

# Modern approaches to artificial gene synthesis: aspects of oligonucleotide synthesis, enzymatic assembly, sequence verification and error correction

G.Y. Shevelev<sup>1, 2</sup>✉, D.V. Pyshnyi<sup>1, 2</sup>

<sup>1</sup> Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia

<sup>2</sup> Novosibirsk State University, Novosibirsk, Russia

Synthetic biology is a rapidly developing field aimed at engineering of biological systems with predictable properties. Synthetic biology accumulates the achievements of modern biological sciences, programming and computational modeling as well as engineering technologies for creation of biological objects with user-defined properties. Evolution of synthetic biology has been marked by a number of technological developments in each of the mentioned fields. Thus, significant reduction in cost of DNA sequencing has provided an easy access to large amounts of data on the genetic sequences of various organisms, and decreased the price of the DNA sequence synthesis, which, analogous to Moore's law, resulted in an opportunity to create a lot of potential genes without the time – consuming and labor – intensive traditional methods of molecular biology. Development of system biology has allowed forming a deeper understanding of the functions and relationship of natural biological models, as well as of the computational models describing processes at the cell and system levels. Combination of these factors has created an opportunity for conscious changes of natural biological systems. In this review the modern approaches to oligonucleotide gene assembly synthesis are discussed, including such aspects as protocols for gene assembly, sequence verification, error correction and further applications of synthesized genes.

**Key words:** oligonucleotide synthesis; artificial gene synthesis; enzymatic gene assembly; polymerase cycling assembly (PCA); ligase cycling assembly (LCA); error correction; sequence verification.

## Современные подходы к синтезу генов: аспекты синтеза олигонуклеотидов, ферментативной сборки, проверки последовательностей и коррекции ошибок

Г.Ю. Шевелев<sup>1, 2</sup>✉, Д.В. Пышный<sup>1, 2</sup>

<sup>1</sup> Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия

<sup>2</sup> Новосибирский государственный университет, Новосибирск, Россия

Синтетическая биология – быстро развивающаяся отрасль науки, нацеленная на создание биологических систем с предсказанными свойствами. При этом она использует достижения современной биологии, программирования и компьютерного моделирования, а также инженерной отрасли для создания биологических объектов, обладающих набором заранее заданных пользовательских свойств. Развитие синтетической биологии было обусловлено множеством технологических разработок в каждой из упомянутых отраслей. Так, значительное снижение стоимости технологии секвенирования ДНК привело к наработке больших объемов данных о генетических последовательностях различных организмов. Снижение стоимости синтеза последовательностей ДНК в соответствии с законом Мура позволило создавать библиотеки синтетических генов, представляющие потенциальный интерес в работе генных инженеров без необходимости использования традиционных и трудоемких методов молекулярной биологии. Благодаря развитию системной биологии сформировано глубокое понимание взаимосвязей и функций природных биологических моделей, а также построены прогностические модели, описывающие молекулярные процессы на клеточном и системном уровнях. Комбинация вышеперечисленных факторов создала возможность осознанного изменения природных биологических систем. В данном обзоре обсуждается современное состояние подходов к синтезу олигонуклеотидов для последующей сборки генных конструкций и к ферментативной сборке генов. Освещены аспекты использования различного программного обеспечения для подбора олигонуклеотидов для последующей сборки генов, проверки точности синтезированной последовательности генов, а также исправления ошибок.

**Ключевые слова:** синтез олигонуклеотидов; синтез искусственных генов; ферментативная сборка генов; полимерная циклическая сборка; лигазная циклическая сборка; коррекция ошибок; проверка последовательности.

### HOW TO CITE THIS ARTICLE:

Shevelev G.Y., Pyshnyi D.V. Modern approaches to artificial gene synthesis: aspects of oligonucleotide synthesis, enzymatic assembly, sequence verification and error correction. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2018;22(5):498-506. DOI 10.18699/VJ18.387

УДК 577.2.08

Поступила в редакцию 02.05.2018

Принята к публикации 09.06.2018

© АВТОРЫ, 2018

Since a first gene was synthesized from scratch in the 70's (Agarwal et al., 1970) and its *in vivo* activity was tested (Ryan et al., 1979) around 50 years has passed, and the technology of artificial gene synthesis has made a significant leap beyond the approaches implemented in the classical works. Nowadays, it relies on the modern methods of oligonucleotide synthesis, the variants of Polymerase Chain Reaction (PCR) (Saiki et al., 1985; Mullis et al., 1986) with high fidelity DNA polymerases (Böhlke et al., 2000), and the sequence verification step using the Sanger sequencing (Sanger, Coulson, 1975; Smith et al., 1986) or high-throughput DNA sequencing (Church, 2006; Hall, 2007; Bosch, Grody, 2008; Schuster, 2008; Shendure, Ji, 2008; Tucker et al., 2009). During the last decade modern approaches to assemble large (up to dozens and hundreds kb) DNA molecules (Gibson et al., 2009; Gibson, 2011) and error correction within synthesized gene have been introduced (Ma et al., 2012b).

### Oligonucleotide synthesis

All the synthesis approaches rely on usage of chemically synthesized oligonucleotides as building blocks for an enzyme-mediated assembly (Ma et al., 2012b; Kosuri, Church, 2014).

