Advancing high-throughput gene synthesis technology

Jingdong Tian*, Kuosheng Ma and Ishtiaq Saaem
Department of Biomedical Engineering and the Institute for Genome Sciences and Policy, Duke University, 136 Hudson Hall, Box 90281, Durham, NC 27708-0281, USA. E-mail: jtian@duke.edu

Received 10th December 2008, Accepted 11th February 2009

First published on the web 6th April 2009

The emerging field of synthetic biology is generating insatiable demands for synthetic genes, which far exceed existing gene synthesis capabilities. This review discusses the current methods of chemical DNA synthesis and gene assembly, as well as the latest engineering tools, technologies and trends which could potentially lead to breakthroughs in the development of accurate, low-cost and high-throughput gene synthesis technology. The capability of generating unlimited supplies of DNA molecules of any sequence or size will transform biomedical research in the near future.

Jingdong Tian received his PhD from Stony Brook University and an LSRF postdoctoral fellowship at Harvard Medical School. Currently, he is an assistant professor in the Department of Biomedical Engineering and the Institute for Genome Sciences and Policy at Duke University. His primary research interests include synthetic biology, bionanotechnology and genome engineering. He is a Beckman Young Investigator and a recipient of a Hartwell Individual Biomedical Research Award.

Jingdong Tian
Kuo-Sheng Ma received his PhD in electrical engineering from the University of California, Irvine in 2005. He is currently a research associate in the laboratory of Dr Jingdong Tian in the Department of Biomedical Engineering at Duke University. His research is focused on MEMS technology and its application in high-throughput gene synthesis.

Ishtiaq Saaem received his BE and ME in computer engineering from Stevens Institute of Technology in 2006. He is currently a PhD candidate in biomedical engineering and a member of Dr Jingdong Tian’s laboratory at Duke University. His research interests involve the development of inkjet techniques for high-throughput gene synthesis and the fabrication of stimuli-responsive polymeric structures.

**Introduction**

The rapid growth of synthetic biology has been recognized by its broad and promising applications in the design and engineering of various kinds of biological systems at different levels, such as biomolecular engineering, gene circuits and networks design, metabolic engineering, and whole chromosome or genome construction.\(^1\) The implications are far-reaching. To achieve these goals, the ability to synthesize *de novo* DNA constructs of any size or sequence rapidly, accurately and economically is crucial. Today, the process of DNA oligonucleotide synthesis has been automated and methods for synthesizing longer DNA molecules, which mainly rely on the assembly of pre-synthesized oligonucleotides, are being improved. However, the overall development of gene and genome synthesis technology has far lagged behind that of gene and genome sequencing. This review outlines the standard methods, chemistry and instrumentation of DNA synthesis, and discusses technological developments aimed at increasing the quality and throughput of gene synthesis, while minimizing the cost and reagent consumption. The development of gene synthesis technology is at a critical turning point and its success will spur an exponential growth of biomedical research.

**Chemical oligodeoxynucleotide synthesis**
Standard chemical DNA synthesis is a cyclical process that elongates a chain of nucleotides from the 3′-end to the 5′-end. The phosphoramidite four-step process, developed in the early 1980s, is the method of choice currently used by most commercial DNA synthesizers.\textsuperscript{15–17} This process couples an acid-activated deoxynucleoside phosphoramidite to a deoxynucleoside on a solid support.

In the first synthesis cycle, the nucleotide chain grows from an initial protected nucleoside tethered to a solid support via its terminal 3′ hydroxyl. The commonly used supports include controlled pore glass (CPG) or polystyrene (PS) beads. Reagents are pumped onto and through the support to induce a stepwise addition of nucleotide monomers to the growing oligonucleotide chain. The addition of each nucleotide monomer to the elongating oligonucleotide chain is carried out in four steps (Fig. 1): (1) deprotection: a weak acid is used to remove the dimethoxytrityl (DMT) ether group from the 5′-end of the growing oligonucleotide chain and generate a reactive 5′-OH group; (2) coupling: the 5′-OH group generated from the deprotection step reacts with an activated monomer created by simultaneously adding the desired phosphoramidite and an appropriate activator (i.e., tetrazole); (3) capping: uncoupled 5′-OH groups are blocked by an acylating capping reagent to minimize deletion products; and (4) oxidation: the unstable phosphite triester internucleotide linkages are oxidized to a more stable phosphotriester. This four-step cycle is repeated for the addition of each nucleotide in the sequence. After the synthesis is complete, the oligonucleotide chain is cleaved from the solid support with a base, i.e., concentrated ammonium hydroxide. The remaining protection groups on the oligonucleotide chain are removed during the same process.

