

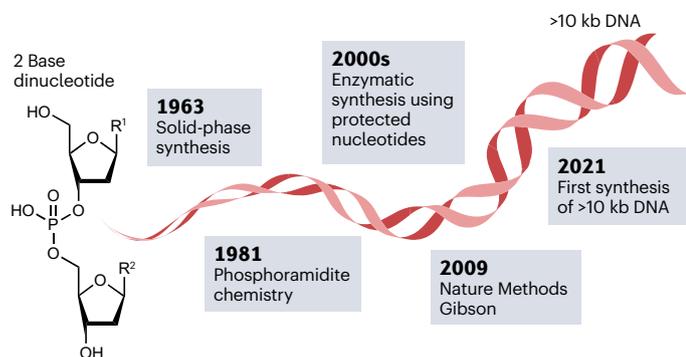
DNA synthesis technologies to close the gene writing gap

Alex Hoose¹, Richard Vellacott², Marko Storch^{3,4}, Paul S. Freemont^{3,4} & Maxim G. Ryadnov¹✉

Abstract

Synthetic DNA is of increasing demand across many sectors of research and commercial activities. Engineering biology, therapy, data storage and nanotechnology are set for rapid developments if DNA can be provided at scale and low cost. Stimulated by successes in next generation sequencing and gene editing technologies, DNA synthesis is already a burgeoning industry. However, the synthesis of >200 bp sequences remains unaffordable. To overcome these limitations and start writing DNA as effectively as it is read, alternative technologies have been developed including molecular assembly and cloning methods, template-independent enzymatic synthesis, microarray and rolling circle amplification techniques. Here, we review the progress in developing and commercializing these technologies, which are exemplified by innovations from leading companies. We discuss pros and cons of each technology, the need for oversight and regulatory policies for DNA synthesis as a whole and give an overview of DNA synthesis business models.

Sections

[Introduction](#)[Why DNA — economy drive](#)[Making DNA: underpinning technologies](#)[Technologies for DNA of unlimited length](#)[Emerging commercialized technologies](#)[Business models and deskilling](#)[Oversight and standardization](#)[Conclusions](#)

¹National Physical Laboratory, Teddington, Middlesex, UK. ²BiologIC Technologies, Cambridge, UK. ³London Biofoundry, Translation and Innovation Hub, Imperial College White City Campus, London, UK. ⁴Section of Structural and Synthetic Biology, Faculty of Medicine, Imperial College London, London, UK. ✉e-mail: max.ryadnov@npl.co.uk

Introduction

DNA is the information repository of life. Since its discovery, it has become an essential research tool for chemistry, biology and materials science. The past two decades have witnessed a remarkable progress in generating biological systems including viable microorganisms from synthetic genomes^{1,2}. As a consequence of this success, the demand for DNA is increasing, driving the development of new technologies to provide DNA in greater purity, quantity and at a reduced cost². These requirements have steered commercial priorities towards supplying synthetic DNA, as opposed to isolation of DNA derived from natural sources.

The ability to sequentially synthesize polynucleotides, nucleotide by nucleotide, allows for control over the composition and size of DNA. Synthetic DNA sequences provide researchers with a versatile tool to probe living systems, rather than relying on natural sequences isolated from organisms. Additionally, for some applications such as the amplification of inaccessible sequences, synthetic DNA is the only practical option. The development of DNA synthesis technologies may also be relevant in materials science and nanotechnology, for example, in DNA origami, to create new types of DNA architectures and functionalities using non-natural nucleotides or non-natural backbones, such as xeno nucleic acids (XNAs)^{3–5}. Similarly, the synthesis of homo-polynucleotides, co-block and arbitrary polynucleotides has gathered momentum in applications in which single-stranded DNA acts as a scaffold or donor material for nanoscale devices or genome engineering^{6,7}.

Innovations in next generation sequencing (NGS) have improved reading and editing DNA⁸ and revolutionized cellular and populational genomic analysis, which are now applied in ‘mega-genomic’ initiatives⁹. DNA can be analysed at scale and low cost. However, the lack of large-scale DNA synthesis remains a barrier to technological advances and the large-scale analysis of genome structure and cellular function. This barrier highlights an existing gap between the well-developed ability to read DNA, identify and sequence genomes, with the less-developed ability to write DNA, and synthesize and produce DNA sequences of unlimited lengths and complexity.

In the current climate of DNA synthesis commercialization, businesses either offer DNA they synthesize themselves or ready-to-use automated synthesizers for researchers to make DNA in their own laboratories. Both routes make DNA synthesis accessible to those end-users who lack expert synthesis skills and as such ‘deskill’ DNA synthesis. However, widespread access to synthetic DNA through deskilling may lead to the misuse of synthetic DNA, which introduces the need for regulation to mitigate potential hazards resulting from the misuse¹⁰.

Here, we review both existing and emerging DNA synthesis technologies, with an emphasis on methodologies developed in industry as a means to accelerate the supply of long synthetic DNA. We also discuss challenges and opportunities that DNA synthesis brings for commercialization.

Why DNA — economy drive

Engineering biology holds promise for providing solutions to the global challenges of resource sustainability. Advances in engineering biology are already addressing industry needs by building effective partnership networks and investing into automation^{11,12}. Indeed, 60% of all manufacturing inputs into the global economy could be produced biologically, whereas 30% of research and development spent is in biology-related industries¹³. DNA synthesis is indispensable in this regard as it provides the essence of engineering biology — DNA molecules of desired composition, complexity and length.

Four waves of DNA technologies

Following the sequencing of the human genome in the early 2000s¹⁴ (the first wave of DNA technologies), the ability to ‘read DNA’ has advanced at a pace that has outstripped even Moore’s law, which predicts that the number of transistors doubles every 2 years¹⁵ (Fig. 1a). As this area matured, a second wave was driven by novel technologies such as de novo DNA synthesis and CRISPR gene editing, which has given an ability to ‘edit and write’ DNA^{15,16}. This has enabled researchers to begin to ‘apply DNA’ by exploiting the abilities to read, edit and write DNA for products such as vaccines¹⁷, data storage⁹ and drug delivery devices⁶ or genome engineering to generate organisms with useful properties, such as heat-resistant plants¹⁸. The improved ability to apply DNA brings the need for synthesis on scale, to provide for industry challenges ranging from health security to environmental sustainability^{17–19}. For instance, the COVID-19 pandemic has shown how quickly the rise in demand for vaccination can overwhelm existing production abilities. As RNA and DNA vaccines continue to get approved for major diseases, including COVID-19 (ref. 17), the demand for mass production of large DNA is growing. Similarly, synthetic DNA can be used in plants able to adapt to climate change, mitigating food security challenges¹⁸. Activated DNA repair increases the tolerance of plants to heat, while introducing synthetic genes make crops able to collect nutrients and water more efficiently in different conditions^{18,19}. Effective DNA synthesis is therefore vital to close the gap between the ability to read and write DNA.

Since the structure of DNA was first understood²⁰ (Box 1), substantial milestones have been achieved, paving the way to a new industry. Over four decades, short but meticulous steps were taken to establish underpinning chemistry for the stepwise synthesis of DNA, nucleotide by nucleotide (Fig. 1b). Chemical methods were developed to reliably provide short <200-nucleotide DNA chains, termed oligonucleotides. These methods were optimized for automatic synthesizers, which became indispensable tools for gene engineering and sequencing. Following this, the development of enzymatic and hybrid approaches to generate DNA that is longer and more complex than oligonucleotides has been achieved (Fig. 1b). Companies have commercialized these approaches, offering services ranging from custom synthesis to benchtop DNA printers, making DNA synthesis accessible to non-expert users. This coincided with an apparent increase in ways that DNA can be applied while exposing the gap in DNA writing capabilities (Fig. 1a). In recent years, DNAs of thousands of nucleotides in length have been produced, highlighting that this gap between the abilities to read and write DNA may close in the near future (Fig. 1b).

Industry landscape for DNA synthesis

The DNA synthesis industry is rapidly growing, with an apparent shift towards greener solutions to reduce the dependence on chemical reagents and organic solvents with potentially adverse effects on the environment to mitigate their costly disposal routes²¹. New industry partnerships have been formed to introduce innovative technologies in the enzymatic DNA synthesis space. This can be exemplified by joint ventures between Codexis and Molecular Assemblies and between Integrated DNA Technologies and Danaher, which aim to advance enzymatic DNA synthesis abilities^{21–23}. Promising technologies include plasmid template approaches, such as rolling circular amplification²⁴, gene assembly approaches, such as Gibson assembly or polymerase cycling assembly (PCA)^{25,26}, and template-independent enzymatic oligonucleotide synthesis (TIEOS), which exploits terminal deoxynucleotidyl transferase (TdT) as a DNA synthesis tool^{27,28}.

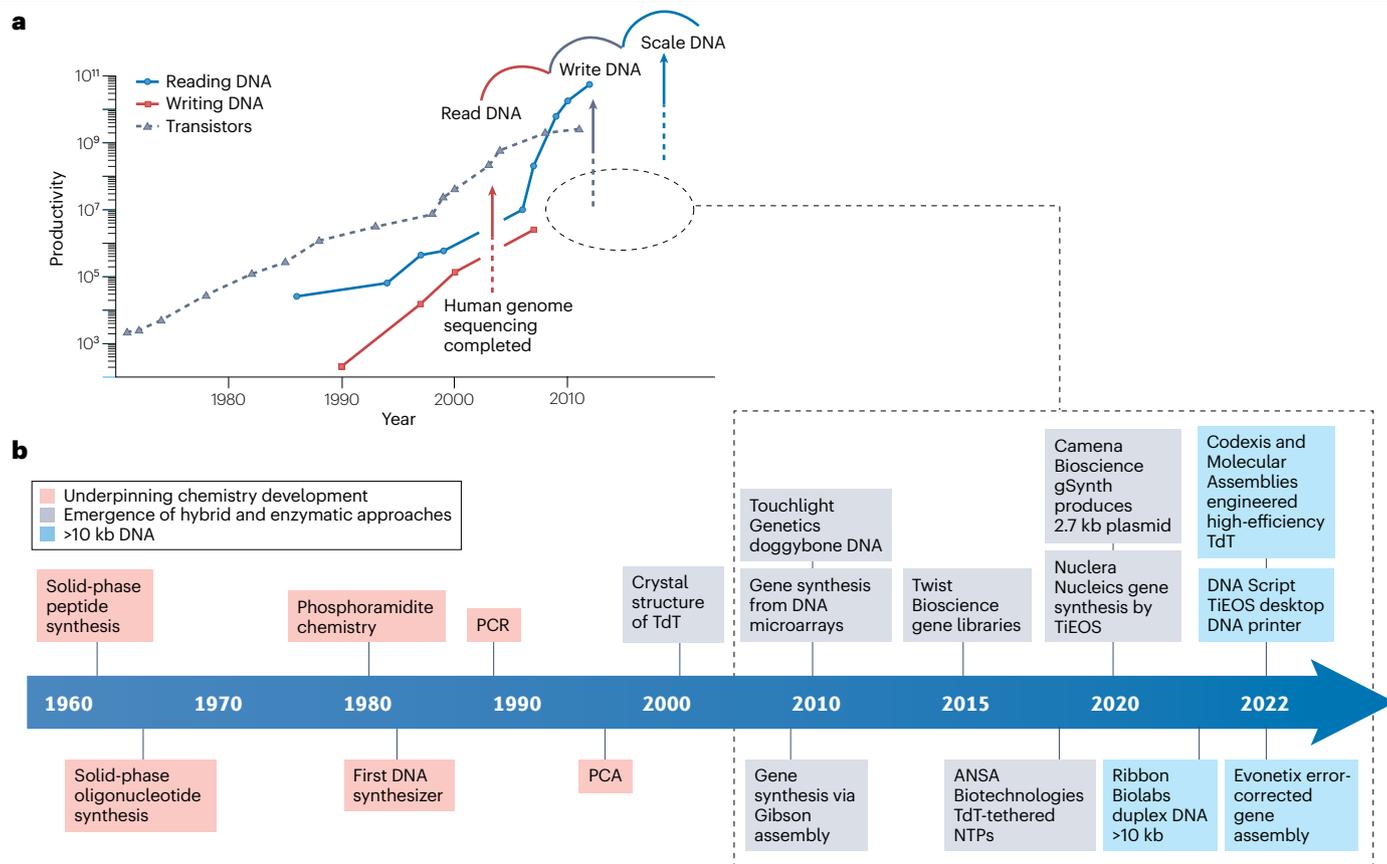


Fig. 1 | The state of the art in DNA synthesis. a, Productivity of DNA reading and DNA writing (synthesis) estimated in the number of nucleotides per person per day¹⁵. The grey arrow denotes the current gap in productivity between reading DNA and writing DNA. The dashed oval outlines the time frame within which the DNA synthesis industry achieved the majority of important milestones to close the gap. DNA synthesis data (red line) are available only for column-based synthesis instruments. The number of transistors per chip

(Moore's law) is shown for comparison. The graph uses the data available in the literature. **b**, Timeline of milestones in DNA synthesis technologies discussed in the report^{20,25,26,34,35,37,42,71,146}. For simplicity not all milestones are shown. NTP, nucleoside 5'-triphosphate; PCA, polymerase cycling assembly; TdT, terminal deoxynucleotidyl transferase; TiEOS, template-independent enzymatic oligonucleotide synthesis. Copyright Wiley-VCH GmbH. Reproduced with permission from ref.¹⁵.

A key challenge in DNA synthesis is the generation of >300-nucleotide DNA, which is limited by the elongation cycle efficiency, that is, the efficiency with which each nucleotide is incorporated in the sequence. For example, with the elongation cycle efficiency of 99%, the theoretical yield for an oligonucleotide comprising 120 nucleotides is ~30% ($0.99^{120} \times 100\%$). However, for a 200 bp polymer/oligonucleotide, this is reduced to just 13%. Attempts to overcome this issue have focused on improving the accuracy and speed of DNA assembly processes. For commercial technologies, when >1 kb sequences are required, chemically produced >300 bp polynucleotides are used as building blocks for larger chains^{29–31}. DNA printers, developed by some vendors, such as DNA Script, have enabled the parallel synthesis of multiple sequences, which can be linked together to produce longer chains. Other companies such as Molecular Assemblies focus on improving synthesis methodologies that might be implemented by developers of DNA printers or providers of synthetic genes.

Other vendors, such as ANSA Biotechnologies and Camena Bioscience, analyse the quality of the DNA they produce, to eliminate the need for the user to perform further sequencing or cloning. This has

also allowed for more oversight to counter potential biosecurity risks. As synthetic DNA is involved in genetic engineering, there is a risk of its use in the production of pathogens and hence it is subject to an oversight or regulatory system. Similarly, companies that manufacture DNA printers use a cloud-based software enabling a degree of oversight for a desktop production mode. Table 1 provides examples of DNA synthesis companies highlighting pros and cons of their core technologies.

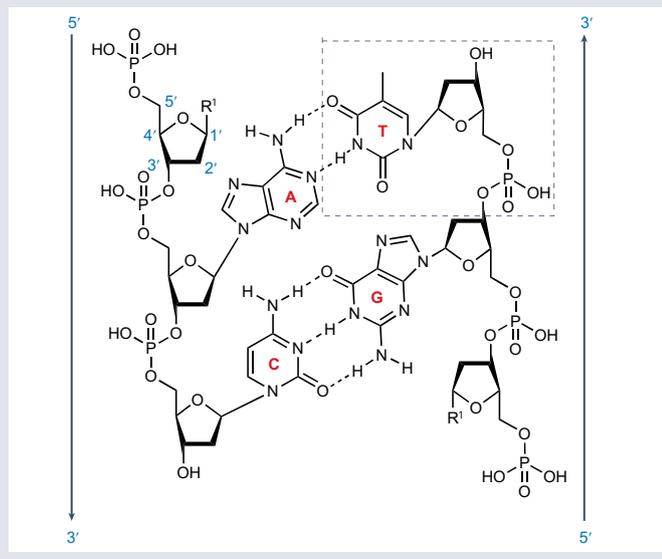
Making DNA: underpinning technologies

Over the past few decades, there has been significant interest in the development of DNA synthesis techniques. Starting with chemically synthesized dinucleotides³², de novo DNA synthesis was made possible and exploited in the process of deciphering the genetic code³³. Advances in solid-phase synthesis inspired further synthetic improvements^{34,35}, which led to the ground-breaking development of phosphoramidite chemistry for DNA synthesis in the 1980s resulting in the introduction of phosphoramidite oligonucleotide synthesis (POS)^{36,37}.