#### Solid-phase oligonucleotide synthesis

The most widely used approach to oligonucleotide synthesis has been the solid-phase phosphoramidite method developed in the 80's (Beaucage, Caruthers, 1981; Matteucci, Caruthers, 1981). The DNA oligonucleotides are synthesized from the 3' to 5' end by consecutive coupling of activated building deoxynucleoside phosphoramidites to an initial deoxynucleoside attached to a solid support (usually the support is a controlled pore glass (CPG) or highly cross-linked polystyrene beads) by its 3'-OH group (Ellington, Pollard, 2000). A solid support matrix is placed into several or several hundred parallel flow-through individual reaction columns, and while synthesis all the reagents necessary for the synthesis cycle flow through the solid support column. A full cycle to add a single nucleotide consists of several stages: (1) deprotection: acid is used to remove a Dimethoxytrityl (DMT) group from the 5'-end of a growing oligonucleotide chain; (2) coupling: the solid-phase 5'-OH group of an attached oligonucleotide reacts with tetrazole-activated nucleoside phosphoramidite producing the elongated product of reaction; (3) capping: treatment of the uncoupled 5'-OH groups by acetic anhydride to minimize the products of deletion mutations; (4) oxidation: conversion of an unstable phosphite triester into a stable phosphate prior to the next step of detritylation. The cycle repeats until a full-length oligonucleotide is obtained. The synthesized oligonucleotide is then cleaved from solid support by treatment with a strong base such as ammonia in aqueous solution and the remaining protection groups are removed during the cleavage. For gene synthesis application high quality and homogeneity of synthesized oligonucleotides is critically important (Xiong et al., 2006) since the synthesized oligonucleotides need to be purified either by the high performance liquid chromatography (HPLC) or by polyacrylamide gel electrophoresis (PAGE) methods.

Synthesis of oligonucleotides with the length up to 100 nt using solid-phase phosphoramidite method is a standard for gene synthesis industry and other applications. The limiting

factor for overcoming barrier over 100 nt during oligonucleotide synthesis is stochastic depurination side reactions which accompany acid deprotection step (Hall et al., 2009). Vendors such as Integrated DNA Technologies (IDT), however, declare possibility to synthesize high quality oligonucleotides up to 200 nt using a special support and synthetic protocol for low-yield and long-oligonucleotide synthesis.

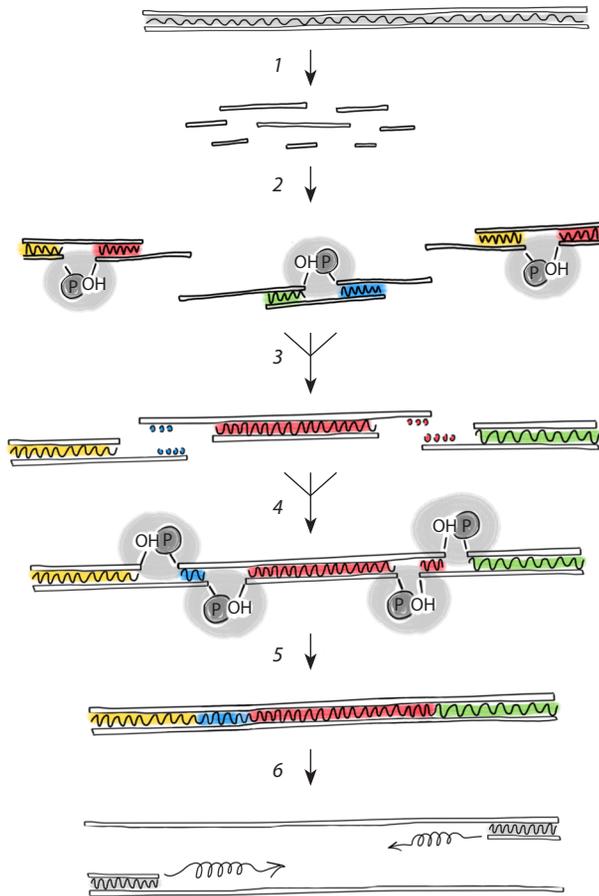
As an alternative to chemical oligonucleotide synthesis, T4 RNA ligase-mediated solid-phase enzymatic oligonucleotide synthesis approach was proposed (Schmitz, Reetz, 1999). The approach is based on coupling reaction between a solid-phase attached oligonucleotide primer and the mononucleoside 3',5'-biphosphate mediated by T4 RNA ligase in a water solution. The terminal phosphate blocking group is removed enzymatically by alkaline phosphatase, and all excess reagents are simply washed off from the resin. The next nucleoside 3',5'-biphosphate can then bind to the end of the chain. Following  $\text{NH}_2\text{OH}$ -induced cleavage from the resin gives the desired elongation of the product. The drawback of this approach is relatively low rate of elongation (48 h) due to the kinetic properties of the enzyme. However, possible automation and molecular biological optimization of enzyme using directed evolution can make the enzymatic approach a viable alternative to today's chemical synthesis approach.

#### Parallel oligonucleotide synthesis using microchips

The oligonucleotide amounts necessary for gene synthesis application are relatively small (picomols) while modern DNA synthesizers can produce oligonucleotides in a higher excess (nanomolar and micromolar) making this approach redundant. Significant improvements in performance and cost-reduction may be achieved by parallelization and miniaturization of oligonucleotide synthesis platforms. Thus, microchip-based approaches previously used for DNA diagnostics and sequencing have been adopted for multi-parallel synthesis of oligonucleotide arrays. Several constructive implementations such as ink-jet DNA printing (Lausted et al., 2004), microfluidic devices (Zhou et al., 2004; Huang et al., 2009; Lee et al., 2010), light-directed (Richmond et al., 2004) and electrochemical (Egeland, Southern, 2005; Chow et al., 2009) microarray synthesis and LED-controlled capillaries (Blair et al., 2006) synthesis have been independently developed. A conditions modification during the detritylation step of chip-based oligonucleotide synthesis lead to improved quality of oligonucleotides with the possible length up to 150 nt (LeProust et al., 2010). However, the limitation of all microarray synthesis methods is the attomoles ( $10^{-18}$  mol) of an individual oligonucleotide that is insufficient for a gene assembly. PCR of selected subpools allows amplifying synthesized oligos at acceptable quantities (Kosuri et al., 2010; Schmidt et al., 2015) for a subsequent multiple gene assembly using barcoded magnetic beads (Plesa et al., 2018).