Recently, an alternative two-step ODN synthesis method has been developed.\textsuperscript{18} The approach utilizes a peroxy anion as the nucleophile in each synthesis cycle to simultaneously remove a 5′-carbonate and oxidize the internucleotide phosphite triester. The cyclical removal of the 5′-protecting group with a peroxy anion under mildly basic conditions is considered essentially irreversible and quantitative. This procedure is designed to completely eliminate depurination and reduce mutation frequencies in cloned, synthetic DNA. The two-step method also simplifies the synthesis procedure by eliminating several reagents. This should allow for simpler and potentially more robust automation, and result in cost savings for the large scale synthesis of oligonucleotides.

Solid-phase synthesis makes automation possible because it eliminates the need to purify synthetic intermediates or unreacted reagents. These reagents are simply rinsed off the column at the end of each
Gene synthesis

Chemical DNA synthesis is typically used for synthesizing oligonucleotides shorter than 120–150 bases. Although longer oligomers of up to 300–600 bases can be directly synthesized, the yield becomes extremely low. For the synthesis of gene-sized DNA sequences, a number of enzymatic methods have been developed over the past decades. Among them, two general methods are most commonly used today for the synthesis of individual genes using pre-synthesized oligonucleotides as building blocks (Fig. 2).

Fig. 2 Enzymatic methods for gene synthesis. A: Ligation-based assembly usually involves two steps, ligation and PCR amplification. B: Two-step PCR-driven assembly, in which the gene-end amplification primer pair (red) is used after the first PCR assembly reaction. C: One-step gene assembly reaction with gene-end amplification primer pair mixed with the gene-construction primers (blue) in a single-step PCR assembly and amplification reaction.
Ligation-based assembly

Joining oligonucleotides using DNA ligase to form longer genes have been used in early examples of gene synthesis and in certain commercial solid-phase gene synthesis platforms. With the discovery of thermostable DNA ligases and the development of the ligase chain reaction (LCR), thermo-ligase or LCR-based gene assembly methods have become very convenient. One advantage of using thermo-ligase is the reduced secondary structure formation at elevated ligation temperatures (~50–65 °C).

In this approach (Fig. 2A), carefully designed overlapping oligos that completely cover both strands of the gene sequence are chemically synthesized and phosphorylated at the 5'-ends. The oligos are mixed together in a buffer with thermo-ligase and heat-denatured. The mixture is then slowly cooled down to a temperature suitable for proper annealing and ligation. The denaturation and annealing/ligation steps can be repeated for a number of cycles. In order to produce enough quantities of the full-length gene product, the ligation reaction is usually followed by an amplification step using the polymerase chain reaction (PCR) with a pair of specific gene-end primers to amplify the full-length gene sequence.

PCR-driven assembly

Without ligation, a procedure similar to PCR can be used to assemble overlapping oligos into full-length gene constructs, either in two steps or a single step. These methods have been given various names, such as “recursive PCR”, “assembly PCR”, or “polymerase cycling assembly” (PCA).

In the two-step procedure (Fig. 2B), overlapping oligonucleotides that together span the whole sequence are mixed together in equal, low concentrations with a PCR mixture, including buffer, dNTPs, and a polymerase. The thermal cycling steps are carried out in a similar way to a normal PCR. During the first PCR assembly reaction, overlapping oligos will anneal and extend using each other as a template to form longer and longer DNA fragments, and eventually reach full length. Then a PCR amplification reaction will use a pair of gene-end primers and a small fraction of the assembly reaction mixture as template to amplify the full-length construct.

In the one-step procedure (Fig. 2C), a pair of gene-end primers is included from the beginning and at a higher concentration than the rest of the gene-construction oligos. Extra cycles may be needed to assemble and amplify the full-length construct. This strategy has been used in a multiplexed fashion to simultaneously assemble multiple genes from a single reaction.