Box 1

Chemical structure of DNA

DNA is a monodisperse polymer of four building blocks, each comprising a base and a sugar residue, deoxyribose. Each block (adenosine (A), thymidine (T), guanosine (G), cytidine (C)) is coupled to a phosphate, which constitutes a nucleotide (thymidine highlighted by a dashed box). Nucleotides joined by phosphodiester bonds form a polynucleotide chain. The bases in this polymer pair up via three hydrogen bonds between G and C and two hydrogen bonds between A and T²⁰. This selectivity, termed Watson–Crick base pairing, enables the formation of a double helix in which two complementary strands run anti-parallel to one another and intertwine with base pairs (A–T, C–G) stacked along the helical axis²⁰. Owing to its self-complementarity, the double helix can be self-replicated from each of the two strands. Arrows indicate the 5'-to-3' direction of the asymmetric ends.



Phosphoramidite synthesis

A typical solid-phase synthesis of oligonucleotides using phosphoramidite chemistry to build up a sequence, nucleotide by nucleotide, is given in Fig. 2A. This approach involves the stepwise addition of building blocks derived from 5'-protected dimethoxytrityl (DMT) nucleotide phosphoramidites **4** (refs. ^{38–40}). This method was used by Applied Biosystems to develop the first automated DNA synthesizer in the 1980s, improving the accessibility of synthetic oligonucleotides^{41–45}. Initial solid-phase methodologies used plastic or glass solid supports, onto which individual oligonucleotide sequences were chemically assembled⁴⁶ (steps c–i). Since then, parallel in situ synthesis of oligonucleotides has been achieved using different microarray formats comprising multiple reaction sites, where one sequence is assembled onto one site, which can be controlled independently of other sites thus providing the synthesis of multiple sequences in a site-specific manner^{47–50}.

Current technologies use silicon as a solid support onto which a million unique oligonucleotides can be written simultaneously^{29,48}.

Microscopic reaction clusters manufactured on a silicon chip decrease the reaction volume and significantly increase the output of DNA compared with single sequence synthesis methods⁵⁰, whereas thermal control provides a means to monitor the incorporation of each nucleotide to enable site-specific DNA synthesis^{51,52}.

However, the phosphoramidite method of DNA synthesis has drawbacks, including poor phosphoramidite bench stability, the need to use large quantities of organic solvents and the inability to synthesize poly-repeat sequences^{53–61}. In addition, the acid required to remove the 5'-DMT protecting group (PG) can catalyse depurination (steps j–m in Fig. 2B), a deleterious side reaction leading to the loss of purine bases (A, G) **13** from the synthesized DNA strand **12**, making this DNA strand susceptible to hydrolysis **14** and **15** and premature release **16** and **18** (refs. ^{62–65}). As a result, depurination reduces the yield and purity of the desired oligonucleotide.

The workflow of oligonucleotide manufacturing, processing and purification is labour-intensive and remains largely the domain of service providers. Therefore, synthesis capabilities have become centralized within specialist reagent manufacturers. Leading vendors such as Agilent Technologies, GenScript, Integrated DNA Technologies, ThermoFisher, TriLink, Dharmacon, Twist Bioscience and others produce custom DNA (and RNA) on demand in a range of formats. For those users who wish to decrease the lead time for such services, there is a range of instruments, for example, Cytiva's ÅKTA oligonucleotide synthesizers, which can be purchased and operated on a daily basis.

Traditionally, molecular biology relied on short DNA sequences, such as primers for PCR or probes for molecular detection, amplification and modification applications. More recently, asymmetric PCR methods⁶⁶ have enabled advances in the amplification of individual DNA strands of thousands of nucleotides in length^{67,68}. Now researchers seek longer sequences of varied composition including entire genomes with a single-base accuracy, which must be assembled from scratch^{31,69}. Such long sequences are incompatible with the phosphoramidite method whose efficacy in the synthesis of pure DNA reduces beyond approximately 200-bp oligonucleotide sequences.

To synthesize long DNA, the elongation cycle efficiency must be increased to improve yields, and the incomplete removal of PGs and side reactions such as depurination must be minimized or avoided^{62,70}. Longer sequences must be assembled from smaller strands in error-correcting stages using an alternative methodology. Enzymatic approaches are most attractive in this regard and are also scalable, stereospecific and environmentally friendly²¹. Enzymes can mediate mismatch recognition enabling the selective annealing of complementary strands, reduce the number of steps in each elongation cycle by eliminating the need for coupling reagents and decrease the dependence on organic solvents. Enzymes can promote synthesis with or without DNA templates, through amplification or in the synthesis of de novo sequences.

Enzymatic oligonucleotide synthesis

Enzymatic synthesis uses the principles of solid-phase synthesis. A short strand of DNA synthesized on a solid support can be extended by DNA polymerases using nucleoside 5'-triphosphates (NTPs)²⁷. DNA polymerases use a template DNA strand that provides base pairing, thereby selecting the incoming nucleotide. This means that although polymerases are effective in amplifying existing DNA templates, they are unable to generate de novo DNA sequences. Therefore, an alternative enzyme is required to efficiently elongate polynucleotide chains in the absence of a template strand^{71–73}. Such a polymerase

has been identified as TdT and is integrated into commercial TiEOS methods^{27,28,71} (Fig. 3A).

Template-independent enzymatic oligonucleotide synthesis. TdT elongates oligonucleotides in the 5'-to-3' direction in a promiscuous manner, accepting any of the four canonical nucleotides, resulting in the concomitant formation of different sequences⁷³. An effective solution is to control the incorporation of nucleotides via a 'reversible termination' mechanism (Fig. 3A). This mechanism uses NTPs modified with a synthesis-interrupting 'terminator' or PG at the 3' position, which ensures the addition of a single nucleotide per reaction step and is subsequently removed to incorporate the next desired nucleotide 24–26 (refs. 74–82). To this effect, TiEOS uses resin beads pre-loaded with a chemically synthesized single-stranded initiator DNA 19 (refs. 72,83), onto which TdT ligates 3'-protected NTPs 20 into a desired sequence 23. At each step of elongation cycle, steps **a** and **b** (Fig. 3A), a washing step is used to remove side-products and surplus reagents, and the deblocking of 3'-PG is performed at the end of each cycle before the next elongation cycle (steps **b** and **c** in Fig. 3A). Initiator DNA 19 incorporates a highly specific deoxyuridine cleavage site at its 3' end which is

enzymatically labile. This site is cleaved by uracil DNA glycosylase upon the completion of the synthesis to release the assembled sequence 23 from the resin^{27,74–78}.

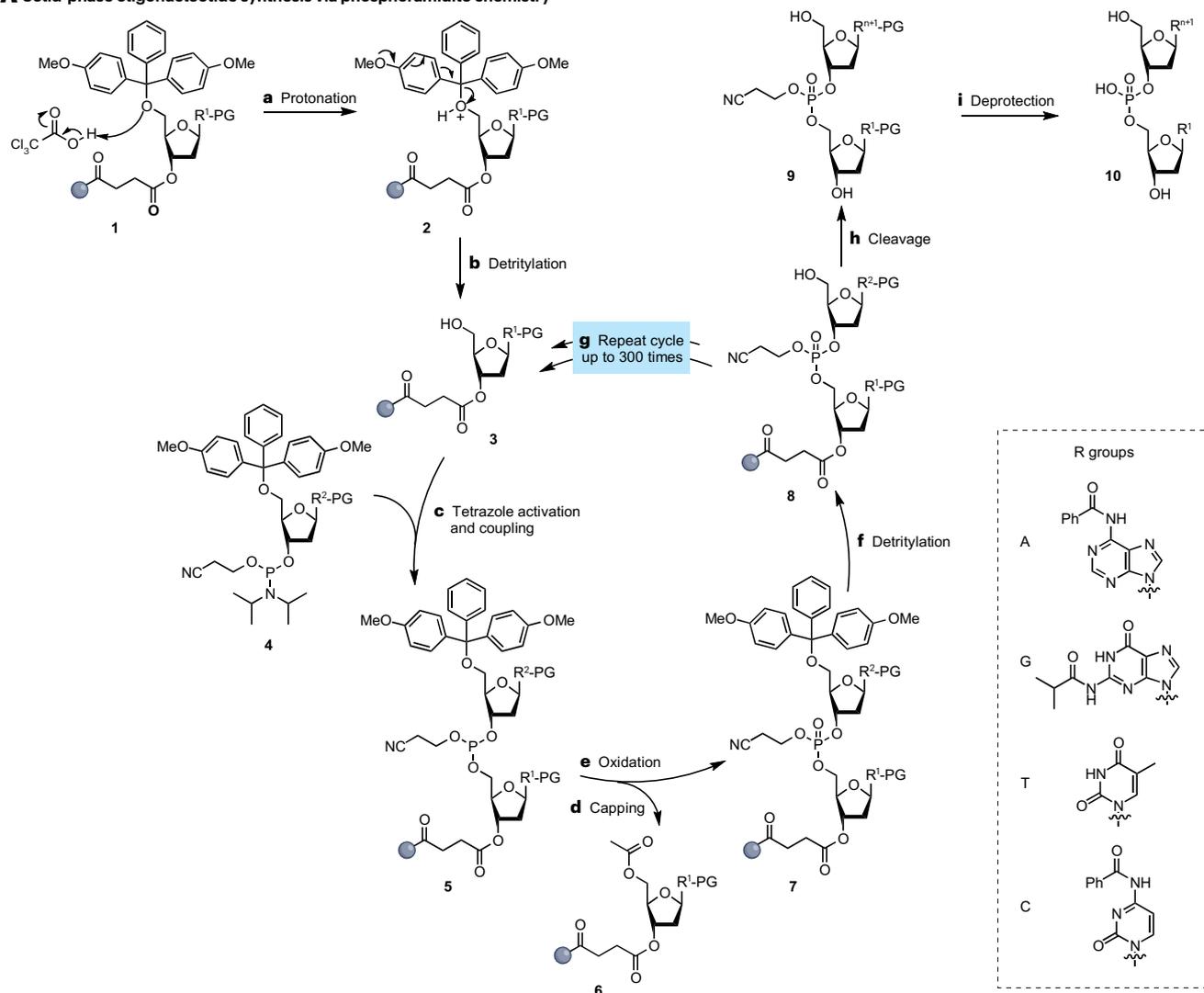
TdT methodologies as a main paradigm in DNA synthesis. As more companies test this approach, several important and unique limitations of TdT have been reported. First, the enzyme demonstrates a preference for the incorporation of some nucleotides over others⁸⁴. This bias could increase the rates of sequence-specific errors. Second, TdT works only on single-stranded DNA. This is attributed to a lariat-like loop in the enzyme, which acts as a steric shield that prevents a double-stranded DNA template accessing the active site of the enzyme^{71,80}. Consequently, the efficiency of the synthesis is reduced if the strand under construction begins to form secondary structures⁸³. Third, like all DNA polymerases, TdT-catalysed phosphoryl transfer requires divalent cations to synthesize DNA from NTPs⁸⁵. However, unlike other DNA polymerases, which typically require Mg²⁺ to catalyse the synthesis of DNA molecules, TdT can use various divalent metal cations, for example, Co²⁺, Mn²⁺, Zn²⁺ and Mg²⁺, with the NTP incorporation tailored by the cation identity. For instance, the use of Mg²⁺

Table 1 | Examples of companies applying enzymatic and hybrid approaches to DNA synthesis

Company	Business type	Synthesis platform	Length ^a (kb) ^b	Pros ^c	Cons ^c
ANSA Biotechnologies ¹³²	Service provider	TdT reversibly tethered to NTPs	0.3–3	Tethering prevents formation of homopolymeric tracts Low reliance on chemistry	Tethers can leave 'chemical scars' upon removal Incomplete tether removal affects downstream applications
Camena Bioscience ^{101,128,133,134}	Service provider	TiEOS with combinations of enzymes	0.3–3	High sequence accuracy (full-length sequences at 85% vs 23% by POS) Combination of different enzymes avoids dependence on TdT	Accuracy limited to <3kb sequences Secondary structure formation of long G-rich sequences
DNA Script ^{23,94,95}	DNA printer	TdT with reversibly terminated 3'-protected NTPs	0.01–3	Customized reagent kits, off the shelf solution for newcomers High-throughput benchtop synthesizers (96 oligos a run)	Printers print sequences of up to 80 nucleotides in length Secondary structure formation of long G-rich sequences
Evonetix ^{51,52,149,150}	DNA printer	Thermally controlled parallel synthesis	0.01–10	Site-selective, highly parallelized, high-density on-chip production Effective for the synthesis of GC-rich and repeat sequences	Insufficient thermolysis can reduce purity Challenging to differentiate product from mismatched sequences for long DNA
Molecular Assemblies ^{129–131}	Service provider	TdT with reversibly terminated 3'-protected NTPs	1–10	Steric control over elongation and random polymerization Option for protected NTPs with free 3'-OH	Subject to compatibility between re-engineered TdT and terminators Limited choice of cleavable terminators compatible with TdT
Ribbon Biolabs ¹⁵³	Service provider	Convergent Gibson assembly	1–10	High accuracy assembly under robotic control Synthesis of large DNA: first >10kb DNA made	Reliance on large libraries of highly pure oligonucleotides High impact of mutations and frameshifts on synthesis
Touchlight Genetics ^{173–176}	Service provider	Templated rolling circle amplification	1–10	Rapid manufacturing of multigram DNA quantities DNA of practically unlimited lengths	Dependence on template production 'Cleavage scars' in resulting dbDNA
Twist Bioscience ^{47–50}	Service provider	Hybrid approach of POS, PCR and Gibson assembly	0.3–5	Low error rates in high-density microarray format Efficient solution to reduce depurination in each cycle	Low quantities of produced oligonucleotides Reliance on PCR to amplify DNA for gene assembly

3'-OH, 3'-hydroxyl group of the deoxyribose ring; C, cytosine; dbDNA, doggybone DNA; G, guanosine; NTP, nucleoside 5'-triphosphate; POS, phosphoramidite oligonucleotide synthesis; TdT, terminal deoxynucleotidyl transferase; TiEOS, template-independent enzymatic oligonucleotide synthesis. ^aLengths depend on sequence composition and complexity. ^bConservative estimates of conceivable lengths by the given technology, and POS and TiEOS remain limited to <0.3 and <3kb, respectively. ^cFew exemplars given in each case for simplicity, the exemplars may be common for different vendors using similar underpinning technologies.