#### Gene assembly

Enzymatic assembly of synthetic oligonucleotides into long double-stranded DNA fragments with the length up to several hundred bp or even several kb (Ma et al., 2012b) is the next step of gene synthesis with numerous approaches developed since this technology was implemented. In this paper we will discuss several protocols for gene assembly and their parameters.



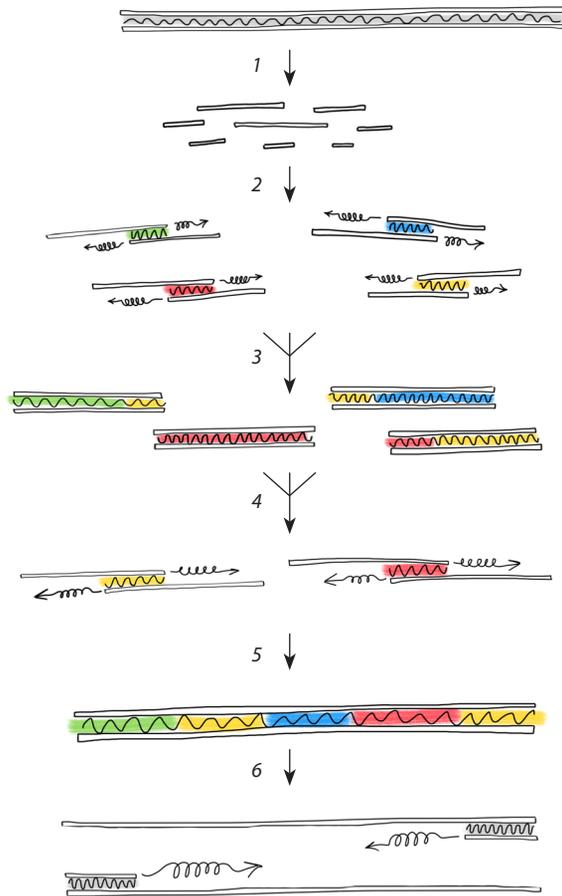
**Fig. 1.** LCA assembly of a DNA construct.

A target sequence is shown on the top of the figure. A set of oligonucleotides synthesized to build the construct (1). A set of 5'-phosphorylated oligonucleotides is assembled in stoichiometric ratio and annealed. Ligase seals the nicks (2). The double stranded DNA is melted and reannealed with extension products and any remained oligos (3) and ligase seals the nicks again (4) until full length product appears (5). Amplification of the full-length product using PCR (6).

### Ligase cycling assembly (LCA)

LCA utilizes the properties of *Taq* DNA ligase to repair single stranded 5'-P, 3'-OH breaks (nicks) within a double stranded DNA at increased temperature (50–60 °C). Conjunction of two oligonucleotides (one 5'-phosphorylated and another with 3'-OH) takes place when they are annealed to a third template oligonucleotide. After joining of oligonucleotides the reaction mixture contains multiple oligonucleotides heated to 95 °C until all the complexes are denatured and further ligation of longer fragments takes place. These steps (annealing, ligation and denaturation) repeat a number of cycles and after that the reaction mixture is amplified with terminal primers and DNA polymerase until the full-length product is obtained (Fig. 1).

One of the alternative implementation of LCA is blunt-end ligation of double stranded DNA blocks using streptavidin-coated magnetic beads and biotin bonded oligonucleotides joined with subsequently added 5'-phosphorylated DNA duplexes (Dietrich et al., 1998). The latest modifications of this approach utilize 5'-phosphorylated single-stranded oligonucleotides (Pengpumkiat et al., 2016).



**Fig. 2.** PCA assembly of a DNA construct.

A target sequence is shown on the top of the figure. A set of oligonucleotides is synthesized to build the construct (1). A set of oligonucleotides is assembled in stoichiometric ratio and annealed (2). Polymerase extends the chain in 3' direction until the end of the template oligo is reached (3). The double stranded DNA is melted and reannealed with extension products and any remained oligos (4). Each extension cycle results in a longer products until the full-length product is synthesized (5). The full-length product is amplified with PCR using terminal primers (6).

### Polymerase cycling assembly (PCA)

PCA is the most widely used method of gene synthesis based on the polymerase chain reaction (PCR) simultaneously carried out in multiple places within the set of short complementary overlapping oligonucleotides forming paired complexes with sticky ends (Fig. 2).

Each step of filling the gaps by polymerase is followed by denaturation and annealing steps and leads to formation of another set of elongated complexes with sticky ends. The process repeats until a full-size product appears (with some side products formed) in the mixture and the final product is amplified with conventional PCR using terminal primers. In the earlier experiments using PCA, in the process of gene amplification with terminal primers gene assembly was performed in a separate tube. Later on this approach was adopted for one-tube assembly (Wu et al., 2006; TerMaat et al., 2009). The desired fragment assembled in a “single-pot” cycled enzymatic reaction or in a several step assembly of separately preassembled structural blocks of the final genetic circuit. Similarity with recursion process gave it its name, so

it is known as recursive PCR. The approach was proposed in 1992 and used for the synthesis of human lysozyme gene (Prodromou, Pearl, 1992). The structure of oligonucleotides overlapping ends was specially designed to obtain the melting temperature within the range 52–56 °C to minimize heterodimers and the oligonucleotides were purified with PAGE. Gap-less implementation of recursive PCR was used for 2703-bp plasmid assembly using 134 unpurified 40nt oligonucleotides with 20nt sticky ends (Stemmer et al., 1995).