The performance of the ligation-based assembly is almost equal to the PCR-based approach for most sequences. In PCR-driven assembly, gaps are allowed between adjacent oligos that belong to the same sense or anti-sense strand. This gives PCR-driven assembly a slight advantage in terms of the total amount of chemical DNA synthesis required over the ligation-based assembly, where no gap is allowed. The speed and convenience of the single-step, PCR-driven assembly is another attractive feature. However, not all sequences can be easily assembled by PCR. For some difficult constructs involving repetitive sequences or excessive DNA secondary structures, thermo-ligation may be the best option.

Error removal

Both chemical oligonucleotide synthesis and enzymatic gene assembly can introduce errors to the final synthetic gene product. Various error removal strategies can be employed to eliminate errors generated during various stages of the gene synthesis process (Fig. 3).
Fig. 3 Error removal strategies for synthetic genes. A: Error removal using mismatch binding proteins. Mismatch binding proteins selectively bind to error sequences with a mismatch (blue band) and are separated from correct sequences (brown band) by either gel-shift assay (left) or affinity column (right). B: Error removal using mismatch cleaving enzymes. Mismatch cleaving enzymes cleave sequences with a mismatch. Cleaved fragments are removed either by size exclusion (left) or repaired by DNA exonuclease activity and extended into error-free full-length sequences (right).

Error removal from synthetic oligonucleotides

Because chemical reactions are rarely 100% efficient, the coupling efficiency for each monomer is typically 98.5–99.5% during chemical nucleic acid synthesis.40 Deletions and insertions are the most common types of errors in oligonucleotide chemical synthesis. Deletions happen largely as a result of incomplete capping or deprotection and can be as frequent as 0.5% per position. Insertions, mostly due to DMT cleavage by tetrazole, can reach approximately 0.4% per position.41 As a consequence, for an oligonucleotide containing 100 bases, typically less than 30–40% of the sequences are correct.

Besides perfecting the DNA synthesis chemistry to improve the oligonucleotide quality, a relatively
An effective method to remove deletions or insertions is by size exclusion using high performance liquid chromatography (HPLC) or polyacrylamide gel electrophoresis (PAGE) purification. With these methods, approximately 90% of the impurities can be removed. However, these methods are not effective against other types of mutations, which do not involve size change.

Error removal from synthetic genes

The errors that remain in the synthetic oligonucleotides will be carried over and accumulated in longer synthetic DNA constructs. Gene assembly reactions are also error-prone and can introduce additional errors. Identifying error-free sequences by cloning and sequencing is time consuming and costly. In some cases, expression or functional screens can be implemented to eliminate errors that will cause frame shift or function loss. However, such targets are limited to protein-coding sequences and functional DNA elements.

A more general approach is to use DNA mismatch recognition proteins. There are two categories depending on the mode of action: (1) mismatch binding proteins, such as MutS, and (2) mismatch cleaving enzymes, such as the MutHLS complex or mismatch-specific endonucleases. The mismatch binding proteins will selectively bind to mismatches generated by hybridization between correct and incorrect sequences. The resulting protein–DNA complexes containing incorrect sequences can then be removed from the sequence pool by gel-shift assay or affinity columns (Fig. 3A). The mismatch cleaving enzymes will bind and cleave at the mismatch sites generated by hybridization between correct and incorrect sequences (Fig. 3B). The cleaved sequences can either be removed by size selection, degradation, or repaired by trimming off the mismatched base pairs with additional enzymes, such as DNA polymerase or exonuclease. The repaired fragments can then be extended using overlapping fragments as templates.

Oligonucleotide synthesis from DNA microarrays

Although synthetic oligonucleotides and genes have become commodities, the costs are still very high and the synthesis throughputs are limited. In searching for an abundant source for inexpensive oligonucleotides for gene construction, attention has recently been focused on DNA microarrays. Different from oligonucleotide synthesis on columns, where reactions occur in large separate compartments or columns, DNA microarray synthesis technologies utilize the surface of a silicon chip or a glass slide, with oligonucleotide growth confined to specific spots or volumes. This is achieved through a variety of different mechanisms, including photolithography with physical or digital masks, inkjet printing, electrochemical array and microfluidics. These mechanisms control whether or not a phosphoramidite monomer will be coupled to the growing oligonucleotide chains on a particular spot during each synthesis cycle.