A Solid-phase oligonucleotide synthesis via phosphoramidite chemistry



B Depurination mechanism

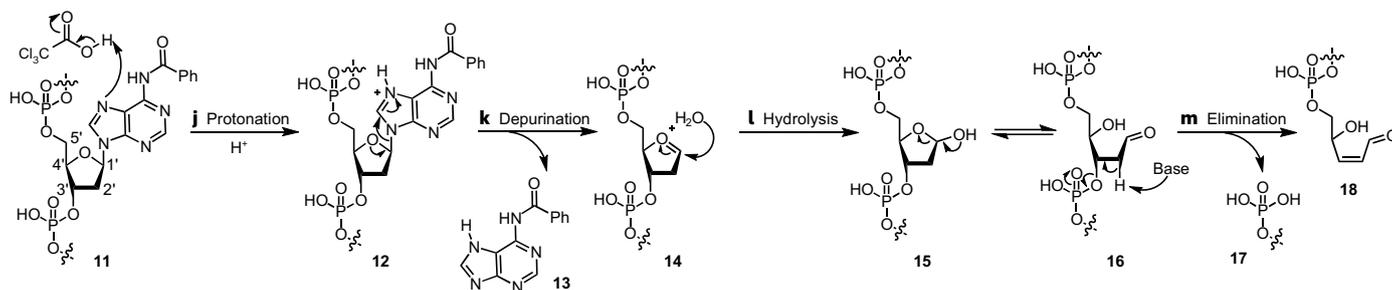


Fig. 2 | Mechanisms for solid-phase oligonucleotide synthesis and depurination. **A**, Schematic representation of phosphoramidite solid-phase oligonucleotide synthesis. The method sequentially adds 5'-dimethoxytrityl-protected nucleoside phosphoramidites **4** upon activation by tetrazole to ensure sequence-specific strand elongation^{35–39}. The steps of the process include a synthesis cycle comprising: protonation (step **a**), detritylation (step **b**), tetrazole activation and coupling (step **c**), capping of unreacted nucleotides on the resin (step **d**), oxidation (step **e**), detritylation (step **f**), which is repeated *n* times (step **g**),

and followed by cleavage from the support (step **h**) and deprotection (step **i**) to yield a desired DNA **10**. **B**, Mechanism of acid-catalysed depurination as a common side reaction in chemical synthesis^{62–64,163}, comprising protonation (step **j**), depurination (step **k**), hydrolysis (step **l**) and elimination (step **m**). Resulting apurinic site **14** owing to the loss of a purine base (for example, adenine **13**) is readily hydrolysed (steps **c** and **d**) during the basic work-up steps required for removing base protecting group (PG).

favours the incorporation of deoxyguanosine triphosphate and deoxyadenosine triphosphate, whereas Co^{2+} promotes the incorporation of deoxycytidine triphosphate and deoxythymidine triphosphate^{84–86}. Crucially, this bias extends to protected NTPs used in DNA synthesis^{27,28,80}, prompting researchers to develop methods to mitigate the bias⁸⁷. Additional features of TdT, which impact on the choice of PGs and synthesis efficiency, include the DNA phosphorylation capacity and phosphatase activity of the enzyme^{88–90}. Thus when a growing oligonucleotide chain is exposed to a mixture of NTPs, TdT would preferentially incorporate certain nucleotides resulting in the synthesis of homopolymeric chains of varying lengths.

To address these shortcomings, different approaches are being explored. By analogy to peptide synthesis⁹¹, microwave irradiation can be tailored to accelerate synthesis using DNA polymerases that work on double-stranded DNA or convert a desired double-stranded DNA into its single-stranded form, which is accessible to TdT^{92,93}.

To avoid the random incorporation of NTPs into a growing DNA chain by TdT, suitable 3'-PGs have been developed for NTPs **24–26** (Fig. 3B), which facilitate a sequential synthesis cycle comprising 3'-PG deblocking, resin washing and the coupling of NTPs (Fig. 3A). This cycle constitutes a technologically optimized TiEOS that is already adopted by several companies such as DNA Script and Nuclera Nucleics^{23,94,95}.

Important optimizations for this approach concern the design of 3'-PGs, for example, DNA Script chooses 3'-ONH₂-protected NTPs **25** (refs. ^{94,96}), whereas Nuclera Nucleics and Molecular Assemblies prefer azidomethyl terminators **24** (refs. ^{97–100}) and Camena Bioscience appears to favour 2-nitrobenzyl **26** as a 3'-PG¹⁰¹. Other PGs are attempted for the protection of 3'-OH and the bases of NTPs^{102–106}, with parallel efforts focusing on PGs for XNA synthesis^{90,107}. However, TdT must be able to accommodate the protected nucleotides in its active site, which limits the choice of PGs or requires the re-engineering of the enzyme for compatibility with 3'-PG. Indeed, 3'-PG NTPs are not natural substrates for TdT owing to the steric hindrance in the active site of the enzyme⁷¹, and their development is closely guarded by vendors^{82,101,108}. Re-engineering of TdT may provide a solution to this issue and also may aid the development of thermostable TdT^{81,95,109–114}. DNA Script, Nuclera Nucleics and Molecular Assemblies are active players in this area^{76,77,98,109,115–128}, whereas Camena Bioscience has developed a proprietary combination of high-fidelity enzymes to achieve template-free DNA synthesis^{101,128} (Table 1).

Other companies adapt an alternative approach to temporarily cap the growing oligonucleotide chain by developing 3'-OH protecting strategies. For example, Molecular Assemblies furnish incoming NTPs with blocking groups to sterically shield its 3'-OH from elongation until removal^{129–131}. In another strategy, ANSA Biotechnologies tether TdT to the base of an incoming NTP via a cleavable linker to prevent the formation of homopolymeric nucleotide tracts¹³² (Fig. 3C). The α -phosphate group of the NTP **28** reacts with the 3'-OH of the growing oligonucleotide **27**, whereas its unprotected 3'-OH remains sterically shielded by the enzyme **30**, which prevents polymerization¹²¹. Cleaving the linker releases TdT **31** and the elongated oligonucleotide **32**. By repeating the cycle, steps **a–c**, the desired sequence can be assembled and released **33**.

The yield, purity and achievable lengths of chemically synthesized oligonucleotides depend on the effective completion of each coupling cycle. Although a two-step cycle used in TiEOS is an improvement to the four steps required in POS, TiEOS is unlikely to provide the cost-effective and time-effective synthesis of full-length genes. A nearly quantitative elongation cycle efficiency of 99.9% results in a <37% yield for a 1,000-bp (or 1 kb) DNA strand. By contrast, the 99.7% efficiencies

reported by DNA Script would result in yields less than 5%²³. However, even with 99.9%, a yield of <5% for 3 kb DNA would be achievable. For example, Camena Bioscience applied their proprietary de novo synthesis and gene assembly technology – gSynth – for the construction of a 2.7 kb plasmid vector, pUC19 (refs. ^{101,128,133,134}). As the synthesis progresses, >3 kb polynucleotide chains can form stabilized secondary structures (for example, hairpins) with detrimental effects on the elongation cycle efficiency^{83,135}. Microwave treatments might mitigate this issue, but still within the 3 kb range⁹².

Despite limitations, TiEOS reduces the complexity of crude oligonucleotides by minimizing the number of possible impurities, uses 'green' reagents, and relies on fewer steps per synthesis cycle when compared with POS. These benefits of TiEOS enhance product purity and quality compared with POS, but still do not achieve quantitative elongations or resolve the detrimental impact of secondary structure formation on DNA synthesis^{83,135}. Therefore, TiEOS is viewed as a promising 'green' methodology for the synthesis of <3 kb DNA. To synthesize larger constructs (that is, gene clusters or chromosomes), TiEOS may be used to generate shorter fragments that can then undergo ligation by Gibson assembly or PCA.

Technologies for DNA of unlimited length

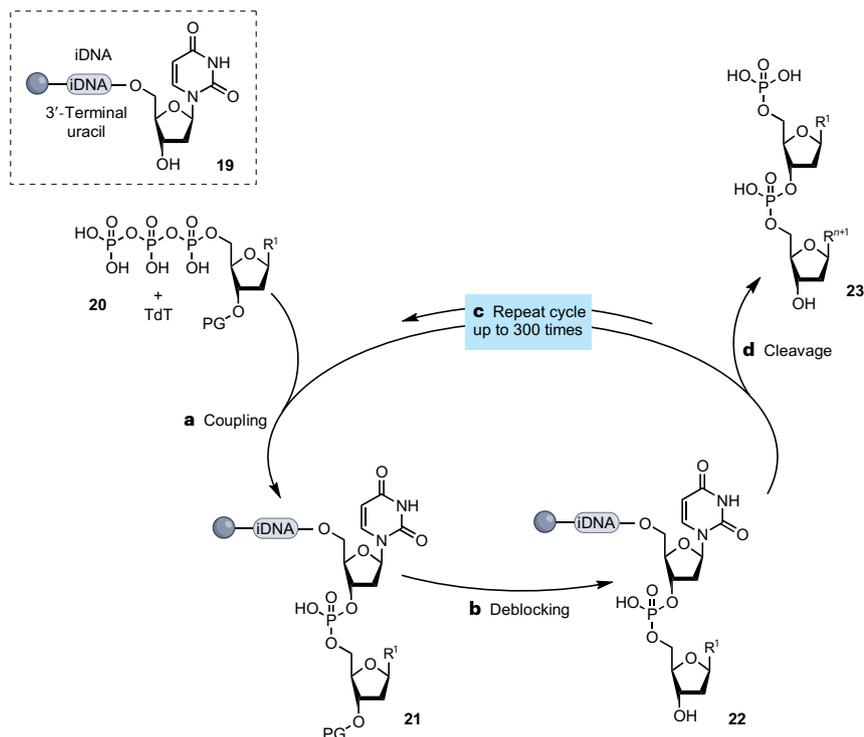
The complementary nature of DNA (Box 1) and the wealth of enzymes capable of polymerizing, cleaving, nicking, ligating and mutating DNA have resulted in the development of various assembly methods. By improving enzymes and assembly standards, the accuracy and number of DNA molecules that can be combined in a single step have improved, which has been applied in the synthesis of a minimal bacterial genome¹³⁶ and synthetic yeast chromosomes⁶⁹. With DNA assembly methods reviewed in detail elsewhere^{137,138}, here we focus on two essential methods for DNA synthesis workflows, namely, Gibson assembly^{25,139,140} and PCA^{26,141}.

Gibson assembly

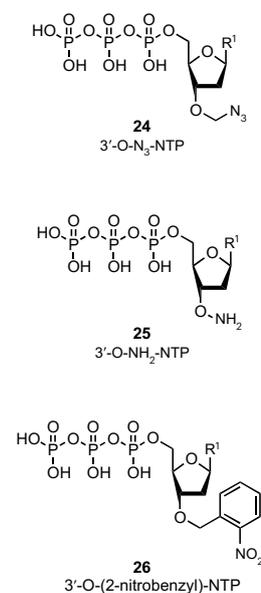
Gibson assembly is an enzymatic approach used to complement POS and TiEOS methods^{25,139}. Although this approach is inefficient for the synthesis of short strands (<100 nucleotides)¹⁴², it is used to assemble large DNA fragments^{143–145} (Fig. 4A). Gibson assembly starts with two DNA duplexes **34** and **35**, which have complementary terminal overlap regions. Each strand of these DNA duplexes is degraded by an exonuclease from the 5'-end, generating the 3'-'sticky' ends of duplexes **36** and **37**. The sticky ends of these two duplexes are then annealed in step **b** and repaired by a polymerase, which adds missing nucleotides to the two strands using base pairing interactions. A DNA ligase then stitches the nucleotides of each strand together to form the desired duplex product **38**.

Multiple rounds of Gibson assembly yield large genetic fragments for a range of applications, such as protein expression to transcriptional control. However, the process remains laborious. A non-automated gene assembly is time-consuming, which is compounded by the need for high-purity oligonucleotides in large quantities. Oligonucleotide purity is also critical to ensure correct assembly, even small percentages of deletions can create substantial frameshift mutations within the open reading frame of a desired DNA – the section of DNA that is transcribed by enzymes into RNA. Even a single deletion can shift the reading frame compromising RNA transcription, which renders the DNA unusable. Because of this, the final gene products are cloned into plasmids and transformed into bacterial strains to confirm the presence of the desired DNA sequence. The synthesis of longer genes

A TiEOS: the 3'-protected NTP approach



B 3'-Protected NTPs



C TiEOS: the tethered NTP approach

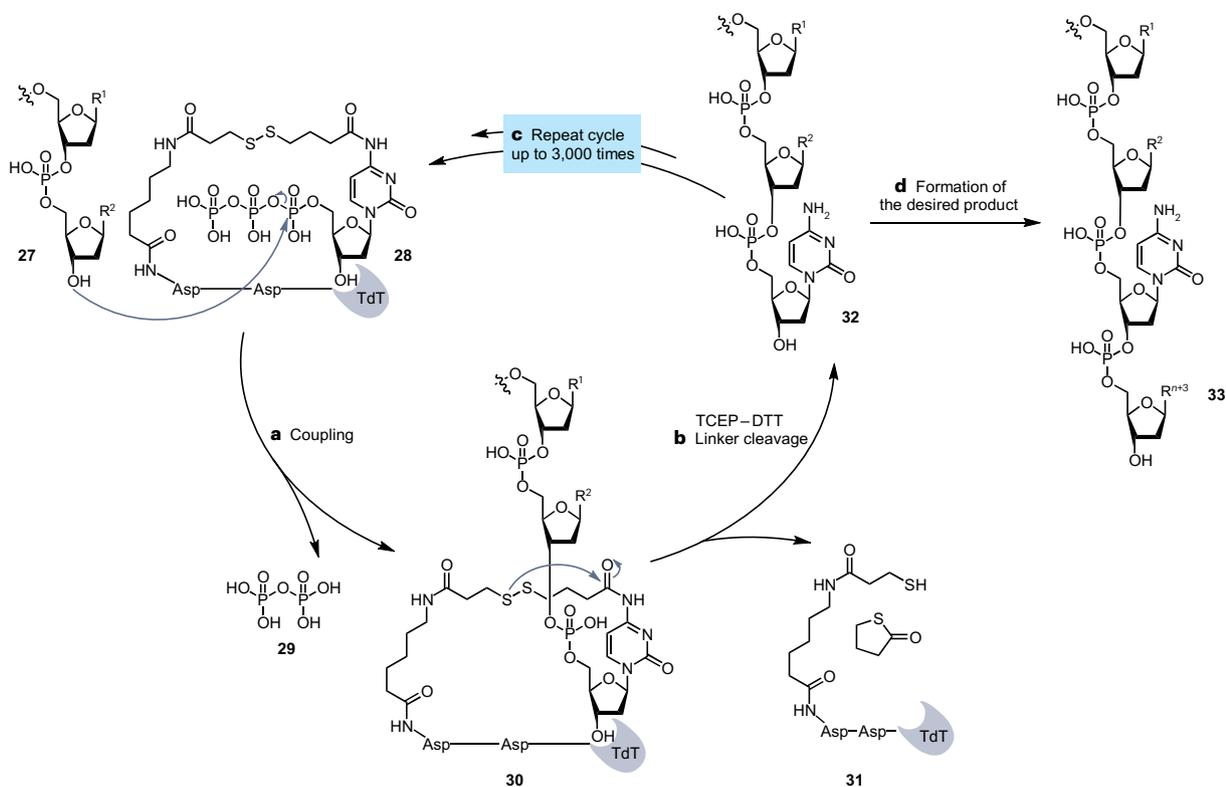


Fig. 3 | Mechanisms for TiEOS. **A**, Schematic representation of 3'-protected nucleoside 5'-triphosphate (NTP) approach. Resin beads are pre-loaded with an initiator DNA (iDNA) **19** to provide a template for binding of terminal deoxynucleotidyl transferase (TdT) and as a post-synthesis cleavage site^{27,72,77,78,95}. Oligonucleotide synthesis then proceeds in a stepwise fashion in the 5'-to-3' direction. TdT ligates NTP **20** to the 3' terminus of the growing oligonucleotide chain with each NTP protected at 3'-OH with a protecting group (PG) **24–26** (refs. ^{27,28,82}). The resin is washed to remove surplus reagents and the pyrophosphate by-product of the ligation. After deblocking or deprotection of the 3'-PG (step **b**), the resin-bound 3'-OH nucleophile of **22** becomes available for the next synthesis cycle (step **c**). The complete sequence is assembled by repeating the cycle of TdT-catalysed NTP(PG) coupling (step **a**) and deblocking (step **b**). On completion, the synthesized oligonucleotide **4** is cleaved from the

solid support (step **d**) by uracil DNA glycosylase. **B**, Examples of NTP(PG)s used in the method – 3'-azidomethyl-protected NTPs **24** by Nuclera Nucleics, Molecular Assemblies, 3'-ONH₂-protected NTPs **25** by DNA Script and 3'-O-2-nitrobenzyl **26** by Camena Bioscience^{97,100,101,108,118}. **C**, Schematic representation for alternative (tethered) protecting strategies. 3'-Unprotected NTPs (cytidine) are supplied pre-immobilized within the TdT-active site **28**, via a short and labile linker^{121,129}. TdT then catalyses the incorporation of this NTP into the growing DNA strand **30** (step **a**) and sterically prevents the uncontrolled polymerization of the NTP until the linker is cleaved (step **b**), releasing the oligonucleotide **32**. The cycle is repeated (step **c**) until the desired oligonucleotide **33** is completed (step **d**). Asp, aspartic acid; DTT, dithiothreitol; TCEP, tris-carboxyethylphosphine; TiEOS, template-independent enzymatic oligonucleotide synthesis.

often requires multiple cloning and repeated Gibson assembly steps causing additional costs and long lead times.