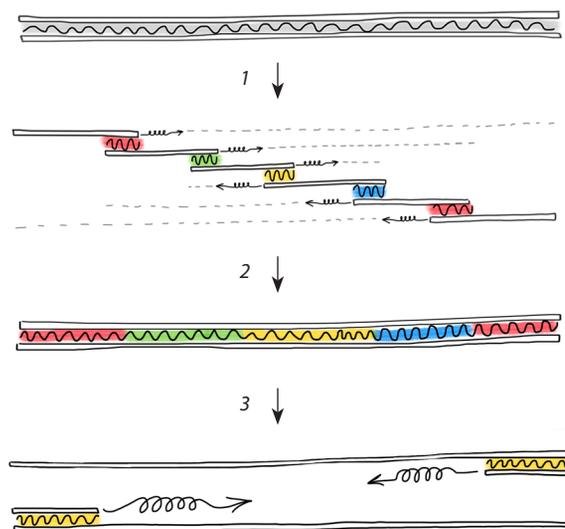
Since the first usage, PCA approach has been thoroughly investigated and several improvements of the approach have been developed, e. g. primeSTAR HS DNA polymerase was proposed as the best-choice polymerase for efficient high-fidelity gene synthesis (Cherry et al., 2008). Utilization of real-time PCR with intercalating dyes has given the insight information concerning the optimal amount of PCR cycles during gene assembly: the authors declared the amount of PCR cycles of about two times more than the theoretically calculated minimal amount of cycles necessary for assembling a full-length product (Ye et al., 2009). Also, the optimal oligonucleotide concentrations while gene assembly was found to be in a range 10–60 nM, external primer concentration – 0.2–1 μM (Wu et al., 2006) and dNTP increased concentration – up to 1 mM each (Ye et al., 2009).

Modification of PCA approach known as TopDown one-step gene synthesis was proposed as advanced technique (Ye et al., 2009). The essential idea of the approach is in a difference between the  $T_m$  values of complexes corresponding to the overlapping sites of oligonucleotides and terminal primers. The outer primers ( $T_m \sim 50$  °C) and inner oligonucleotides ( $T_m \sim 65$  °C) were designed with a melting temperature difference of  $\sim 15$  °C. A higher annealing temperature (65 °C) was used for the first 20 cycles of product assembly and after that the annealing temperature was lowered to 50 °C for the next cycles of full-length product amplification using outer primers. Further improvement of the TopDown synthesis led to the automatic kinetics switch approach (Cheong et al., 2010). The additional feature of this kinetic approach is flanking tail at the outer primer sequence. Primers has two regions with melting temperatures  $T_{p1}$  and  $T_{p2}$ , where  $T_{p1}$  is the temperature of a gene-specific region without a flanking tail, and  $T_{p2}$  is the melting temperature of a primer including a flanking tail.  $T_{p1}$  is lower than the melting temperature of the oligonucleotides assembly and  $T_{p2}$  is  $\sim 72$  °C. When the full-length template appears, the outer primers first created full-length DNA with flanked tails, causing the shift in melting temperatures of outer flanked template to  $T_{p2}$ . The approach provided kinetic switch from inner oligonucleotides assembly to full-length template amplification during one-step gene assembly.

#### Thermodynamically balanced inside-out (TBIO) approach

Another PCR-based approach is the TBIO gene synthesis (Gao et al., 2003). The approach utilizes the process of primer dimers extension under PCR conditions in the absence of template. The process of gene synthesis starts in the center of gene sequence between two overlapping oligonucleotides (Fig. 3).

After extension by PCR, the annealing and extension of the next pair of primers between previously synthesized DNA duplex hybridized through specially designed overlapped sequences with adjusted melting temperatures (i. e. thermody-



**Fig. 3.** TBIO assembly of a DNA construct.

A target sequence is shown on the top of the figure. A set of synthesized oligonucleotides is mixed in concentration increasing from inside to the outside primers (1). Polymerase extends the chain in 3' direction until the end of the oligo template is reached. The double stranded DNA is melted and reannealed with external primers for further elongation by polymerase. Each extension cycle results in a longer products until the full-length product is synthesized (2). The full-length product is amplified with PCR using terminal primers (3).

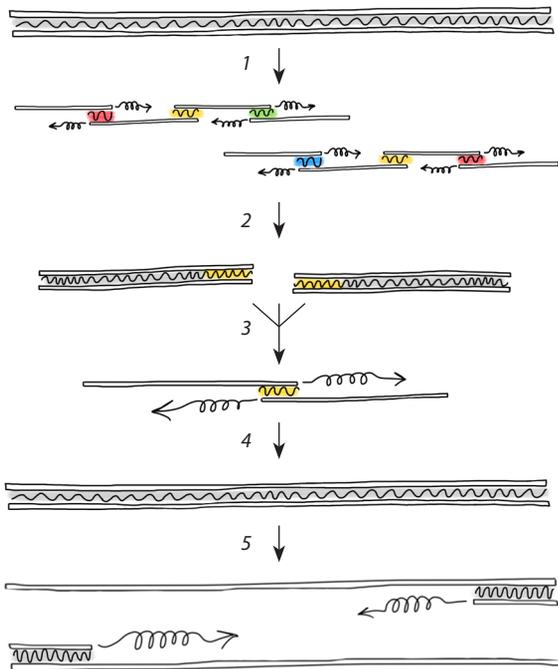
namically balanced) begins. The next iteration with the third pair of primers leads to further elongation of the fragment. The primers in a TBIO assembly are added with gradient concentration: the most internal primers are added in the lowest concentration and the most external primers – in the highest concentration facilitating reaction in the direction inside out. Several gene fragments up to 400–500 bp are independently assembled using a TBIO approach and gel-purified to remove PCR byproducts. Then, these fragments are assembled in one larger gene circuit.