Photolithography

Photolithography using physical masks and photolabile nucleoside monomers is an early system used to synthesize oligonucleotide microarrays. The masks are employed to generate light patterns that dictate which areas on the array need to be activated for chemical coupling. A stack of masks needs to be prefabricated according to the oligonucleotide sequences to be synthesized on a chip. Light exposure in specified areas removes photolabile blocking groups on the growing chains. After deblinking, a selected phosphoramidite monomer is added onto the entire surface, but coupling reactions only occur in areas under light exposure. The cycles repeat until the whole synthesis is completed. Affymetrix Inc. applied this technology to the large-scale fabrication of high-density GeneChip probe arrays for gene expression and nucleic acid sequence analysis.

The cost of using large numbers of pre-fabricated photomasks is high, probably only justifiable for large-scale, high-volume gene chip fabrication. A relatively low-cost and flexible alternative is digital...
photolithography using Texas Instruments’ Digital Micromirror Device (DMD™) based on the Digital Light Processing (DLP™) technology (Fig. 4). The DMD™ is a reflective display device consisting of an electromechanically controlled array of micromirrors. A major advantage of using DMD™ over physical masks is that it is easily programmable and very flexible. High resolution, precisely controllable light patterns can be generated in an automated manner.

![Fig. 4 DNA array synthesis using digital photolithography. (a) The ends of the linkers on the microarray slides are protected with photo- or acid-labile protection groups. (b) Digitally controlled light patterns remove protection groups from specified locations either directly or via photogenerated acids (PGA). (c) Phosphoramidite monomers (N-) are added and the cycles repeat to synthesize different oligonucleotide sequences on chip surface. (d) Digital micromirror array from Texas Instruments.](image)

Different photochemistry can be used with digital photolithography for DNA microarray synthesis. (Fig. 4(a) and (b)) Some designs directly use a photolabile protecting group (PLPG), such as (R,S)-1-(3,4-(methylenedioxy)-6-nitrophenyl)ethyl chloroformate (MeNPOC) or 2-(2-nitrophenyl)propanoyl (NPPOC) to block functional groups on the linker or the phosphoramidite monomers. Others use photogenerated acid (PGA) in solution to perform the deblocking step in conventional nucleotide phosphoramidite chemistry. The PGA solution at mM concentrations can effectively remove the DMT group to free the 5'-OH group of the nucleosides or nucleotides.

**Inkjet printing**

Inkjet printer heads, controlled by high quality motion controllers, can deliver small droplets of reagents to specific spots on a chemically modified slide surface, where they react to synthesize DNA (Fig. 5). A piezoelectric inkjet head consists of a small reservoir with an inlet port and a nozzle at the other end. One
wall of the reservoir consists of a thin diaphragm with an attached piezoelectric crystal. When a voltage is applied to the crystal, it contracts laterally thus deflecting the diaphragm and ejecting a small drop of fluid from the nozzle. The reservoir then refills via capillary action through the inlet.

Fig. 5 DNA array synthesis with piezoelectric inkjet technology. The microarray synthesis platform consists of a controller PC, DAQ-electronics, and a servo controller, which together control the printing from the print head onto the slides. The inset shows the working mechanism of the piezoelectric inkjet head.

In 1996, the first proof-of-principle oligonucleotide synthesis using the inkjet mechanism was demonstrated with microfabricated piezoelectric inkjet pumps. The inkjet head contained six fluid channels, with four delivering phosphoramidite precursors, one for an activator, and another one for an optional linker or modified base. Commercial piezoelectric inkjet heads were used in later inkjet microarray printing platforms.

To minimize the evaporation of picolitre-sized droplets during printing, the selection of a suitable solvent is critical. Acentonitrile, the solvent of choice in standard DNA synthesis, is too volatile to work well during inkjet synthesis. This high volatility results in frequent misplaced droplets, crystals of reagent sitting dry on the slide, and a very high rate of clogging the piezoelectric nozzles. Therefore, propylene carbonate (PC) and another solvent composed of a 1:1 mixture of 2-methyl glutaronitrile (MGN) and 3-methoxypropionitrile (3MP) were selected because they showed low volatility, higher boiling point and higher vapour pressure than acetonitrile. They were also compatible with phosphoramidite synthesis and not harmful to the inkjet head. The droplets were confined to specific positions on the slide via pre-patterned surface tension.