Polymerase cycling assembly

Owing to the development of PCR, the amplification and sequencing of DNA are now routine^{146–148}. Watson–Crick base pairing was used in conjunction with PCR to develop a method for stitching together pools of synthetic oligonucleotides in a technology termed PCA²⁶. In PCA, target oligonucleotides, which are referred to as 'sense', are annealed via complementary overhangs to oligonucleotides corresponding to a complementary, antisense strand (Fig. 4B). Each oligonucleotide, with the exception of those positioned at the 5' termini of each strand, hybridizes with two complementary oligonucleotides within the opposite strand. This produces an annealed construct **39** and **43**, with alternating 'gaps' present within the sense and antisense strands. The gaps are then filled in using a polymerase to generate a duplex DNA template **40** and **41** for PCR amplification. Following this assembly phase, external primers, which are complementary to the 5' ends of the duplex DNA template, are introduced to perform a PCR reaction, which amplifies the target sequence to yield the final product **42**. Using this approach, a plasmid of >2.5 kb has been produced from short, chemically synthesized oligonucleotides²⁶.

Like for Gibson assembly, the performance of PCA can be compromised by impurities of synthetic oligonucleotides. Other disadvantages of PCA include the dependence of the method on sequence confirmation from an individual clone and reliance on high-fidelity proof-reading PCR enzymes which must be used to copy constructed genes to prevent mutations during amplification. Yet, owing to the limitations on the length of iteratively synthesized polynucleotides, Gibson assembly and PCA remain the main practical options for making large DNA.

Emerging commercialized technologies

Companies developing novel ways to make DNA focus on meeting one of two main requirements: longer DNA sequences or greater numbers of DNA constructs made in parallel. Both templated and template-independent approaches are developed for large-scale production and the assembly of long DNA. Increasingly, companies place an emphasis on DNA synthesis services, which remain highly competitive and necessitate tighter control over the distribution of synthetic DNA. Automation offers opportunities to minimize expert involvement in DNA synthesis and is being realized by the supply of benchtop DNA printers. Typically, industry tailors synthetic methods for specific DNA targets, in terms of both complexity and length. This is driven

by challenging and topical applications such as the synthesis of DNA vaccines or gene therapeutics. These applications demonstrate the value of providing DNA products in high yield and purity. A number of exciting developments in industry are discussed subsequently to exemplify the progress in the field of DNA synthesis.

Thermally controlled synthesis

A progressive solution to parallel DNA synthesis, proposed by Evonetix, is thermally controlled synthesis. This method is compatible with both phosphoramidite and TiEOS approaches^{29,51,52,149,150} and offers the synthesis of DNA libraries, with sequences immobilized on discrete thermally controlled reaction sites of silicon chips. Thermal heating allows to selectively cleave PGs (5' for phosphoramidite or 3' for TiEOS, step **a** in Fig. 5), from the termini of specific reaction sites for elongation **47**. The entire chip can then be exposed to a TiEOS or phosphoramidite elongation cycle, selectively elongating only oligonucleotides immobilized on the heated reaction **49**. Unheated sites retain their thermally labile terminal PGs rendering these chains unavailable for elongation **46** and **48** (refs. ^{150–152}) (Fig. 5).

As with other DNA synthesis approaches, the elongation cycle efficiencies are the limiting factor. In each reaction site, a percentage of insufficient thermolysis of PGs is expected. With every elongation cycle, deletion sequences would accumulate creating impurities similar to the desired product. Evonetix addressed this issue by tethering each immobilized oligonucleotide to the chip via a linker that is labile to thermally assisted chemical cleavage. Once the desired strands **53** are assembled, the site on the chip to which they are immobilized is heated resulting in the cleavage of the linker and the liberation of these strands into solution. These liberated strands can be made complementary to oligonucleotides **52**, which remain immobilized on the chip and can be subsequently annealed together to yield double-stranded DNA molecules **54** (ref. ¹⁵⁰). Any imperfectly annealed oligonucleotide pairs **55**, for example, owing to truncated sequences, can be thermally denatured at lower temperatures than the desired DNA **54**. Such a process of thermal purification removes incorrect sequences **56**, yielding a double-stranded DNA product with the desired base pairing **57** (refs. ^{150,152}). If these duplex DNA pairs have sticky ends complementary to strands **58** immobilized on another site of the chip, then sequential pairs can be annealed into a 'nicked' construct **59**. Repetition of this process yields a double-stranded product, the length of which is virtually unlimited. The nicks present in the strand could be repaired by a DNA ligase into the double-stranded DNA of a desired length and the construct may be amplified by PCR. Evonetix is anticipated to offer desktop DNA printers based on this technology. These plug-and-play instruments will feature

A Gibson assembly

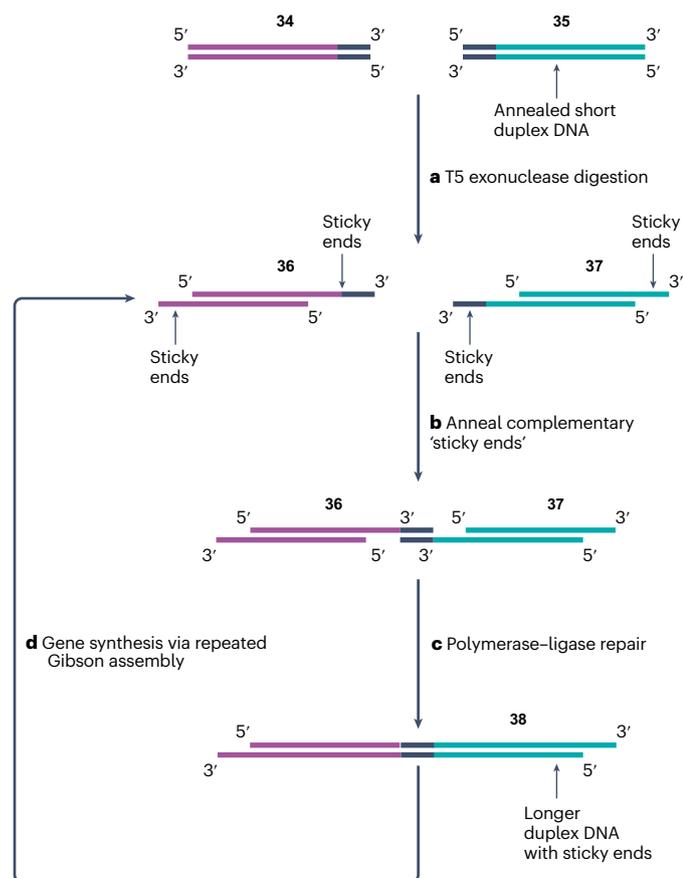
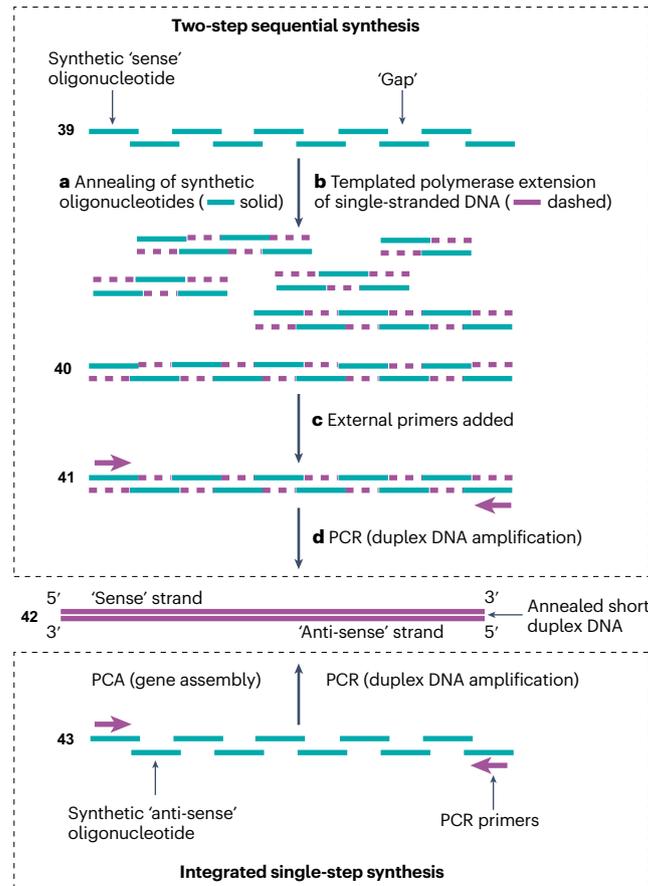


Fig. 4 | Schematic representations of assembly methods for DNA synthesis. **A**, Gibson assembly^{25,139,140}. Two duplex DNA strands **34** and **35** are selected with a complementary terminal overlap region (black). Digestion with T5 exonuclease (step **a**) degrades each strand of the DNA duplexes in the 5′-to-3′ direction, yielding the sticky ends of **36** and **37**, followed by the annealing of complementary sticky ends between the two DNA duplexes¹⁴³ (step **b**). Phusion polymerase and Taq ligase are then combined (step **c**) to ligate the two short DNA duplexes into a single, long DNA duplex construct **38** (refs. ^{144,145}). The process is repeated (step **d**) for gene assembly. **B**, Polymerase cycling assembly (PCA)^{26,141}. High-purity synthetic oligonucleotides **39** are designed, such that annealing of complementary overlaps generates the desired long duplex DNA construct (step **a**). The desired construct is then assembled in either a single step from **43** or two steps from **39** using a DNA polymerase (step **b**) to yield template **40**. PCR amplification (steps **c** and **d**) of **41** amplifies the desired long duplex DNA construct **42**.

B PCA



long DNA duplex construct **38** (refs. ^{144,145}). The process is repeated (step **d**) for gene assembly. **B**, Polymerase cycling assembly (PCA)^{26,141}. High-purity synthetic oligonucleotides **39** are designed, such that annealing of complementary overlaps generates the desired long duplex DNA construct (step **a**). The desired construct is then assembled in either a single step from **43** or two steps from **39** using a DNA polymerase (step **b**) to yield template **40**. PCR amplification (steps **c** and **d**) of **41** amplifies the desired long duplex DNA construct **42**.

user interfaces and design algorithms implemented in the cloud to enable control over biosecurity of gene synthesis⁵².

Gene synthesis from libraries

Ribbon Biolabs has developed a convenient synthesis of long (>10 kb) duplex DNA, using the convergent assembly of double-stranded oligonucleotide pools¹⁵³ (Fig. 6A). The methodology requires the synthesis of a library of tens of thousands 5′-phosphorylated single-stranded oligonucleotides of high purity and 8–26 nucleotides in length, encompassing all the necessary building blocks for DNA synthesis¹⁵⁴. Each oligonucleotide has a designated 5′-phosphorylated reverse complement strand in the library, with annealing overhangs of four nucleotides designed for each strand at the 5′-end. The assembly process requires the denaturation and annealing of pair of complementary oligonucleotides **60** and **61** and **62** and **63** to generate a library of duplex DNA constructs **64** and **65**, each with two four-nucleotide sticky ends at the 5′ termini of both strands¹⁵³.

Duplex DNA fragments **64** and **65** with 5′-phosphorylated, four-nucleotide sticky ends are then annealed in step **b** and ligated in step **c** together in a convergent synthesis giving rise to larger duplex DNA **67** for further assembly. Repeated cycles of the annealing and ligation of these building blocks give the desired duplex construct¹⁵⁵. Terminal duplex DNA blocks have a single ‘blunt end’ and a single ‘sticky end’ to yield linear duplex DNA products. Once the final DNA duplex is obtained, it can be amplified by PCR using a high-fidelity polymerase to provide product yield for the customer. This technology from Ribbon Biolabs is therefore analogous to a convergent Gibson assembly approach²⁵.

Gene synthesis from DNA microarrays

A similarly effective approach has been developed by Twist Bioscience through miniaturizing and performing gene synthesis onto a silicon microarray chip^{47–50}. A grid of 25,000 discrete reaction sites is generated using an ink jet printing. Specialist reagents are then delivered to each site. The method enables the selective elongation of several

desired sequences out of a library of tens of thousands with improved elongation efficiencies (Fig. 6B). Low concentrations and volumes used in the method (approximately femtomole) permit starting reagents to be used in a large excess, whereas the acidic 5' detritylation solution is neutralized by basic oxidation to prevent depurination^{47,156–163}. Assembled sequences are produced in relatively low quantities, which necessitates the use of PCR to generate sufficient DNA for gene assembly³⁰. Microarrays provide complex pools of DNA 68, which can include both strands of a complementary duplex. After annealing, the duplex is used for template-specific 69 and 70 amplification by PCR to generate larger quantities of selected sequences 71 and 72, respectively, which

hybridize efficiently with primers^{164–167}. In this format, duplex DNA can be selectively amplified from a complex pool of sequences in parallel. The assembly subpools of double-stranded DNAs 71 and 72 amplified by PCR are then digested by type IIS restriction endonucleases to generate sticky-ended duplexes 73 and 74, respectively. These duplexes are then used as building blocks for Gibson assembly to assemble desired genes 75 and 76 at a fraction of the cost of traditional column synthesized oligonucleotides^{30,168}. Miniaturization has benefitted other areas too. Notably, microarrays have proven instrumental for optimizing the parallel synthesis of oligonucleotides on TiOES platforms, including the impact of initiating strands and chemically modified NTPs on enzymatic