Another strategy for gene synthesis utilizes a combination of Dual Asymmetric (DA) and Overlap Extension PCR (OE-PCR) (Young, Dong, 2004). The oligonucleotides are designed in the same manner as it is done in the PCA approach to contain overlapping regions with or without a gap. Two pairs of complementary oligonucleotides are mixed in one tube and extended by *Pfu* polymerase to double-stranded blunt-ended blocks (Fig. 4).

Several blocks are assembled independently and then mixed with each other to make a whole gene assembly. Reduction of the gap size or its total elimination leads to higher fidelity of the gene assembly during PCR process. To further decrease the mutation product ratio, the full length PCR products are denatured, reannealed and cleaved by T7 endonuclease I. In this manner several genes from 470 bp to 1.2 kb are synthesized. The approach is relatively low-cost because it does not require gel purification of enzymatic phosphorylation.

#### Error correction

The source of the errors presented in a synthetic gene can be either the product of oligonucleotide synthesis or of enzymatic gene assembly. Various strategies can be utilized to reduce the error occurrence during different stages of gene synthesis.



**Fig. 4.** DA OE-PCR assembly of a DNA construct.

A target sequence is shown on the top of the figure. Two pairs from the set of synthesized oligonucleotides are mixed in one tube and extended by polymerase to double-stranded blunt-ended blocks (1). Several blocks are assembled independently then mixed with each other to make a whole gene assembly (2–4). The full-length product is amplified with PCR using terminal primers (5).

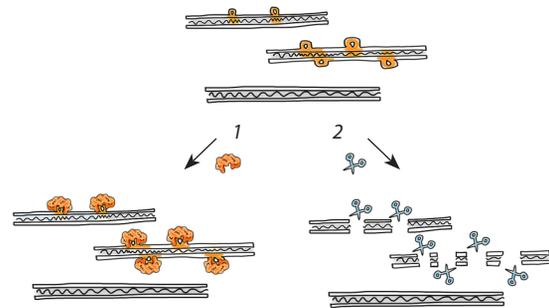
### Correction of errors in synthetic oligonucleotides

Since the efficacy of solid-phase oligonucleotide synthesis is less than 100 %, the most common types of errors are deletions and insertions. The deletions occur due to an incomplete capping step of solid-phase oligonucleotide synthesis with the rate up to 0.5 % per nucleotide, while the insertions happen due to tetrazole cleavage of DMT group with the rate below 0.4 % per nucleotide. Another error, depurination (cleavage of the N-glycosidic bond), occurs under acidic conditions during detritylation and causes apurinic sites within an oligonucleotide chain (Ellington, Pollard, 2000). Such modifications of synthesis protocol as increased coupling time of the phosphoramidite, additional methylene chloride wash steps prior and subsequent to deblocking, using of dichloroacetic acid (DCA) instead of trichloroacetic acid (TCA) for deblocking, help to increase the product yield.

Purification of oligonucleotides with HPLC or PAGE helps to minimize the amount of side products containing deletions and insertions (Andrus, Kuimelis, 2001). However, depurination-type errors does not alter the mobility of long oligonucleotide at PAGE and retention time during HPLC purification.

### Correction of errors from synthetic genes

Synthetic oligonucleotides accumulate the errors remained after solid-phase synthesis. Additionally, enzymatic gene assembly may introduce more errors. One of the possible solutions is cloning of a synthesized gene and identifying a correct sequence. If a correct clone cannot be isolated, site-directed



**Fig. 5.** Error correction in synthesized DNA constructs using mismatch recognition proteins.

MutS binds to gene sequences contained mismatches (1). The T7 endonuclease cleaves mismatched sequences (2).

mutagenesis needs to be applied to correct all the remaining errors. This step is costly and time-consuming. Another approach is to use DNA mismatch recognition proteins to remove incorrect products from synthetic genes. There are two types of such proteins differ in action: (1) mismatch binding proteins, such as *Taq* MutS and (2) mismatch cleaving proteins, for instance, T7 endonuclease or MutHLS complex (Fig. 5).

The MutS protein is a part of bacterial MutHLS DNA repair system (Smith, Modrich, 1997). It detects and binds to a DNA molecule containing single-strand loops and mismatches *in vivo*. In one approach MutS protein isolated from *Thermus aquaticus* was utilized for selection of error-free gene sequences using gel-shift assay (Carr et al., 2004). This method has reduced errors by >15-fold relative to conventional gene synthesis techniques, yielding DNAs with one error per 10 kb. The mismatch cleaving proteins (2) cleave mismatch regions within incorrect DNA complexes. The cleaved complexes can be built into an error-free complex by a polymerase chain assembly (Binkowski et al., 2005) or removed by size selection or exonuclease degradation (Bang, Church, 2008). A combination of mismatch-specific endonucleases such as single-strand-specific nucleases S1 and P1, mung bean nuclease and CEL I nuclease (Desai, Vepatu, 2003), mismatch repair endonuclease MutH (Smith, Modrich, 1997), and resolvases, such as T7 endonuclease I, *E. coli* endonuclease V and T4 endonuclease VII (Ma et al., 2012a) can be used for error removal from synthetic genes. This approach was also applied for error correction in both column and microarray synthesized oligonucleotides using Surveyor nuclease (Saaem et al., 2012; Currin et al., 2014). However, it is unclear which type of mismatch-specific endonucleases is the most effective for error removal (Sequeira et al., 2016).

Another approach for correct gene harvesting is functional selection. The gene of purpose could be fused to a selection marker, such as luciferase gene (Yarimizu et al., 2015). In the case when a marker gene is expressed after cloning, the luciferin-producing clones could be sequenced further. This approach, however, works only for protein-coding genes and is not effective for non-functional mutations.

### Blocks assembly

The typical length of a gene fragment that can be synthesized using the above-mentioned methods is about 1 kb. The

synthesized fragments need to be cloned into a plasmid for sequence verification, the verified fragments combined into a larger gene circuit.