Since then, the inkjet printing method for microarray synthesis has mainly been developed and commercialized by Agilent Technologies, and used for gene-expression profiling applications. Currently, Agilent manufactures both custom and catalogue oligonucleotide microarrays on the same 1 × 3 in. microscope slide format that can contain up to 244 000 features.

Electrochemical array
Instead of using photo-chemical methods to generate the acid for the deprotection step, localized electrochemical reactions can also be used for DNA synthesis. Acid is produced only at specified sites by electrochemical oxidation using an array of individually addressable microelectrodes (Fig. 6). The electrodes can be made by thin-film photolithography of iridium metal. The electrolyte used in one design is hydroquinone and benzoquinone with tetrabutylammonium hexafluorophosphate in anhydrous acetonitrile. When a current is applied to the microelectrodes, the electrolyte is oxidized at the anodes and therefore releases acid, which diffuses to the substrate where the oligonucleotides are synthesized. The acid is confined in the region by the adjacent cathodes, which consumes acid by reduction (Fig. 6(c)). Alternatively, the synthesis can be performed on a porous polymer layer, which slows down the diffusion of the acid generated from the electrodes (Fig. 6(b)). The synthesis of short oligonucleotide sequences has been demonstrated using electrochemical arrays.

**Fig. 6** DNA synthesis on an electrochemical array. (a) Diagram of a microelectrode array on top of a DNA synthesis slide with an acid-labile protecting group. Two designs are shown here in order to generate localized acid; in (b), the acid generated from the anodes is localized by using a polymer material to block its diffusion, while in (c), the diffused acid is consumed by the adjacent cathodes. (d) Phosphoramidite monomers (N-) are added and
the cycles repeat to synthesize different oligonucleotide sequences on the chip surface.

**Microfluidics**

In addition to the above platform technologies, a number of other concepts have been developed using microfluidic devices for microscale parallel DNA synthesis. Key technical challenges involved in microfluidic DNA synthesis include individual addressability of each synthesis unit, chemical resistance of the microfluidic materials, as well as robustness and reliability of the whole synthesis process and device. A few of these concepts and prototypes are discussed below.

**Microfluidic oligonucleotide synthesizers**

Using elastomeric materials for making microfluidic devices is popular, due to their versatility, ease of fabrication, and cost-saving benefits over silicon or glass. However, because of the variety of chemicals and organic solvents used in standard chemical DNA synthesis, the materials used for fabricating microfluidic DNA synthesis devices need to have excellent chemical resistance. In an attempt to synthesize DNA using the popular poly(dimethylsiloxane) (PDMS), which swells in DNA synthesis reagents, Moorcroft *et al.* modified the conventional DNA synthesis protocol by substituting solvents for the oxidation and the deprotection steps. In this work, a simple method to functionalize the PDMS surface for linker immobilization was developed. The PDMS microchannel surface was initially oxidized through a UV–ozone system to increase silanol (SiOH) density. Vapour-phase silanization of the oxidized PDMS substrate was performed using 3-glycidoxypropyltrimethoxysilane (GPTMS) in a 175 °C furnace under low pressure (10 mBar). PEG was used as a spacer in the subsequent treatment. Using this PEG–silane–PDMS chip, 21-mer oligonucleotides were successfully synthesized.

A number of strategies have been developed in order to avoid the material swelling issue involved in microfluidic oligonucleotide synthesis. The microvalve array, which will be discussed in the next section, used chemical resistant parylene to coat PDMS, to form a parylene–PDMS hybrid film. Alternatively, chemical-resistant polymer materials, such as perfluoropolyether (PFPE) could be used. The Quake group fabricated a microfluidic oligonucleotide synthesis device using PFPE, in which the synthesis chamber was filled with porous silica beads as in the conventional phosphoramidite DNA synthesis. One microfluidic column could yield 60 pmol of 20-mers oligonucleotides while consuming 60-fold less chemicals as compared with conventional column synthesis.