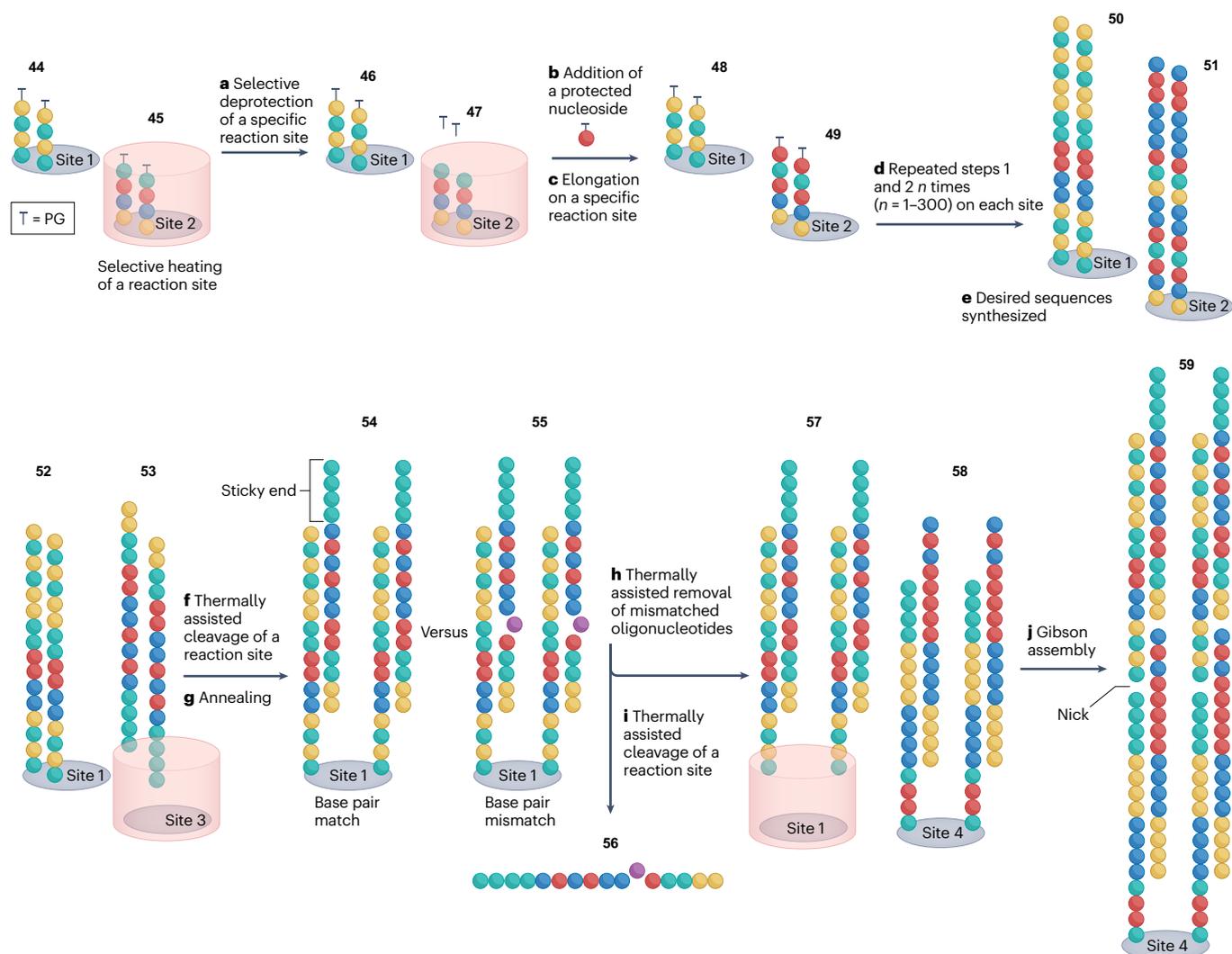
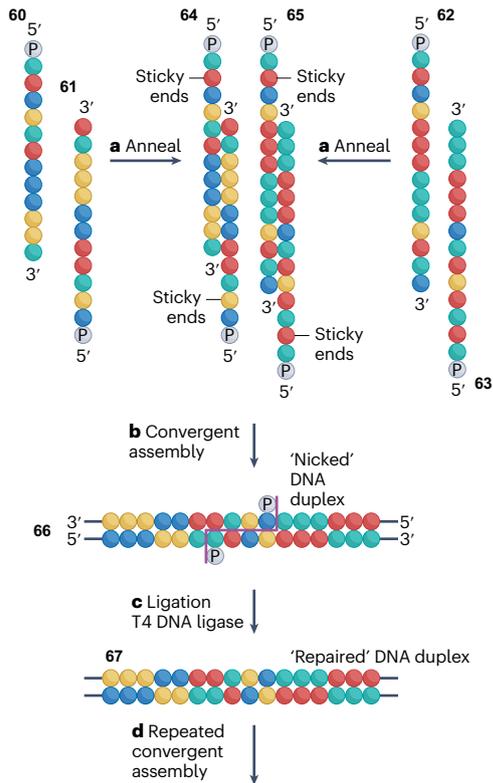


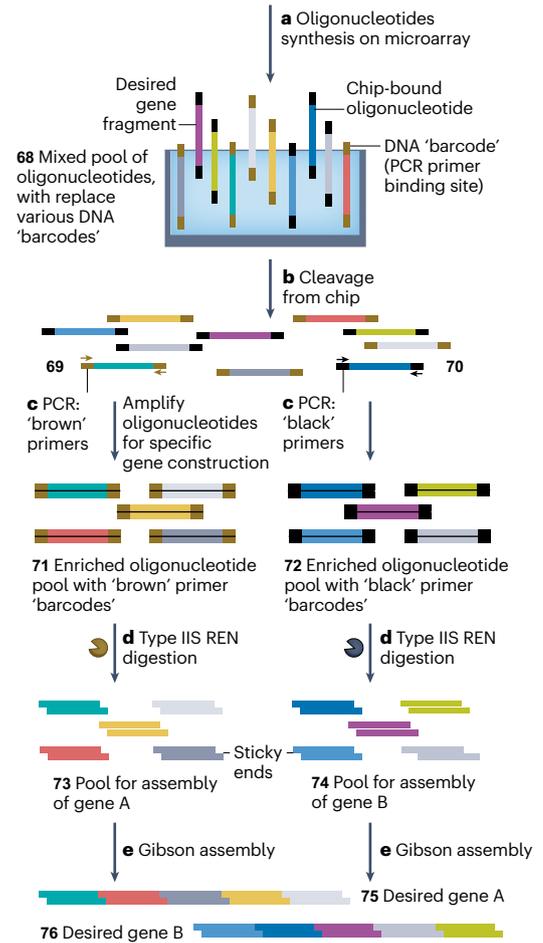
Fig. 5 | Schematic representation of thermally controlled DNA synthesis. Terminally protected oligonucleotide strands 44 and 45 are immobilized on discrete reaction sites (sites 1 and 2). Thermal heating of a chosen site (site 2) cleaves terminal protecting groups (PGs) 47 (step a), enabling selective elongation of strands on this site^{29,31,52,151} (steps b and c). Elongation of a desired oligonucleotide is performed via TiOES or the phosphoramidite method to selectively generate 49 (refs. 27,37). Repeated steps a–c on other selected reaction sites (for example, site 1) sequentially produce bespoke oligonucleotides 50 (steps d and e). Thermally assisted reagent treatment of selected sites cleaves safety catch linkers and liberates oligonucleotides 53 from a chosen

site^{29,151} (step f). The liberated oligonucleotides (step g) anneal to complementary chip-bound oligonucleotides 52 producing perfectly annealed double-stranded DNA 54, which has a higher denaturation temperature than that of DNA duplexes formed by oligonucleotides with mismatches 55. The heating of reaction site 1 allows the mismatched oligonucleotides 56 to be washed away¹⁵⁰ (step h). Thermally assisted reagent treatment of site 1 cleaves safety catch linkers and releases the desired duplex DNA 57 from the chip^{29,151} (step i). This liberated DNA is annealed to a chip-bound complementary DNA duplex 58 to form the nicked construct 59 (step j). The process is repeated to elongate the double-stranded DNA until the desired gene is assembled.

A Gene synthesis from diverse oligonucleotide libraries



B Gene synthesis from DNA microarrays



C Rolling circle amplification technique

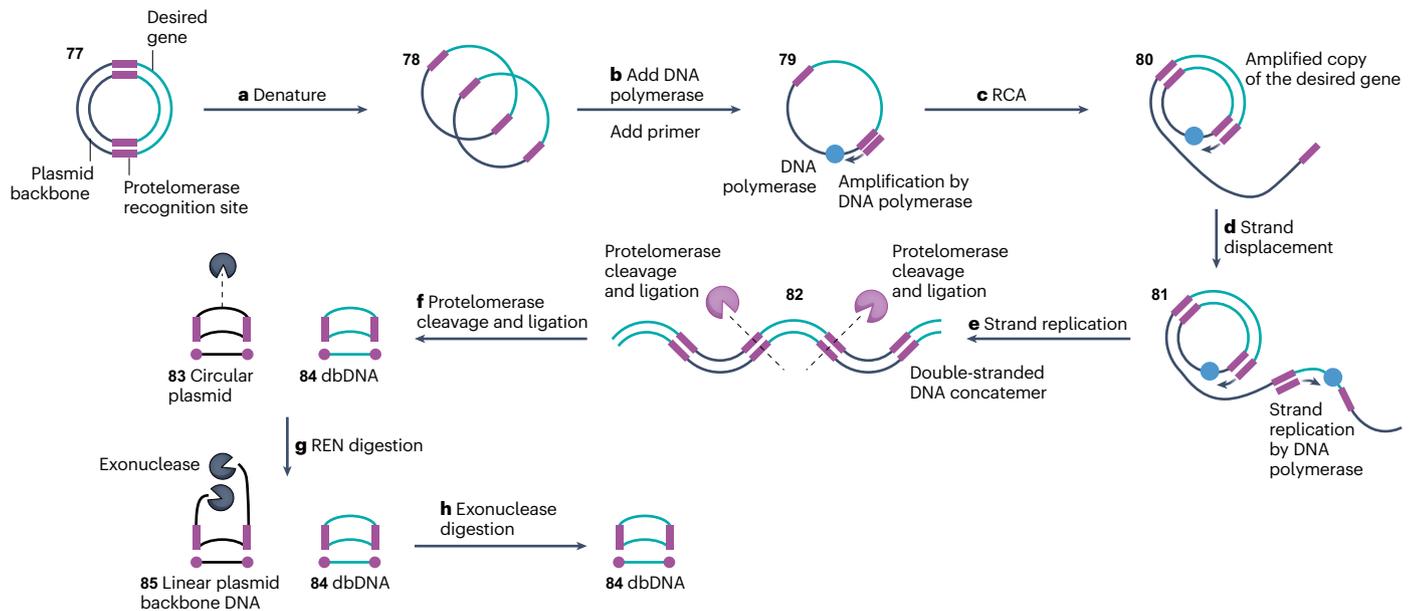


Fig. 6 | Gene synthesis strategies. **A**, Gene synthesis from diverse oligonucleotide libraries. 5'-Phosphorylated sense oligonucleotide strands **60** and **61** are annealed (step **a**) to complementary 5'-phosphorylated antisense strands **62** and **63**. The resulting DNA duplexes **64** and **65** have two 5'-overhangs or sticky ends that are used to anneal (step **b**) the duplexes into an extended, 'nicked' duplex **66** ('nick' highlighted in magenta). T4 DNA ligase is then used (step **c**) to stitch the oligonucleotides at the nick site into an elongated, larger DNA duplex **67**. Cycles of annealing and ligation are repeated until the desired gene is assembled^{153–155} (step **d**). **B**, DNA microarrays³⁰. A library of bespoke single-stranded oligonucleotides **68** is generated with 3'-terminal and 5'-terminal DNA 'barcodes' on a miniaturized chip^{30,47,48,50,164,166,167} (step **a**). These sequences are cleaved (step **b**) from the microchip to yield a pool of template oligonucleotides with a range of DNA 'barcodes' (only two, black and brown, 'barcodes' are shown for clarity)^{164,167}. Primers selectively anneal to either the 'brown' **69** or 'black' **70** DNA barcodes and specifically amplify oligonucleotides via PCR (step **c**), according to the identity of the barcode at its 3' and 5' termini^{146–148}. The resulting duplex DNA constructs

71 and **72** still contain the DNA barcodes at their termini, which must be removed prior to gene assembly. DNA barcodes are cleaved (step **d**) from the duplex DNA **71** and **72** by type IIS restriction endonucleases (REN), giving rise to assembly pools of sequences **73** and **74** with sticky ends³⁰. Duplex DNA fragments are annealed (step **e**) via complementary sticky ends and assembled into desired genes **75** and **76** via Gibson assembly^{25,139,140}. **C**, Rolling circle amplification (RCA)¹⁷⁵. Template plasmid DNA **77** with a desired gene cassette (green) and protelomerase sites (magenta) is thermally denatured (step **a**) to create a single-stranded template **78** (ref. ¹⁷⁸). A complementary primer binds to the protelomerase sites **79** (step **b**) and the template is amplified via RCA (steps **c–e**), to produce double-stranded concatemeric DNA **82** with alternating copies of the desired cassette (green) and the unwanted plasmid backbone (black)^{176,181–183}. Protelomerase then cuts (step **f**) the duplex at its recognition sites and ligates the cut ends generating covalently closed 'doggybone' DNA (dbDNA) **84** and a circular plasmid DNA **83** as a by-product. The circular backbone of the plasmid DNA is subsequently cut (step **g**) by REN and digested (step **h**) by exonucleases¹⁷⁹.

DNA synthesis. Microarrays have also prompted early considerations for DNA nanofabrication, synthesis multiplexing and compatibility of enzymes with alternative polymerization methods^{169–172}.

Rolling circle amplification technique

Touchlight Genetics has commercialized a technology to scale up the manufacturing of large DNA using a linear closed 'doggybone' DNA (dbDNA), named after its structure **84** resembling a doggy bone (Fig. **6C**). dbDNA is produced via rolling circle amplification from a plasmid template **77** (refs. ^{173–176}). The template must be engineered to contain desired expression cassettes, for example, inverted terminal repeats, directly between two 28 nucleotide protelomerase recognition sites^{177–179}. The denatured plasmid template **78** is amplified by DNA polymerase, in the presence of a primer that binds to the protelomerase recognition sites **79**. Once the plasmid has been replicated, the DNA polymerase continues to repeatedly replicate the plasmid template via rolling circle amplification, displacing any pre-existing synthesized strands from the template **81**. The polymerase then binds to these liberated single strands of DNA and replicates the complementary strand to generate concatemeric double-stranded DNA **82**. Protelomerase is added to generate a double-stranded DNA break and form a hairpin loop to re-seal the ends resulting in dbDNA **84** and a circularized plasmid DNA by-product **83** (refs. ^{179–183}).

Restriction endonucleases are carefully selected such that they can digest the unwanted plasmid backbone **83**, but their restriction sites are not present in dbDNA **84** (ref. ¹⁸⁰). Digestion of the products of the reaction liberates dbDNA **84** and the undesired linearized plasmid backbone **85**. A subsequent digestion of this mixture with exonucleases produces a mixture of nucleotides, enzymes and buffers, which can be readily separated from the desired product. dbDNA **84** is then purified to provide a minimal, linear DNA vector encoding virtually any long sequence of interest. These sequences can be either complex or unstable and can be re-amplified using the same process to rapidly generate multigram quantities of large DNA free of bacterial or endotoxin contaminations¹⁷⁶. As a manufacturing platform, this approach permits the production of large DNA five times faster than traditional fermentation methods (Table **1**).

Business models and deskilling

Approaches combining chemical and enzymatic syntheses, sequence selection and assembly are set to undergo continuous development. However, as the underpinning chemistry for synthetic DNA is unlikely to

change markedly, the elongation cycle efficiency remains the main limiting factor. This has prompted companies to develop complementary capabilities such as highly parallelized, miniaturized and automated synthesis, while promoting user autonomy in producing DNA (Fig. **7**).

Automation and services

Research focuses on improving the synthesis of DNA sequences in parallel. This is known to increase the probability of errors, in particular for sequences that are difficult to amplify, such as repeat or GC-rich sequences. Automation provides a compelling direction. Companies exploit advances in other areas, such as electronics and microfluidics to improve DNA synthesis. This has aided the identification and removal of errors, increased accuracy, scale and speed to a far greater extent than non-automated approaches^{29,51,52}. For example, Evonetix developed a platform for high fidelity and rapid gene synthesis, which is controlled by electrochemical processing of each of many thousands of independent reaction sites on a silicon chip, in a highly parallelized fashion. The combination of parallel synthesis and site-specific thermal control has the potential to address limitations of difficult sequences. For instance, sequences with a high GC content, which require higher melting temperatures than other sequences and can form stable secondary structures, can be synthesized at elevated temperatures. However, when using such temperatures, high site specificity is necessary to prevent mis-annealing (Table **1**).