One of the approaches to combine several genetic fragments is the BioBrick strategy allowing assembling a standardized DNA fragment flanked with restriction sites for restriction/ligation reactions (Knight, 2003). Different BioBrick fragments represent various biologically valuable modular parts such as promoters, ribosome-binding sites, coding sequences and transcriptional terminators that can be assembled in a combinatorial way (Shetty et al., 2008). The approach allows assembling several genes and regulatory elements. The method's drawback is the "scars" between modules caused by restriction/ligation assembly.

Utilization of the type II restriction enzymes can be a solution for scarless assembly of gene clusters with the length of about 32 kb (Kodumal et al., 2004). This family of enzymes cut outside of a non-palindromic recognition site has a 4nt overhang, which can be specially chosen for assembling of neighboring fragments without scars (Szybalski et al., 1991). This method is called Golden Gate and has been utilized for assembly of multiple inserts (Engler et al., 2008).

An alternative to restriction/ligation methods is a group of overlap assembly methods (Quan et al., 2009). Similarly to OE-PCR, the overlap methods calls for homologous DNA fragments between blocks to be joined by *in vitro* recombination. There are several commercial cloning kits such as Gateway (Thermo) using proprietary recombinase based on site-specific recombination system of bacteriophage lambda to shuttle sequences between plasmids bearing specific flanking compatible recombination attachment (att) sites (Liang et al., 2013). However, carrying out of a multiple assembly projects with this technique is not convenient because it requires specific sites to be present in each act of assembly. Another kit for cloning into plasmids is In-Fusion (Clontech) which is more promising since it only needs a 15nt sequence overlap between digested plasmid and gene fragment or between gene fragments, which significantly expands the possibility of its application for large circuits assembly (Sleight et al., 2010). Sequence and ligation independent cloning (SLIC) approach utilizes 3'-5' proofreading exonuclease activity of the T4 DNA polymerase in the absence of dNTP to create overhangs of 30 nt after 30 min incubation with double stranded DNA fragments. DNA overhangs anneal with homologous excised fragments. Recombinase RecA is used to insert the fragment into a plasmid. As a demonstration of applicability of this approach for parallel assembly, 10 fragments with the size ranges between 275 and 980 bp with 40 bp overlaps were inserted into a 3.1 kb vector (Li, Elledge, 2007).

The recombinase-free approach named Gibson Assembly is quite useful for multiple assembly of dsDNA fragments with overlapping ends (Gibson et al., 2009). The approach utilizes a cocktail of three enzymes: (1) exonuclease that chews back the end of the fragments and exposes ssDNA overhangs that are specifically annealed; (2) polymerase that fills the gaps within the annealed products; (3) ligase that covalently sews the fragments together. The authors proposed one-step implementation of this technique using exonuclease III preserving its functional properties in the presence of dNTP, antibody-bound *Taq* polymerase preventing competition with exonucle-

ase for 3' ends of DNA and *Taq* ligase placed in a single tube in a thermocycler. Three 5-kb DNA fragments with 40 bp overlaps have been assembled and cloned into *E. coli* given 15 kb insert. Moreover, an assembly of 25 % *M. genitalium* genome using two over-100 kb DNA fragments with 257 bp overlap, and its further complete genome has made the Gibson assembly the most universal and convenient method for large scale gene assembly. It has also been adopted for oligonucleotide-scale to gene-scale assembly: the mouse mitochondrial genome was assembled from 60-mer oligonucleotides using this approach (Gibson et al., 2010).

### Oligonucleotides design software

One of the main limitations in gene synthesis is designing a set of oligonucleotides that should meet several criteria: (1) similar values of  $T_m$  between overlapping fragments; (2) the length of oligonucleotides within the set should not exceed 50–60 nt; (3) oligonucleotides must be annealed only with their neighborhood oligonucleotides in the set and all abnormal intra and intermolecular annealing must be excluded by a splitting algorithm. Designing oligonucleotides for gene synthesis purposes manually is time-consuming because many factors such as GC content, restriction sites, overlapping fragments, codon frequency, and so on need to be taken into account. Various different software packages have been developed in order to optimize the gene design and synthesis process.

One of the most convenient programs used for these purposes is DNAWorks (<https://hpcwebapps.cit.nih.gov/dnaworks/>) (Hoover, Lubkowski, 2002). This program allows designing oligonucleotides for gapped and gapless overlap extension PCR assembly as well as TBIO assemblies with specified  $T_m$  range and oligonucleotide length. It is also capable to transfer from amino acid to nucleotide sequence using about 10 different codon frequency tables as well as custom user defined table.

Gene2Oligo (<http://berry.engin.umich.edu/gene2oligo/>) (Rouillard et al., 2004) allows one to design oligonucleotides for LCA and gapless PCA assemblies. GeneDesign (<http://54.235.254.95/gd/index.html>) (Louw et al., 2011) provides oligos design for gapped PCA assembly.

GeneDesign (<http://54.235.254.95/gd/index.html>) (Villalobos et al., 2006; Richardson et al., 2010) is a program enabling one to divide long DNA sequence into ~500 bp fragments joined by restriction sites and to chop this fragments into sets of overlapping oligonucleotides. This approach can be successfully applied for very long DNA sequences up to several kb. GeneDesign also provides the "Codon Juggling" feature which is capable to produce sequence that is codon-optimized for expression and, as different as possible from the original sequence (and still coding the same protein). Either, addition and removing of restriction sites within sequence is possible.

TmPrime (<http://prime.ibn.a-star.edu.sg>) (Bode et al., 2009) and Genecomposer (<http://www.genecomposer.net>) (Lorimer et al., 2009) are other software solutions for gene design with expanded set of functions analogous to GeneDesign. Currently, not available on the Internet.