**Microvalve array**

Microvalves are used in microfluidic systems to control the flow of fluidics, which is analogous to the use of transistors in microelectronics circuits. A microfluidic integration of microvalves can therefore conceivably be used to control the flow of reagents for parallel DNA oligonucleotide synthesis. In an initial demonstration of this concept, Hua *et al.* developed a lab-on-a-chip microreactor controlled by a chemical-resistant microvalve array and performed addressable oligonucleotide synthesis. In this work, a silicon–PDMS hybrid pneumatic microvalve array was fabricated with a parylene–PDMS double layer valve membrane. Trenches in PDMS layers served as air lines, while all microfluidic channels, microreactors and microvalve seats were located on the silicon layer. The membrane was sandwiched between the top PDMS and the bottom silicon substrate. To address specific reactors in the array, air was first pumped into air channels on top of the other valve seats. Thus, when fluid filled the bottom channels under those unpressurized air channels, the valves were opened by hydraulic pressure on the flexible parylene membrane. For any valve seat under a pressurized air channel, the valve membrane would be pushed down and the valve would be closed. Using this strategy, it was calculated that 16 air channels were enough to control (or address) 12 870 reactors, which was more efficient than conventional valves based
on binary tree multiplexing. A 30-mer oligonucleotide synthesis was performed on a prototype device, which achieved a step-wise synthesis yield of ~99.5%.

LED-controlled capillary synthesis

Instead of using expensive UV lasers or lamps and a complex optical control system, which is normally required for digital photolithography, Blair et al. explored the use of ultraviolet light emitting diodes (UV-LEDs) as an alternative light source for directing oligonucleotide synthesis inside glass capillaries. The UV-LED used in this study was made of a gallium nitride (GaN) chip encased in a plastic package with a 10° opening, generating a beam with a wavelength of 365 nm. A string of these UV-LEDs were placed on top along the length of a glass capillary, with the surface of its inside wall functionalized for DNA synthesis. Oligonucleotides were grown in capillaries using standard photolabile 2-nitrophenyl propoxycarbonyl (NPPOC) chemistry. One advantage of using LEDs is the spectral purity of the light source. The spectrum of the UV-LED illumination in this study had a half-width of 10 nm and contained no emission below 360 nm, therefore avoiding damage to the DNA from short wavelength radiation. The authors reported the successful synthesis of multiple oligonucleotides up to 70 bases and use of the amplified oligos for gene assembly.

Multiplex gene synthesis from DNA microchips

To reduce the cost and increase the throughput for gene synthesis, successful attempts have been made to use oligonucleotides synthesized from DNA microchips as building blocks for gene assembly. Here, DNA microchips were mainly used as an inexpensive source for generating large numbers of different oligonucleotide sequences. In these attempts, two different DNA microarray fabrication platforms were used, all relying on digital photolithography, but using different photochemistries for the deprotection step in DNA synthesis. The strategy of using chip synthesized oligonucleotide pools for gene synthesis is illustrated in Fig. 7.

![Fig. 7 Gene assembly from a DNA microchip.](image)

Gene-construction oligos flanked by universal primer sequences are synthesized in situ on the microchip surface. After cleavage, the oligos are amplified by PCR and the PCR primer sequences are removed by digestion with type II restriction enzymes. Clean oligos are used for the subsequent gene construction steps.

In order to harvest oligos made on DNA microchips, cleavable linkers need to be used in DNA microarray synthesis to anchor the oligos to the surface. After the synthesis is complete, treatment with ammonium hydroxide releases the oligos from the chip. The oligos are then collected and purified as a pool, and used for downstream processing. Current DNA microarrays can synthesize $10^3$–$10^6$ different
oligo sequences, but in very low yields, e.g. $\sim 10^6$ molecules for each sequence. For a typical microlitre-scale gene assembly reaction, an oligo pre-amplification step is usually necessary in order to reach optimal oligo concentrations for the reaction. In this case, sequence features necessary for post-synthesis enzymatic amplification need to be designed into the oligo sequences, such as short universal PCR primers flanking the sequences, and these primer sequences need to be removed after amplification. One way to do this is by using type IIs restriction enzyme digestion.\textsuperscript{10}