With several enzyme companies now active in this space, various business models have emerged with an increasing emphasis on DNA assembly and benchtop printers. For example, the size of a DNA Script's Syntax instrument offered as a benchtop DNA synthesizer is similar to that of a HiSeq sequencer developed by Illumina. This synthesizer can generate 60 bp oligonucleotides in a pure form for immediate use within 6 h.

As there is moderate progress in increasing elongation cycle efficiencies, the use of microarray technologies to produce multiple sequences in parallel is being developed. Although further investment may be required to develop these technologies, they will improve DNA synthesis. Although new, more effective technologies can be expected to boost and dominate the market, the emergence of one winning technology that will be pursued by one vendor is unlikely. Ultimately, the development of several similarly effective technologies will ensure that DNA is priced similarly by all vendors, for example, per gene or length, making the supply of DNA of unlimited lengths affordable for the end user. Every technology is a matter of specialist developments

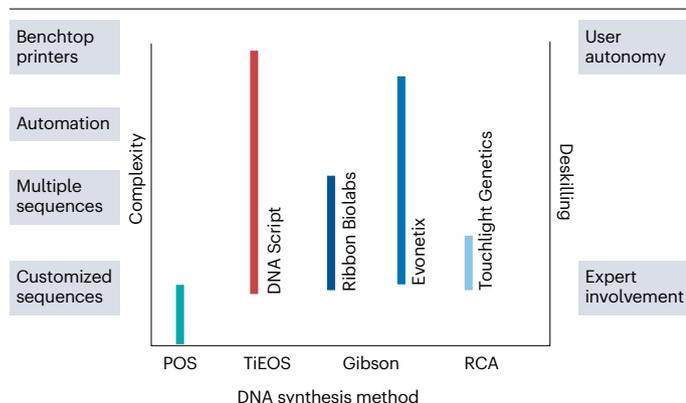


Fig. 7 | User autonomy in DNA synthesis. DNA synthesis technologies plotted versus complexity levels of the current offerings (left axis) by different companies who develop these methods and versus user autonomy (deskilling) that these companies offer (right axis), from expert involvement to complete autonomy for the end user. For example, template-independent enzymatic oligonucleotide synthesis (TIEOS) is the basic technology for DNA Script who offer customized and multiple sequences, while integrating automation into their synthesis workflows, which results in the development of benchtop DNA printers that the end user can buy and use with a minimum expert involvement. Gibson, Gibson assembly approaches; POS, phosphoramidite oligonucleotide synthesis; RCA, rolling circle amplification.

but must eventually subject to automation, reducing the dependence of the end user on the expert involvement and deskilling DNA synthesis (Fig. 7).

Barriers to entry for customers

Custom DNA synthesis remains an expensive endeavour (for example, US\$300–1,000 per 3 kb gene or \$0.1–0.3 b^{-1}). Prices vary depending on vendor, sequence composition and length. A general trend is observed towards the decrease of price to \$0.01 b^{-1} for gene synthesis over several years¹⁵, for example, the current price offered by Twist Bioscience is \$0.07 b^{-1} for gene fragments. More significant funding is required to aid research aiming to make large DNA. More specialized equipment is required for the end users to make DNA that is more complex than plasmids. The provision of such complex and large DNA can be outsourced to DNA synthesis providers (for example, Ribbon Biolabs for assembly). The complexity of custom DNA made for a particular application defines the skill barrier required for the synthesis. There are general trends for reducing the dependence on expert involvement by reducing the need to troubleshoot the DNA synthesis, which is achieved by advances in the performance of enzymes and DNA assembly methods.

Improved access to DNA in bulk quantities and enhanced information capacity of genome-sized DNA may promote further demand. Therefore, the limitation on the lengths available remains the main area of improvement to scale up. The demand for large DNA is anticipated to increase once the length limit of sequences has been overcome. Approaches exploiting automated on-chip gene assembly are promising solutions. Longer DNA will be more costly to produce. However, it is reasonable to expect that with more technologies able to break the size limit and more companies able to supply large DNA, the prices for synthetic DNA will be driven down. DNA storage applications may provide exceptions as these require substantial amounts of starting materials ($g\ kg^{-1}$) to produce larger DNA than required for a biological

application. The quantity and length of DNA needed are related to the amount of information to be stored¹⁸⁴.

Laboratory requirements

An increasing range of benchtop printers will make in-house DNA synthesis viable: Cytiva's ÄKTA Oligopilot provides up to 8 oligonucleotides per 3–4-h run, Kilobaser affords 2 oligonucleotides per 2-h run and Syntax, the first enzymatic printer, produces 96 oligonucleotides in parallel within 6 h. DNA ordered via service providers and manipulated in house may be assembled in small volume reactions, helping to scale down experimentation via miniaturization. Further benefits to reduce costs and lead times for DNA constructs have arisen from automation, miniaturization and parallelization of assembly methods, whereas the accurate sequence verification of assembled DNA benefits from the efficacy of NGS. However, a conservative estimate for an entry-level DNA synthesis laboratory starts at \$200,000, which may increase depending on the length and production scale of the desired DNA.

In this regard, biofoundries provide a complementary infrastructure support for the end-users. Built on strong high-throughput handling and analysis technology platforms, these facilities may establish rapid in-house production pipelines for long double-stranded DNA and diverse variant libraries. Biofoundries typically host a version of a SynBio stack – an ecosystem of technologies, which allows to tackle a complex task by breaking it down to smaller tasks, providing the context and purpose of these tasks in the automated workflows to manipulate, assemble, analyse and organize DNA in small volumes and high throughput¹⁸⁵.

DNA storage and accessibility

Starting with the first book written in DNA¹⁸⁶, there has been steady interest in applying DNA to store and preserve data generated across different sectors of the society^{186–192}. The concept of permanent, compact and low energy data storage in DNA gains traction, notably through the DNA Data Storage Alliance – an industry association which seeks to create a data storage ecosystem using DNA as a medium¹⁹⁰. The data of interest are encoded in the four-letter alphabet of DNA (that is, A, C, G, T), whereas a set of high-fidelity enzymes is used to create copies of these data and an accurate sequencing technology is used to retrieve it. For example, Catalogue Technologies and Cambridge Consultants built a DNA synthesizer, which was able to encode 16 GB of Wikipedia data¹⁸⁹. The ability to design encoding rules for data storage at which may offer an elegant way to accommodate error rates or avoid specific sequence motifs that might be difficult for a given synthesis or sequencing technology. For instance, the recent expansion of the genetic alphabet by Hachimoji bases¹⁹³ creates the prospect of encoding data with eight instead of four letters of the DNA alphabet. This allows for an exponential rise in data density, with the help of engineered enzymes to incorporate, copy and read bases of such an expanded genetic code. These developments mean that DNA data storage products are possible and might rival biological applications as the main use of DNA synthesis technologies.

Oversight and standardization

The relatively new ability to make long pieces of DNA may prove impactful in genetic manipulation and control over living systems, which requires oversight and regulation worldwide. Regulatory policies have been developed in related areas; for example, through the Asilomar Conference on Recombinant DNA research, where communities have introduced self-regulatory processes for biosafety regulations¹⁹⁴.

Ultimately, the ability to write DNA will become accessible to non-experts. Therefore, there is a growing recognition that oversight policies are needed to mitigate the biosecurity risks of misusing DNA technologies. The introduction of new policies helps adapt existing mechanisms to assess these and evolving risks¹⁹⁵. For example, an International Gene Synthesis Consortium was formed by industry to develop a common protocol to screen synthetic sequences, as well as the customers who order these sequences, thus self-regulating sequence identity¹⁹⁶. However, DNA technology has advanced more rapidly than the ability to understand, monitor and regulate the risks. This is similar to CRISPR gene editing, in which the ability to engineer genomes was taken up by the end users before regulators had understood the consequences of misusing the technology for human genome editing¹⁹⁷.

Natural DNA can become self-sustaining when incorporated into organisms which can subsequently embed into ecosystems permanently, raising issues around horizontal gene transfer as a key part of evolution¹⁹⁸. Therefore, there is a need for a considered response and oversight of ethical factors, regulation and other risks associated with DNA synthesis. Similarly, maximizing the reproducibility and reliability of DNA synthesis is vital. Of particular importance is meeting emerging regulations which require industry to demonstrate the traceability of their products and technologies. This increasing focus on reproducibility and traceability prioritizes the need for standardization. The lack of reference materials and methods against which the performance and quality of synthetic DNA and synthesis methods can be evaluated including novel chemistries such as XNA is a key challenge. Reference materials can include DNA sequences, individual or libraries, which are traceable to the International System of Units (Système International d'Unités – the SI). Reference methods may provide synthesis procedures to benchmark the performance of commercial methods, for example, in relation to elongation cycle efficiency. Encouragingly, reference materials have become available, such as the first 'human genome' DNA reference material (RM 8398) from the National Institute of Standards and Technology, which can evaluate the accuracy of NGS assays¹⁹⁹. New metrology, which will provide the basis of comparison and reproducibility for DNA synthesis, is required to support existing and emerging DNA synthesis technologies.

Conclusions

With the rapid development of technologies able to read DNA, the ability to write DNA has lagged behind. DNA synthesis technologies developed to date may differ in their ability to bridge the DNA writing gap (Table 1 and Fig. 1). However, their continuous development is driven by two main factors: the lack of methods to routinely make DNA of unlimited lengths at scale and cost and an increasing demand for DNA from different and unrelated sectors. Alongside the success of NGS, these two factors stimulate the search for innovative technologies and guarantee commercialization success for any strategy able to overcome the barrier of size-limited DNA synthesis. As a consequence, DNA makers tightly guard their knowledge and are cautious in their claims of what their technologies can deliver. This is notable given that most of the existing methodologies are similar, using the same starting materials, suggesting that innovation progresses at a marginal pace. Conversely, close competition prompts companies to search for an application niche at early stages or demonstrate the use of their technologies in producing challenging DNA molecules.

Advances in automation will make DNA synthesis increasingly more accessible to non-experts. Most vendors, especially those who provide synthetic DNA as a service, appreciate the need for oversight and

regulatory policies to protect their commercial and reputational interests and may, in turn, contribute to the development of such policies. Once DNA synthesis is affordable for small hackerspaces of enthusiasts collaborating on making new DNA molecules, the uses of the produced DNA will be difficult to contain. Risk governance designed to monitor the use and distribution of synthetic DNA in accordance with applicable policies and ethics will reduce the likelihood of adverse events.

To conclude, new synthesis methods will continue to emerge with a persistent focus on providing greener solutions, mitigating potentially harmful consequences for the environment owing to the use of organic solvents and hazardous chemicals. With the limitations of existing synthetic approaches, it is unlikely that a routine methodology to effectively synthesize size-unlimited DNA will soon be available. Yet, there remains plenty of space in the gene writing gap for breakthroughs in the foreseeable future.

Published online: 23 January 2023

References

- Hughes, R. A. & Ellington, A. D. Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb. Perspect. Biol.* **9**, a023812 (2017).
- Kosuri, S. & Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **11**, 499–507 (2014).
- Rothmund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
- Pinheiro, V. B. & Holliger, P. Towards XNA nanotechnology: new materials from synthetic genetic polymers. *Trends Biotechnol.* **32**, 321–328 (2014).
- Lee, I. & Berdis, A. J. Non-natural nucleotides as probes for the mechanism and fidelity of DNA polymerases. *BBA Proteins Proteom.* **1804**, 1064–1080 (2010).
- Douglas, S. M., Bachelet, I. & Church, G. M. A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **335**, 831–834 (2012).
- Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
- Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17**, 333–351 (2016).
- Schatz, M. C. & Phillippy, A. M. The rise of a digital immune system. *Gigascience* **1**, 14 (2012).
- McLennan, A. *Regulation of Synthetic Biology* (Edward Elgar, 2018).
- Schwab, K. The fourth industrial revolution: what it means and how to respond. *Foreign Affairs* <https://www.foreignaffairs.com/world/fourth-industrial-revolution> (2015).
- Leurent, H. & Abbosh, O. Shaping the sustainability of production systems: fourth industrial revolution technologies for competitiveness and sustainable growth (World Economic Forum, 2019).
- Chui, M., Evers, M., Manyika, J., Zheng, A. & Nisbet, T. *The Bio Revolution* (McKinsey Global Institute, 2020).
- Collins, F. S. et al. Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945 (2004).
- Carlson, R. in *Synthetic Biology: Parts, Devices and Applications* Ch. 1 (ed. Smolke, C.) 3–13 (Wiley-Blackwell, 2018).
- The first detailed analysis of gene sequencing and gene synthesis technologies in light of Moore's law highlighting the gap between our ability to read and write DNA.**
- Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
- Blakney, A. K. & Bekker, L. G. DNA vaccines join the fight against COVID-19. *Lancet* **399**, 1281–1282 (2022).
- Han, S. H., Park, Y. J. & Park, C. M. HOS1 activates DNA repair systems to enhance plant thermotolerance. *Nat. Plants* **6**, 1439–1446 (2020).
- Brophy, J. A. N. et al. Synthetic genetic circuits as a means of reprogramming plant roots. *Science* **377**, 747–751 (2022).
- Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
- Cumbers, J. A New Way of Making DNA is About to Revolutionize the Biotech Industry (Forbes, 2020).
- Reisch, M. S. Danaher buys oligonucleotide maker Integrated DNA Technologies. *Chem. Eng. News* **96**, 10–10 (2018).
- Eisenstein, M. Enzymatic DNA synthesis enters new phase. *Nat. Biotechnol.* **38**, 1113–1115 (2020).
- Mohsen, M. G. & Kool, E. T. The discovery of rolling circle amplification and rolling circle transcription. *Acc. Chem. Res.* **49**, 2540–2550 (2016).
- Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
- The seminal paper introducing a revolutionary method enabling the assembly of multiple DNA fragments in a single reaction — Gibson assembly.**