Another algorithm based on reliable gradient optimization and the derivative objective function approximated with a central difference operator was proposed for gene synthesis

application by Louw et al. (2011). The authors compared the results of their gradient optimization approach against the one using Gene2Oligo and concluded their approach produces higher purity and yield of assembled genes.

DNASynth (<http://dnasynth.sourceforge.net/>) (Nowak et al., 2015) is a relatively new algorithm that designs the whole artificial gene synthesis process, developing the optimal nucleotide sequence encoding a given peptide for a given host organism and determining the best long DNA LCA-based assembly protocol.

Microchip-based gene assembly protocols needs special software for oligonucleotide arrays design. PICKY (Birla, Chou, 2015) is software that can be applied for multiple one-pot gene assemblies from dsDNA fragments using the Gibson Assembly protocol. It provides thermodynamic analysis that identifies all unique junctions in gene, where consecutive DNA fragments are specially designed to connect to each other.

## Perspectives and applications

Current applications of synthetic DNA combine synthesis of genetic constructs and metabolic pathways as well as synthesis of artificial genomes and production of artificial genome-based microorganisms (Gibson et al., 2008). Last improvements of long DNA assembly approaches lead to the possibility of creation artificial living system based on *M. genitalium* with minimized functional genome (Hutchison et al., 2016). Recently, the Human Genome Wright project has been announced by a group of US scientists (Boeke et al., 2016). The goal of the project is to synthesize a human genome to improve understanding of the interconnections between genes. For this sake, the additional aim of the project is to catalyze the reduction in price per base, which is now floats around 0.2 \$ per 1 bp, to as low as a few cents per bp. Further improvements in microchip-based DNA synthesis technologies will help to drop down the current price per 1 bp.

Another significant application of synthetic biology is metabolic engineering. By designing metabolic systems working in parallel with host metabolic pathways, scientists can program cells for practical applications such as bioremediation (Wasilkowski et al., 2012; Garbisu et al., 2017) and production of industrial compounds (Adrio, Demain, 2010). For example, engineering a pathway for biosynthesis of antimalarial compound Artemisinin in yeast has lowered its price by about 10 times if compared to the previously used method (Paddon, Keasling, 2014).

Application of next generation sequencing technologies (NGS) for synthesis and amplification of error-free oligonucleotides (Schwartz et al., 2012; Lim et al., 2018; Plesa et al., 2018) leads to expanding the application of chip-synthesized oligonucleotide pools into *in vitro* directed evolution methods. For instance, development of multiplex automated genome engineering (MAGE) platform for large-scale programming and evolution of cells and its application to optimization of 1-deoxy-D-xylulose-5-phosphate (DXP) biosynthesis pathway in *E. coli* for overproduction of the industrially important isoprenoid lycopene has allowed for fivefold increase in lycopene amounts which is quite remarkable achievement (Wang et al., 2009). Recent developments of several chip-based oligonucleotide synthesis platforms will enable us to amplify error-free subpools of oligos with further independent

assembly of homologous gene libraries allowed authors to rationally explore sequence-function relationships at unprecedented scale (Plesa et al., 2018).

## Acknowledgments

This work has been supported by Russian Science Foundation (no. 17-74-10157).

## Conflict of interest

The authors declare to have no conflict of interest.

## References

- Adrio J.-L., Demain A.L. Recombinant organisms for production of industrial products. *Bioeng. Bugs.* 2010;1(2):116-131. DOI 10.4161/bbug.1.2.10484.
- Agarwal K.L., Büchi H., Caruthers M.H., Gupta N., Khorana H.G., Kleppe K., Kumar A., Ohtsuka E., Rajbhandary U.L., Van De Sande J.H., Sgaramella V., Weber H., Yamada T. Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast. *Nature.* 1970;227:27-34. DOI 10.1038/227027a0.
- Andrus A., Kuimelis G.R. Analysis and purification of synthetic nucleic acids using HPLC. *Curr. Protoc. Nucleic Acid Chem.* 2001;1(1):unit 10.5. DOI 10.1002/0471142700.nc1005s01.
- Bang D., Church G.M. Gene synthesis by circular assembly amplification. *Nat. Methods.* 2008;5(1):37-39. DOI 10.1038/nmeth1136.
- Beaucage S.L., Caruthers M.H. Deoxynucleoside phosphoramidites – a new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* 1981;22(20):1859-1862. DOI 10.1016/S0040-4039(01)90461-7.
- Binkowski B.F., Richmond K.E., Kaysen J., Sussman R.M., Belshaw P.J. Correcting errors in synthetic DNA through consensus shuffling. *Nucleic Acids Res.* 2005;33(6):1-8. DOI 10.1093/nar/gni053.
- Birla B.S., Chou H.H. Rational design of high-number dsDNA fragments based on thermodynamics for the construction of full-length genes in a single reaction. *PLoS One.* 2015;10(12):e0145682. DOI 10.1371/journal.pone.0145682.
- Blair S., Richmond K., Rodesch M., Bassetti M., Cerrina F. A scalable method for multiplex LED-controlled synthesis of DNA in capillaries. *Nucleic Acids Res.* 2006;34(16). DOI 10.1093/nar/gkl641.
- Bode M., Khor S., Ye H., Li H.M., Ying J.Y. TmPrime: Fast, flexible oligonucleotide design software for gene synthesis. *Nucleic Acids Res.* 2009;37(Web Server issue):W214-W221. DOI 10.1093/nar/gkp461.
- Boeke J.D., Church G., Hessel A., Kelley N.J., Arkin A., Cai Y., Carlson R., Chakravarti A., Cornish V.W., Holt L., Isaacs F.J., Lajoie M., Lessor T., Lunshof J., Maurano M.T., Mitchell L.A., Rine J., Rosser S., Sanjana N.E., Silver P.A., Valle D., Wang H., Way J.C., Yang L. The Genome Project-Write. *Science.* 2016;353(6295):126-127. DOI 10.1126/science.aaf6850.
- Bosch J.R., Grody W.W. Keeping up with the next generation: massively parallel sequencing in clinical diagnostics. *J. Mol. Diagn.* 2008;10(6):484-492. DOI 10.2353/jmoldx.2008.080027.
- Böhlke K.F., Pisani M., Vorgias C.E., Frey B., Sobek H., Rossi M., Antranikian G. PCR performance of the B-type DNA polymerase from the thermophilic euryarchaeon *Thermococcus aggregans* improved by mutations in the Y-GG/A motif. *Nucleic Acids Res.* 2000;28(20):3910-3917. DOI 10.1093/nar/28.20.3910.
- Carr P.A., Park J.S., Lee Y., Yu T., Zhang S., Jacobson J.M. Protein-mediated error correction for *de novo* DNA synthesis. *Nucleic Acids Res.* 2004;32(20):e162. DOI 10.1093/nar/gnh160.
- Cheong W.C., Lim L.S., Huang M.C., Bode M., Li M.H. New insights into the *de novo* gene synthesis using the automatic kinetics switch approach. *Anal. Biochem.* 2010;406:51-60. DOI 10.1016/j.ab.2010.06.036.
- Cherry J., Nieuwenhuijsen B.W., Kaftan E.J., Kennedy J.D., Chanda P.K. A modified method for PCR-directed gene synthesis from