As a preliminary measure to remove errors from chip-synthesized and PCR-amplified oligos, a chip hybridization-based method was designed.\textsuperscript{10} In this method, two pools of error-correction oligos were synthesized from two DNA chips. Each pool consists of short oligos complementary to approximately one half of the “gene-construction” oligos released from the first chip. After hybridization and appropriate washes, mismatch sequences were selectively reduced from the pool and the correct sequences were preferentially enriched. A multiplex gene assembly reaction could be used to assemble multiple genes from the same pool of oligos. Bioinformatic measures may be taken during the sequence design and optimization stage in order to reduce cross hybridization between similar sequences. These gene fragments can be further assembled into longer and longer sequences, either \textit{in vitro} or \textit{in vivo}.

A typical DNA chip with $10^3$–$10^6$ different oligos is capable of constructing mega-bases of DNA sequences, equivalent to or surpassing the lengths of some microbial genomes. Special measures need to be taken to fully utilize this capacity. In addition to multiplexing, a combination of spatial separation, selective releasing, selective amplification, microfluidics or bioinformatics can be explored. It has been demonstrated that using microfluidics can reduce the reaction volume and potentially eliminate the need for oligo amplification before gene assembly.\textsuperscript{70}

**Conclusions and perspectives**

Gene synthesis, combined with recombinant DNA techniques, has been used in a wide range of applications, including biomolecular engineering, DNA nanotechnology and computing, gene circuits and networks construction, metabolic engineering, and genome synthesis.\textsuperscript{1–4,71} Due to the high cost and low throughput nature of conventional gene synthesis, many of the applications only involved oligonucleotide synthesis to make primers for PCR or mutagenesis. Even in the engineering of complex gene circuits or metabolic pathways, \textit{de novo} synthesis accounted for a relatively small fraction of the total sequences constructed. In those cases, most of the protein-coding genes were PCR copied from natural sources.

However, the high costs of conventional gene synthesis have not deterred research groups from breaking the length records of \textit{de novo} synthesized DNA sequences (Fig. 8), which became more frequent in recent years as the sequence information of hundreds of genomes became available. In 2002, the chemical synthesis of a functional poliovirus genome was achieved ($\sim 7500$ bp).\textsuperscript{13} In 2003, Smith, \textit{et al.} demonstrated the synthesis of a functional bacteriophage genome ($5386$ bp) in two weeks\textsuperscript{12} as compared to the months which would have normally been required using previous technologies. In 2004, the syntheses of a 14-kb ribosomal gene cluster\textsuperscript{10} and a 32-kb polyketide synthase gene cluster\textsuperscript{2} were reported, and in 2008, the synthesis and cloning of the \textit{Mycoplasma genitalium} genome ($582970$ bp) was accomplished.\textsuperscript{11} Being one of the smallest non-viral genomes, its synthesis paved the way towards a top-down approach of defining a minimal genome and a deeper understanding of the meaning of life. The paper noted that this remarkable feat was achieved by using $\sim 10^4$ synthetic oligonucleotides, each $\sim 50$ bases in length, which were individually synthesized and then assembled into $5$–$7$ kb gene cassettes by commercial providers.\textsuperscript{11} If the DNA microchip technology were used, it would take less than a single chip to make all the oligos for the \textit{Mycoplasma} genome.
With further drops in cost and increases in throughput, automation and platform integration, the time will come when de novo DNA writing will eventually become a routine and standard method for molecular biology and bioengineering. Combining high-throughput gene synthesis with biomolecular and systems design and evolution would make a powerful engine for creating new biological machines and phenotypes. This capability could potentially transform biomedical research in the near future. Entwined with it are legal, ethical, and societal ramifications of gene synthesis. Though it is out of the scope of this review, it is of utmost importance that researchers in the arena of gene and genome synthesis are aware of these developing issues.

Acknowledgements

The authors would like to thank members of the Tian laboratory for critical readings of the manuscript. J.T. is a Beckman Young Investigator and a recipient of a Hartwell Foundation Individual Biomedical Research Award.

References


---

**Footnote**

† This article is part of a Molecular BioSystems themed issue on Synthetic Biology

---

This journal is © The Royal Society of Chemistry 2009