26. Stemmer, W. P., Cramer, A., Ha, K. D., Brennan, T. M. & Heyneker, H. L. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **164**, 49–53 (1995).
27. Jensen, M. A. & Davis, R. W. Template-independent enzymatic oligonucleotide synthesis (TIEOS): its history, prospects, and challenges. *Biochemistry* **57**, 1821–1832 (2018).
28. Sarac, I. & Hollenstein, M. Terminal deoxyribonucleotidyl transferase in the synthesis and modification of nucleic acids. *ChemBioChem* **20**, 860–871 (2018).
29. Crosby, S. et al. Oligonucleotide and nucleic acid synthesis. Patent WO/2019/145713 (2019).
30. Kosuri, S. et al. Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nat. Biotechnol.* **28**, 1295–1299 (2010).
31. Kodumal, S. J. et al. Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proc. Natl. Acad. Sci. USA* **101**, 15573–15578 (2004).
32. Michelson, A. M. & Todd, A. R. Nucleotides part XXXII. Synthesis of a dithymidine dinucleotide containing a 3': 5'-internucleotidic linkage. *J. Chem. Soc.* **1**, 2632–2638 (1955).
33. Nishimura, S., Jones, D. S. & Khorana, H. G. Studies on polynucleotides. 48. The in vitro synthesis of a co-polypeptide containing two amino acids in alternating sequence dependent upon a DNA-like polymer containing two nucleotides in alternating sequence. *J. Mol. Biol.* **13**, 302–324 (1965).
34. Merrifield, R. B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
35. Letsinger, R. L. & Mahadevan, V. Oligonucleotide synthesis on a polymer support. *J. Am. Chem. Soc.* **87**, 3526–3527 (1965).
- Inception paper for chemical oligonucleotide synthesis.**
36. Beaucage, S. L. & Caruthers, M. H. Deoxyribonucleoside phosphoramidites — a new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **22**, 1859–1862 (1981).
- Introduction of phosphoramidite method for oligonucleotide synthesis that continues to underpin DNA synthesis today.**
37. Matteucci, M. D. & Caruthers, M. H. Synthesis of deoxyoligonucleotides on a polymer support. *J. Am. Chem. Soc.* **103**, 3185–3191 (1981).
38. McBride, L. J. & Caruthers, M. H. An investigation of several deoxyribonucleoside phosphoramidites useful for synthesizing deoxyoligonucleotides. *Tetrahedron Lett.* **24**, 245–248 (1983).
39. Vinayak, R. Chemical synthesis, analysis, and purification of oligoribonucleotides. *Methods* **5**, 7–18 (1993).
40. Caruthers, M. H. et al. Chemical synthesis of deoxyribonucleotides by the phosphoramidite method. *Method. Enzymol.* **154**, 287–313 (1987).
41. Letsinger, R. L., Caruthers, M. H. & Jerina, D. M. Reactions of nucleosides on polymer supports. Synthesis of thymidylthymidylthymidine. *Biochemistry* **6**, 1379–1388 (1967).
42. Caruthers, M. H. The chemical synthesis of DNA/RNA: our gift to science. *J. Biol. Chem.* **288**, 1420–1427 (2013).
43. Caruthers, M. H. & Beaucage, S. L. Phosphoramidite compounds and processes. US Patent 4,415,732 (1983).
44. Caruthers, M. H. & Matteucci, M. D. Nucleosides useful in the preparation of polynucleotides. US Patent 4,500,707 (1985).
45. Caruthers, M. H. & Beaucage, S. L. Phosphoramidite nucleoside compounds. US Patent 4,668,777 (1987).
46. Pon, R. T. Solid-phase supports for oligonucleotide synthesis. *Curr. Protoc. Nucleic Acid Chem.* **3**, 1–28 (2000).
47. LeProust, E. M. et al. Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process. *Nucleic Acids Res.* **38**, 2522–2540 (2010).
- High impact methodology for generating libraries of small DNA sequences while avoiding depurination during synthesis.**
48. Hughes, T. R. et al. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.* **19**, 342–347 (2001).
49. LeProust, E., Zhang, H., Yu, P., Zhou, X. & Gao, X. Characterization of oligodeoxyribonucleotide synthesis on glass plates. *Nucleic Acids Res.* **29**, 2171–2180 (2001).
50. Cleary, M. A. et al. Production of complex nucleic acid libraries using highly parallel in situ oligonucleotide synthesis. *Nat. Methods* **1**, 241–248 (2004).
51. Ferguson, A. J. et al. Thermofluidic chip containing virtual thermal wells. *Eng. Biol.* **3**, 20–23 (2019).
52. Crosby, S. Unlocking synthetic biology through DNA synthesis. *Chim. Oggi* **38**, 22–24 (2020).
53. Letsinger, R. L. & Ogilvie, K. K. Nucleotide chemistry. XIII. Synthesis of oligothymidylates via phosphotriester intermediates. *J. Am. Chem. Soc.* **91**, 3350–3355 (1969).
54. Letsinger, R. L., Finnan, J. L., Heavner, G. A. & Lunsford, W. B. Nucleotide chemistry. XX. Phosphite coupling procedure for generating internucleotide links. *J. Am. Chem. Soc.* **97**, 3278–3279 (1975).
55. Letsinger, R. L. & Lunsford, W. B. Synthesis of thymidine oligonucleotides by phosphite triester intermediates. *J. Am. Chem. Soc.* **98**, 3655–3661 (1976).
56. Smith, M., Rammler, D. H., Goldberg, I. H. & Khorana, H. G. Studies on polynucleotides. XIV. Specific synthesis of the C3'-C5' interribonucleotide linkage. Syntheses of uridylyl-(3'→5')uridine and uridylyl-(3'→5')adenosine. *J. Am. Chem. Soc.* **84**, 430–440 (1962).
57. Schaller, H., Weimann, G., Lerch, B. & Khorana, H. G. Studies on polynucleotides. XXIV.1 The stepwise synthesis of specific deoxyribopolynucleotides. Protected derivatives of deoxyribonucleosides and new syntheses of deoxyribonucleoside-3' phosphates. *J. Am. Chem. Soc.* **85**, 3821–3827 (1963).
58. Krotz, A. H. et al. Solution stability and degradation pathway of deoxyribonucleoside phosphoramidites in acetonitrile. *Nucleosides Nucleotides Nucleic Acids* **23**, 767–775 (2004).
59. Caruthers, M. & Matteucci, M. D. Process for preparing polynucleotides. US Patent 4,458,066 (1984).
60. Caruthers, M. H. & Beaucage, S. L. Process for oligonucleotide synthesis using phosphoramidite intermediates. US Patent 4,973,679 (1990).
61. Koster, H. & Sinha, N. D. Process for the preparation of oligonucleotides. US Patent 4,725,677 (1988).
62. Efcavitch, J. W. & Heiner, C. Depurination as a yield decreasing mechanism in oligodeoxynucleotide synthesis. *Nucleosides Nucleotides* **4**, 267 (1985).
63. An, R. et al. Non-enzymatic depurination of nucleic acids: factors and mechanisms. *PLoS ONE* **9**, e0115950 (2014).
64. Suzuki, T., Ohsumi, S. & Makino, K. Mechanistic studies on depurination and apurinic site chain breakage in oligodeoxyribonucleotides. *Nucleic Acids Res.* **22**, 4997–5003 (1994).
65. Ravikumar, V., Andrade, M., Mulvey, D. & Cole, D. L. Carbocation scavenging during oligonucleotide synthesis. US Patent 5,510,476A (1996).
66. Wooddell, C. I. & Burgess, R. R. Use of asymmetric PCR to generate long primers and single-stranded DNA for incorporating cross-linking analogs into specific sites in a DNA probe. *Genome Res.* **6**, 886–892 (1996).
67. Veneziano, R. et al. Designer nanoscale DNA assemblies programmed from the top down. *Science* **352**, e1534 (2016).
68. Veneziano, R. et al. In vitro synthesis of gene-length single-stranded DNA. *Sci. Rep.* **8**, e6548 (2018).
69. Mitchell, L. A. et al. Synthesis, debugging, and effects of synthetic chromosome consolidation: synVI and beyond. *Science* **355**, e1045 (2017).
70. Reddy, M. P., Hanna, N. B. & Farooqui, F. Ultrafast cleavage and deprotection of oligonucleotides synthesis and use of C^{ac} derivatives. *Nucleosides Nucleotides* **16**, 1589–1598 (1997).
71. Delarue, M. et al. Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase. *EMBO J.* **21**, 427–439 (2002).
72. Bollum, F. J. Chemically defined templates and initiators for deoxypolynucleotide synthesis. *Science* **144**, 560–560 (1964).
73. Boulé, J. B., Rougeon, F. & Papanicolaou, C. Terminal deoxynucleotidyl transferase indiscriminately incorporates ribonucleotides and deoxyribonucleotides. *J. Biol. Chem.* **276**, 31388–31393 (2001).
74. Lee, H. H., Church, G. M. & Kalhor, R. Enzymatic nucleic acid synthesis. Patent WO/2017/176541 (2017).
75. Ybert, T. & Garieff, S. Method for synthesizing nucleic acids, in particular long nucleic acids, use of said method and kit for implementing said method. US Patent 2021/0130863 (2021).
76. Ybert, T. & Garieff, S. Modified nucleotides for synthesis of nucleic acids, a kit containing such nucleotides and their use for the production of synthetic nucleic acid sequences or genes. US Patent 2020/0231619 (2020).
77. Creton, S. Efficient product cleavage in template-free enzymatic synthesis of polynucleotides. Patent WO/2020/165137 (2020).
78. Loftie-Eaton, W. et al. Novel variants of endonuclease V and uses thereof. Patent WO/2022/090057 (2022).
79. Schott, H. & Schrade, H. Single-step elongation of oligodeoxynucleotides using terminal deoxynucleotidyl transferase. *Eur. J. Biochem.* **143**, 613–620 (1984).
80. Motea, E. A. & Berdis, A. J. Terminal deoxynucleotidyl transferase: the story of a misguided DNA polymerase. *Biochim. Biophys. Acta* **1804**, 1151–1166 (2010).
81. Efcavitch, J. W. & Sylvester, J. E. Modified template-independent enzymes for polydeoxynucleotide synthesis. US Patent 2016/0108382 (2016).
- Invention exacting an effective reversible termination mechanism with NTPs modified with terminators to add a single, defined nucleotide per reaction step.**
82. Hiatt, A. C. & Rose, F. D. Enzyme catalyzed template-independent creation of phosphodiester bonds using protected nucleotides. US Patent 5,990,300 (1999).
83. Barthel, S., Palluk, S., Hillson, N. J., Keasling, J. D. & Arlow, D. H. Enhancing terminal deoxynucleotidyl transferase activity on substrates with 3' terminal structures for enzymatic de novo DNA synthesis. *Genes* **11**, e102 (2020).
84. Deibel, M. R. & Coleman, M. S. Biochemical properties of purified human terminal deoxynucleotidyltransferase. *J. Biol. Chem.* **255**, 4206–4212 (1980).
85. Chang, L. M. S. & Bollum, F. J. Multiple roles of divalent cation in the terminal deoxynucleotidyltransferase reaction. *J. Biol. Chem.* **265**, 17436–17440 (1990).
86. Chang, L. M. S. & Bollum, F. J. Deoxynucleotide-polymerizing enzymes of calf thymus gland. V. Homogeneous terminal deoxynucleotidyl transferase. *J. Biol. Chem.* **246**, 909–916 (1971).
87. Bhan, N. et al. Recording temporal signals with minutes resolution using enzymatic DNA synthesis. *J. Am. Chem. Soc.* **143**, 16630–16640 (2021).
88. Arzumanov, A. A., Victorova, L. S., Jasko, M. V., Yesipov, D. S. & Kravetsky, A. A. Terminal deoxynucleotidyl transferase catalyzes the reaction of DNA phosphorylation. *Nucleic Acids Res.* **28**, 1276–1281 (2000).
89. Kravetsky, A., Victorova, L. S., Arzumanov, A. A. & Jasko, M. V. Terminal deoxynucleotidyl transferase catalysis of DNA (oligodeoxynucleotide) phosphorylation. *Pharmacol. Ther.* **85**, 165–173 (2000).
90. Flamme, M. et al. Evaluation of 3'-phosphate as a transient protecting group for controlled enzymatic synthesis of DNA and XNA oligonucleotides. *Commun. Chem.* **5**, e68 (2022).

