucleotide phosphorylase. The monoaddition of 2' (3')-O-isovaleryl-nucleoside diphosphate to an initiator oligonucleotide. *Eur J Biochem*, 1971, 24(1): 4-11.
- [66] Gillam S, Waterman K, Doel M, et al. Enzymatic synthesis of deoxyribo-oligonucleotides of defined sequence. Deoxyribo-oligonucleotide synthesis. *Nucleic Acids Res*, 1974, 1(12): 1649-1664.
- [67] Gillam S, Smith M. Enzymatic synthesis of deoxyribo-oligonucleotides of defined sequence. Properties of the enzyme. *Nucleic Acids Res*, 1974, 1(12): 1631-1647.
- [68] Gilham S, Smith M. Enzymatic synthesis of deoxyribo-oligonucleotides of defined sequence. *Nat New Biol*, 1972, 238(86): 233-234.
- [69] Gillam S, Rottman F, Jahnke P, et al. Enzymatic synthesis of oligonucleotides of defined sequence: synthesis of a segment of yeast iso-1-cytochrome c gene. *PNAS*, 1977, 74(1): 96-100.
- [70] Cardenas PP, Carzaniga T, Zangrossi S, et al. Polynucleotide phosphorylase exonuclease and polymerase activities on single-stranded DNA ends are modulated by RecN, SsbA and RecA proteins. *Nucleic Acids Res*, 2011, 39(21): 9250-9261.
- [71] Hinton DM, Baez JA, Gumpport RI. T4 RNA ligase joins 2'-deoxyribonucleoside 3',5'-bisphosphates to oligodeoxyribonucleotides. *Biochemistry*, 1978, 17(24): 5091-5097.
- [72] Hinton DM, Gumpport RI. The synthesis of oligodeoxyribonucleotides using RNA ligase. *Nucleic Acids Res*, 1979, 7(2): 453-464.
- [73] Hinton DM, Brennan CA, Gumpport RI. The preparative synthesis of oligodeoxyribonucleotides using RNA ligase. *Nucleic Acids Res*, 1982, 10(6): 1877-1894.
- [74] Schmitz C, Reetz MT. Solid-phase enzymatic synthesis of oligonucleotides. *Org Lett*, 1999, 1(11): 1729-1731.
- [75] Nick McElhinny SA, Havener JM, Garcia-Diaz M, et al. A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining. *Mol Cell*, 2005, 19(3): 357-366.
- [76] Chen YN, Podlevsky JD, Logeswaran D, et al. A single nucleotide incorporation step limits human telomerase repeat addition activity. *EMBO J*, 2018, 37(6): e97953.
- [77] Chen W, Guidotti G. Soluble apyrases release ADP during ATP hydrolysis. *Biochem Biophys Res Commun*, 2001, 282(1): 90-95.
- [78] J.W. Efcavitch JES. Modified template-independent enzymes for polydeoxynucleotide synthesis: US, 11390858B2. 2020-06-13.
- [79] Mathews AS, Yang HK, Montemagno C. Photo-cleavable nucleotides for primer free enzyme mediated DNA synthesis. *Org Biomol Chem*, 2016, 14(35): 8278-8288.
- [80] Palluk S, Arlow DH, De Rond T, et al. *De novo* DNA synthesis using polymerase-nucleotide conjugates. *Nat Biotechnol*, 2018, 36(7): 645-650.
- [81] Gardner AF, Wang JC, Wu WD, et al. Rapid incorporation kinetics and improved fidelity of a novel class of 3'-OH unblocked reversible terminators. *Nucleic Acids Res*, 2012, 40(15): 7404-7415.
- [82] Efcavitch JW SJ. Modified template-independent enzymes for polydeoxynucleotide synthesis: US, 14918212. 2015-10-20.
- [83] Thomas Y MD. Variants of a DNA polymerase of the polx family: FR, 16055475. 2016-06-14.
- [84] Efcavitch JW SS. Methods and apparatus for synthesizing nucleic acids: US, 14056687. 2013-10-17.
- [85] Chen MC, Lazar RA, Huang JH, et al. A process for the preparation of nucleic acid by means of 3'-O-azidomethyl nucleotide triphosphate: US, 15555232. 2016-03-03.
- [86] Lu XY, Li JL, Li CY, et al. Enzymatic DNA synthesis by engineering terminal deoxynucleotidyl transferase. *ACS Catalysis*, 2022, 12(5): 2988-2997.

- [87] Cello J, Paul AV, Wimmer E. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science*, 2002, 297(5583): 1016-1018.
- [88] Tumpey TM, Basler CF, Aguilar PV, et al. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science*, 2005, 310(5745): 77-80.
- [89] Screening framework guidance for providers of synthetic double-stranded DNA. *Biotechnol Law Report*, 2011, 30(2): 243-257.
- [90] Harmonized screening protocol v2.0 gene sequence & customer screening to promote biosecurity[EB/OL]. [2017-11-19]. <https://genesynthesisconsortium.org/wp-content/uploads/IGSCHARmonizedProtocol11-21-17.pdf>.
- [91] Biosecurity innovation and risk reduction: a global framework for accessible, safe and secure DNA synthesis[EB/OL]. [2020-01-08]. <https://www.weforum.org/reports/biosecurity-innovation-and-risk-reduction-a-global-framework-for-accessible-safe-and-secure-dna-synthesis-582d582cd4/>.
- [92] 何蕊, 曹芹, 陈洁君, 等. 涉及基因操作的前沿生物技术风险及其法律法规应对. *生物安全学报*, 2021, 30(1): 3-10.
- He R, Cao Q, Chen JJ, et al. Risks of gene manipulation in frontier biotechnology and response of laws and regulations. *J Biosaf*, 2021, 30(1): 3-10 (in Chinese).
- [93] 中华人民共和国生物安全法[EB/OL]. [2020-10-17]. <http://www.npc.gov.cn/npc/c30834/202010/bb3bee5122854893a69acf4005a66059.shtml>.
- [94] 《天津指南》成为全球生物安全高级别原则[EB/OL]. [2022-09-20]. http://digitalpaper.stdaily.com/http_www.kjrb.com/kjrb/html/2022-09/20/content_541833.htm?div=-1.
- [95] 科学家生物安全行为准则天津指南 [EB/OL]. [2021-09-03]. <https://www.cas.cn/zcjd/202111/P020211118397643265029.pdf>.

(本文责编 陈宏宇)