91. Collins, J., Singh, S. & Vanier, G. Microwave technology for solid phase peptide synthesis. *Chim. Oggi* **30**, 26–29 (2012).
92. Hari Das, R., Ahirwar, R., Kumar, S. & Nahar, P. Microwave-assisted rapid enzymatic synthesis of nucleic acids. *Nucleosides Nucleotides Nucleic Acids* **35**, 363–369 (2016).
93. Yoshimura, T., Sugiyama, J., Mineki, S. & Ohuchi, S. *Effect of Microwaves on DNA and Proteins* (Springer, 2017).
94. DNA Script. Enzymatic DNA synthesis: technical note (DNA Script, 2021).
95. Soskine, M., Champion, E. & Mchale, D. Stabilized N-terminally truncated terminal deoxynucleotidyl transferase variants and uses thereof. Patent WO/2022/063835 (2022).
96. Hutter, D. et al. Labeled nucleoside triphosphates with reversibly terminating aminoalkoxyl groups. *Nucleosides Nucleotides Nucleic Acids* **29**, 879–895 (2010).
97. Chen, M. C., Huang, J., Lazar, R. & Mclnroy, G. Azidomethyl ether deprotection method. US Patent 2018/0201968 (2018).
98. Hoff, K., Halpain, M., Garbagnati, G., Edwards, J. S. & Zhou, W. Enzymatic synthesis of designer DNA using cyclic reversible termination and a universal template. *ACS Synth. Biol.* **9**, 283–293 (2020).
99. Bentley, D. R. et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).
100. Guo, J. et al. Four-color DNA sequencing with 3'-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc. Natl Acad. Sci. USA* **105**, 9145–9150 (2008).
101. Stemple, D. L., Mankowska, S. A. & Harvey, S. A. Compositions and methods for template free enzymatic nucleic acid synthesis. Patent WO/2018/152323 (2018).
102. Mathews, A. S., Yang, H. & Montemagno, C. Photo-cleavable nucleotides for primer free enzyme mediated DNA synthesis. *Org. Biomol. Chem.* **14**, 8278–8288 (2016).
103. Metzker, M. L. et al. Termination of DNA synthesis by novel 3'-modified deoxyribonucleoside 5'-triphosphates. *Nucleic Acids Res.* **22**, 4259–4267 (1994).
104. Wu, W. et al. Termination of DNA synthesis by N6-alkylated, not 3'-O-alkylated, photocleavable 2'-deoxyadenosine triphosphates. *Nucleic Acids Res.* **35**, 6339–6349 (2007).
105. Bowers, J. et al. Virtual terminator nucleotides for next-generation DNA sequencing. *Nat. Methods* **6**, 593–595 (2009).
106. Takeshita, L., Yamada, Y., Masaki, Y. & Seio, K. Synthesis of deoxypseudouridine 5'-triphosphate bearing the photoremovable protecting group at the N1 position capable of enzymatic incorporation to DNA. *J. Org. Chem.* **85**, 1861–1870 (2020).
107. Flamme, M. et al. Towards the enzymatic synthesis of phosphorothioate containing LNA oligonucleotides. *Bioorg. Med. Chem. Lett.* **48**, e128242 (2021).
108. Wojciechowski, F. & Ybert, T. Method for preparing 3'-O-amino-2'-deoxyribonucleoside-5'-triphosphates. Patent WO/2021/198040 (2021).
109. Champion, E., Soskine, M., Ybert, T. & Delarue, M. Variants of terminal deoxynucleotidyl transferase and uses thereof. Patent WO/2019/135007 (2019).
110. **Underpinning technology by DNA Script for template-independent DNA synthesis using re-engineered Tdt.**
110. Ybert, T. Template-free enzymatic polynucleotide synthesis using dismutationless terminal deoxynucleotidyl transferase variants. Patent WO/2021/122539 (2021).
111. Champion, E., Soskine, M., Jaziri, F. & Mchale, D. Chimeric terminal deoxynucleotidyl transferases for template-free enzymatic synthesis of polynucleotides. Patent WO/2021/122539 (2021).
112. Soskine, M. & Champion, E. Terminal deoxynucleotidyl transferase variants and uses thereof. Patent WO/2021/213903 (2021).
113. Lu, X. et al. Enzymatic DNA synthesis by engineering terminal deoxynucleotidyl transferase. *ACS Catal.* **12**, 2988–2997 (2022).
114. Chua, J. P. S. et al. Evolving a thermostable terminal deoxynucleotidyl transferase. *ACS Synth. Biol.* **9**, 1725–1735 (2020).
115. Chen, M. C. & Mclnroy, G. R. Method of oligonucleotide synthesis. Patent WO/2020/178604 (2020).
116. Mclnroy, G. R., Ost, T. W. B. & Lovedale, D. Methods relating to de novo enzymatic nucleic acid synthesis. Patent WO/2022/034331 (2022).
117. Chen, M. C., Mclnroy, G. R., Fox, M. E. & Matuszewski, M. R. Nucleic acid polymer with amine-masked bases. Patent WO/2020/229831 (2020).
118. Chen, M. C., Lazar, R. A., Huang, J. & Mclnroy, G. R. A process for the preparation of nucleic acid by means of 3'-o-azidomethyl nucleotide triphosphate. Patent WO/2016/139477 (2016).
119. Ost, T. W. B., Mclnroy, G. R., Gaber, Z. B., Swerdlow, H. & Tognoloni, C. Methods of nucleic acid synthesis. Patent WO/2021/148809 (2021).
120. Champion, E. et al. High efficiency template-free enzymatic synthesis of polynucleotides. Patent WO/2021/094251 (2021).
121. Palluk, S. et al. De novo DNA synthesis using polymerase–nucleotide conjugates. *Nat. Biotechnol.* **36**, 645–650 (2018).
122. Chang, L. M. S., Bollum, F. J. & Gallo, R. C. Molecular biology of terminal transferase. *Crit. Rev. Biochem.* **21**, 27–52 (1986).
123. Knapp, D. C. et al. Fluoride-cleavable, fluorescently labelled reversible terminators: synthesis and use in primer extension. *Chem. Eur. J.* **17**, 2903–2915 (2011).
124. Chen, M. C. & Mclnroy, G. R. Nucleotide derivatives containing amine masked moieties and their use in a templated and non-templated enzymatic nucleic acid synthesis. Patent WO/2019/097233 (2019).
125. Chen, M. C., Lazar, R. A., Huang, J. & Mclnroy, G. R. Compositions and methods related to nucleic acid synthesis. US Patent 11236377 (2020).
126. Mclnroy, G. R., Cook, I. H., Chen, M. C. & Chen, S. Modified terminal deoxynucleotidyl transferase (Tdt) enzymes. Patent WO/2022/029427 (2022).
127. Chen, M. C., Huang, J. & Mclnroy, G. R. Use of terminal transferase enzyme in nucleic acid synthesis. Patent WO/2018/215803 (2018).
128. Bell, N. M., Mankowska, S. A., Harvey, S. A. & Stemple, D. L. gSynth: synthesis and assembly of whole plasmids (Camena Bioscience, 2020).
129. Efcavitch, W. & Siddiqi, S. Methods and apparatus for synthesizing nucleic acids. US Patent 8,808,989 (2014).
130. **Invention for template-free de novo DNA synthesis, nucleotide by nucleotide, by Molecular Assemblies.**
130. Aggarwal, S. Reusable initiators for synthesizing nucleic acids. Patent WO/2021/207158 (2021).
131. Efcavitch, J. W. & Tubbs, J. L. Nucleic acid synthesis using DNA polymerase theta. Patent WO/2018/175436 (2018).
132. Arlow, D. H. & Palluk, S. Nucleic acid synthesis and sequencing using tethered nucleoside triphosphates. US Patent 2019/0112627 (2019).
133. **Invention of exemplar reversible termination by ANSA Biotechnologies to sterically control prescribed polymerization using Tdt-tethered NTPs.**
133. Bell, N. M., Mankowska, S. A., Harvey, S. A., Stemple, D. L. & Fraser, A. gSynth: a highly accurate, enzymatic, de novo synthesis and gene assembly technology (Camena Bioscience, 2019).
134. Stemple, D. L., Fraser, A. G., Mankowska, S. & Bell, N. Compositions and methods for template-free geometric enzymatic nucleic acid synthesis. Patent WO/2020/150143 (2020).
135. **Exemplar methodology by Camena Bioscience for template-free nucleic acid synthesis using reversible terminating NTPs.**
135. Horgan, A., Sarvac, I., Niyomchon, S. & Godron, X. Increasing long-sequence yields in template-free enzymatic synthesis of polynucleotides. Patent WO/2021/018921 (2021).
136. Hutchison, C. A. III et al. Design and synthesis of a minimal bacterial genome. *Science* **351**, e6253 (2016).
137. Casini, A., Storch, M., Baldwin, G. S. & Ellis, T. Bricks and blueprints: methods and standards for DNA assembly. *Nat. Rev. Mol. Cell Biol.* **16**, 568–576 (2015).
138. Young, R., Haines, M., Storch, M. & Freemont, P. S. Combinatorial metabolic pathway assembly approaches and toolkits for modular assembly. *Metab. Eng.* **63**, 81–101 (2021).
139. Gibson, D. G. et al. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215–1220 (2008).
140. Gibson, D. G., Smith, H. O., Hutchison, C. A. III, Venter, J. C. & Merryman, C. Chemical synthesis of the mouse mitochondrial genome. *Nat. Methods* **7**, 901–903 (2010).
141. TerMaat, J. R., Pienaar, E., Whitney, S. E., Mamedov, T. G. & Subramanian, A. Gene synthesis by integrated polymerase chain assembly and PCR amplification using a high-speed thermocycler. *J. Microbiol. Methods* **79**, 295–300 (2009).
142. Roth, T. L., Milenkovic, L. & Scott, M. P. A rapid and simple method for DNA engineering using cyclized ligation assembly. *PLoS ONE* **9**, e107329 (2014).
143. Sayers, J. R. & Eckstein, F. Properties of overexpressed phage T5 D15 exonuclease. Similarities with *Escherichia coli* DNA polymerase I 5'-3' exonuclease. *J. Biol. Chem.* **265**, 18311–18317 (1990).
144. Housby, J. N., Thorbjarnardóttir, S. H., Jónsson, Z. O. & Southern, E. M. Optimised ligation of oligonucleotides by thermal ligases: comparison of *Thermus scotoductus* and *Rhodothermus marinus* DNA ligases to other thermophilic ligases. *Nucleic Acids Res.* **28**, e10 (2000).
145. Takahashi, M., Yamaguchi, E. & Uchida, T. Thermophilic DNA ligase. Purification and properties of the enzyme from *Thermus thermophilus* HB8. *J. Biol. Chem.* **259**, 10041–10047 (1984).
146. Saiki, R. K. et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491 (1988).
147. Jayaraman, K., Fingar, S. A., Shah, J. & Fyles, J. Polymerase chain reaction-mediated gene synthesis: synthesis of a gene coding for isozyme c of horseradish peroxidase. *Proc. Natl Acad. Sci. USA* **88**, 4084–4088 (1991).
148. Saiki, R. K. et al. Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anaemia. *Science* **230**, 1350–1354 (1985).
149. Hayes, M. J., Ferguson, A. J., Juncu, V. D. & Temple, S. Temperature control device. Patent WO/2018/104698 (2018).
150. Hayes, M. J., Sanches-Kuiper, R. M. & Bygrave, D. A. Error detection during hybridization of target double-stranded nucleic acid. Patent WO/2019/064006 (2019).
151. Crosby, S. R., Jennison, M. & Brennan, J. Thermally-cleavable protecting and linker groups. Patent WO/2018/189546 (2018).
152. Brennan, J., Bygrave, D., Aditya, S. & Sanches-Kuiper, R. Method for producing double stranded polynucleotides based on oligonucleotides with selected and different melting temperatures. Patent WO/2018/167475 (2018).
153. **Enabling methodology of oligonucleotide selectivity for highly parallelized DNA synthesis under thermal control developed by Evonetix.**
153. Vldar, H. P., Redondo, F. & Aparecido, R. A novel method for synthesis of polynucleotides using a diverse library of oligonucleotides. Patent WO/2019/073072 (2019).
154. **Invention of a highly selective methodology by Ribbon Biolabs for DNA synthesis from diverse oligonucleotide libraries culminating in the first synthesis of >10 kb DNA.**
154. Vldar, H. P., Redondo, F. & Aparecido, R. A novel method for synthesis of polynucleotides using a diverse library of oligonucleotides. US Patent 2020/0283756 (2020).

155. Rouillard, J. M. et al. Gene2Oligo: oligonucleotide design for in vitro gene synthesis. *Nucleic Acids Res.* **32**, W176–W180 (2004).
156. Banyai, W., Chen, S., Fernandez, A., Indermuhle, P. & Peck, B. J. De novo synthesized gene libraries. Patent WO/2015/021080 (2015).
157. Peck, B. J., Noe, M., Pitsch, S. & Weiss, P. A. Highly accurate de novo polynucleotide synthesis. Patent WO/2020/139871 (2020).
158. Horgan, A., Lachaize, H., Verado, D. & Godron, X. Massively parallel enzymatic synthesis of polynucleotides. Patent WO/2022/013094 (2022).
159. Gao, X. et al. Oligonucleotide synthesis using solution photogenerated acids. *J. Am. Chem. Soc.* **120**, 12698–12699 (1998).
160. Gao, X. et al. A flexible light-directed DNA chip synthesis gated by deprotection using solution photogenerated acids. *Nucleic Acids Res.* **29**, 4744–4750 (2001).
161. Fodor, S. P. A. et al. Light-directed, spatially addressable parallel chemical synthesis. *Science* **251**, 767–773 (1991).
162. Zhou, X. et al. Microfluidic PicoArray synthesis of oligodeoxynucleotides and simultaneous assembling of multiple DNA sequences. *Nucleic Acids Res.* **32**, 5409–5417 (2004).
163. Septak, M. Kinetic studies on depurination and detriylation of CPG-bound intermediates during oligonucleotide synthesis. *Nucleic Acids Res.* **24**, 3053–3058 (1996).
164. Warner, J. R., Reeder, P. J., Karimpour-Fard, A., Woodruff, L. B. A. & Gill, R. T. Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat. Biotechnol.* **28**, 856–862 (2010).
165. Tian, J. et al. Accurate multiplex gene synthesis from programmable DNA microchips. *Nature* **432**, 1050–1054 (2004).
166. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **3**, e3647 (2008).
167. Nugent, R., Chen, S., Kettleborough, R. & Raynard, N. Barcode-based nucleic acid sequence assembly. Patent WO/2020/257612 (2020).
168. Smith, H. O., Hutchison, C. A. III, Pfannkoch, C. & Venter, J. C. Generating a synthetic genome by whole genome assembly: ϕ X174 bacteriophage from synthetic oligonucleotides. *Proc. Natl Acad. Sci. USA* **100**, 15440–15445 (2003).
169. Schaudy, E., Lietard, J. & Somoza, M. M. Sequence preference and initiator promiscuity for de novo DNA synthesis by terminal deoxynucleotidyl transferase. *ACS Synth. Biol.* **10**, 1750–1760 (2021).
170. Chow, D. C., Lee, W. K., Zauscher, S. & Chilkoti, A. Enzymatic fabrication of DNA nanostructures: extension of a self-assembled oligonucleotide monolayer on gold arrays. *J. Am. Chem. Soc.* **127**, 14122–14123 (2005).
171. Lee, H. et al. Photon-directed multiplexed enzymatic DNA synthesis for molecular digital data storage. *Nat. Commun.* **11**, e5246 (2020).
172. Deshpande, S., Yang, Y., Chilkoti, A. & Zauscher, S. Enzymatic synthesis and modification of high molecular weight DNA using terminal deoxynucleotidyl transferase. *Method. Enzymol.* **627**, 163–188 (2019).
173. Scott, V. L. et al. Novel synthetic plasmid and Doggybone™ DNA vaccines induce neutralizing antibodies and provide protection from lethal influenza challenge in mice. *Hum. Vacc. Immunother.* **11**, 1972–1982 (2015).
174. Walters, A. A. et al. Comparative analysis of enzymatically produced novel linear DNA constructs with plasmids for use as DNA vaccines. *Gene Ther.* **21**, 645–652 (2014).
175. Banér, J., Nilsson, M., Mendel-Hartvig, M. & Landegren, U. Signal amplification of padlock probes by rolling circle replication. *Nucleic Acids Res.* **26**, 5073–5078 (1998).
176. Hill, V. Production of closed linear DNA. Patent WO/2010/086626 (2010).
Enzymatic methodology for manufacturing dbDNA as the first plasmid-free vector cassette encoding the sequence of interest invented by Touchlight Genetics.
177. Kendirgi, F. et al. Novel linear DNA vaccines induce protective immune responses against lethal infection with influenza virus type A/H5N1. *Hum. Vaccines* **4**, 410–419 (2008).
178. Karbowiczek, K. et al. Doggybone™ DNA: an advanced platform for AAV production. *Cell Gene Ther. Insights* **3**, 731–738 (2017).
179. Huang, W. M., Joss, L., Hsieh, T. & Casjens, S. Protelomerase uses a topoisomerase IB/Y-recombinase type mechanism to generate DNA hairpin ends. *J. Mol. Biol.* **337**, 77–92 (2004).
180. Smith, H. O. & Wilcox, K. W. A restriction enzyme from hemophilus influenzae. *J. Mol. Biol.* **51**, 389–391 (1970).
181. Porter, N., Rothwell, P. J., Gehrig-Giannini, S., Babula, A. & Adie, T. A. J. Synthesis of DNA with improved yield. Patent WO/2020/035698 (2020).
182. Porter, N., Rothwell, P. & Extnance, J. Synthesis of DNA. Patent WO/2016/034849 (2016).
183. Porter, N. Production of closed linear DNA using a palindromic sequence. Patent WO/2012/017210 (2012).
184. Extnance, A. How DNA could store all the world's data. *Nature* **537**, 22–24 (2016).
185. Hillson, N. et al. Building a global alliance of biofoundries. *Nat. Commun.* **10**, e2040 (2019).
186. Church, G., Gao, Y. & Kosuri, S. Next-generation digital information storage in DNA. *Science* **337**, e1226355 (2012).
The first book written using DNA demonstrating DNA as a digital medium for information storage.
187. Lee, H. H., Kalthor, R., Goela, N., Bolot, J. & Church, G. M. Terminator-free template-independent enzymatic DNA synthesis for digital information storage. *Nat. Commun.* **10**, e2383 (2019).
188. Ceze, L., Nivala, J. & Strauss, K. Molecular digital data storage using DNA. *Nat. Rev. Genet.* **20**, 456–466 (2019).
189. Vitak, S. Technology alliance boosts efforts to store data in DNA. *Nature* <https://doi.org/10.1038/d41586-021-00534-w> (2021).
190. Xu, C., Zhao, C., Ma, B. & Liu, H. Uncertainties in synthetic DNA-based data storage. *Nucleic Acids Res.* **49**, 5451–5469 (2021).
191. Song, L. F., Deng, Z. H., Gong, Z. Y., Li, L. L. & Li, B. Z. Large-scale de novo oligonucleotide synthesis for whole-genome synthesis and data storage: challenges and opportunities. *Front. Bioeng. Biotechnol.* **9**, e689797 (2021).
192. Meiser, L. C. et al. Reading and writing digital data in DNA. *Nat. Protoc.* **15**, 86–101 (2020).
193. Hoshika, S. et al. Hachimoji DNA and RNA: a genetic system with eight building blocks. *Science* **363**, 884–887 (2020).
194. Berg, P. Asilomar 1975: DNA modification secured. *Nature* **455**, 290–291 (2008).
195. Fatehi, L. & Hall, R. F. Synthetic biology in the FDA realm: toward productive oversight assessment. *Food Drug Law J.* **70**, 339–369 (2015).
196. International Gene Synthesis Consortium. About IGSC. *International Gene Synthesis Consortium* <https://genesynthesisconsortium.org>.
197. National Academy of Medicine, National Academy of Sciences & the Royal Society. Heritable human genome editing (National Academies, 2020).
198. Burmeister, A. R. Horizontal gene transfer. *Evol. Med. Public Health* **1**, 193–194 (2015).
199. Kalman, L. V. et al. Development and characterization of reference materials for genetic testing: focus on public partnerships. *Ann. Lab. Med.* **36**, 513–520 (2016).
200. Spencer, M. The stereochemistry of deoxyribonucleic acid. II. Hydrogen-bonded pairs of bases. *Acta Crystallogr.* **12**, 66–71 (1959).

Acknowledgements

The authors acknowledge the UK Defence Science and Technology Laboratory for initiating and supporting this review. The authors thank P. Oyston for the critical assessment of the work.

Author contributions

All authors researched data for the article. All authors contributed substantially to discussion of the content. A.H. and M.G.R. wrote the article. All authors reviewed and edited the manuscript before submission.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to Maxim G. Ryadnov.

Peer review information *Nature Reviews Chemistry* thanks Marcel Hollenstein, De-Ming Kong, Vincent Noireaux and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Crown 2023, corrected publication 2023