- large number of overlapping oligodeoxyribonucleotides. *J. Biochem. Biophys. Methods*. 2008;70:820-822. DOI 10.1016/j.jprot.2007.12.009.
- Chow B.Y., Emig C.J., Jacobson J.M. Photoelectrochemical synthesis of DNA microarrays. *Proc. Natl. Acad. Sci. USA*. 2009;106(36):15219-15224. DOI 10.1073/pnas.0813011106.
- Church G.M. Genomes for all. *Sci. Am*. 2006;294(1):47-54.
- Currin A., Swainston N., Day P.J., Kell D.B. SpeedyGenes: An improved gene synthesis method for the efficient production of error-corrected, synthetic protein libraries for directed evolution. *Protein Eng. Des. Sel.* 2014;27(9):273-280. DOI 10.1093/protein/gzu029.
- Desai N.A., Vepatu S. Single-strand-specific nucleases. *FEMS Microbiol. Rev.* 2003;26:457-91.
- Dietrich R., Wirsching F., Opitz T., Schwienhorst A. Gene assembly based on blunt-ended double-stranded DNA-modules. *Biotechnol. Tech.* 1998;12(1):49-54. DOI 10.1023/A:1008855526226.
- Egeland R.D., Southern E.M. Electrochemically directed synthesis of oligonucleotides for DNA microarray fabrication. *Nucleic Acids Res.* 2005;33(14):1-7. DOI 10.1093/nar/gni117.
- Ellington A., Pollard J.D., Jr. Introduction to the synthesis and purification of oligonucleotides. *Curr. Protoc. Nucleic Acid Chem.* 2000; App.3:A.3C.1-A.3C.22. DOI 10.1002/0471142700.nca03cs00.
- Engler C., Romy K., Sylvestre M. A one pot, one step, precision cloning method with high throughput capability. *PLoS One*. 2008;3(11):e3647. DOI 10.1371/journal.pone.0003647.
- Gao X., Yo P., Keith A., Ragan T.J., Harris T.K. Thermodynamically balanced inside-out (TBIO) PCR-based gene synthesis: a novel method of primer design for high-fidelity assembly of longer gene sequences. *Nucleic Acids Res.* 2003;31(22):2-11. DOI 10.1093/nar/gng143.
- Garbisu C., Olatz G., Epelde L., Grohmann E., Alkorta I. Plasmid-mediated bioaugmentation for the bioremediation of contaminated soils. *Front. Microbiol.* 2017;8:1966. DOI 10.3389/fmicb.2017.01966.
- Gibson D.G. Enzymatic Assembly of Overlapping DNA Fragments. In: Voigt C. (Ed.). *Synthetic Biology. Pt. B: Computer Aided Design and DNA Assembly*. (Ser. Methods in Enzymology. Vol. 498). Acad. Press, 2011;349-361. DOI 10.1016/b978-0-12-385120-8.00015-2.
- Gibson D.G., Benders G.A., Andrews-Pfannkoch C., Denisova E.A., Baden-Tillson H., Zaveri J., Stockwell T.B., Brownley A., Thomas D.W., Algire M.A., Merryman C., Young L., Noskov V.N., Glass J.I., Venter J.C., Hutchison III C.A., Smith H.O. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*. 2008;319:1215-1220. DOI 10.1126/science.1151721.
- Gibson D.G., Smith H.O., Hutchison III C.A., Venter J.C., Merryman C. Chemical synthesis of the mouse mitochondrial genome. *Nat. Methods*. 2010;7(11):901-903. DOI 10.1038/nmeth.1515.
- Gibson D.G., Young L., Chuang R.-Y., Venter J.C., Hutchison C.A., Smith H.O., Hutchison III C.A. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*. 2009;6(5):343-345. DOI 10.1038/nmeth.1318.
- Hall B.H., Micheletti J.M., Satya P., Ogle K., Pollard J., Ellington A.D. Design, synthesis, and amplification of DNA pools for *in vitro* selection. *Curr. Protoc. Mol. Biol.* 2009;88(1):24.2.1-24.2.27. DOI 10.1002/0471142727.mb2402s88.
- Hall N. Advanced sequencing technologies and their wider impact in microbiology. *J. Exp. Biol.* 2007;210:1518-1525. DOI 10.1242/jeb.001370.
- Hoover D.M., Lubkowski J. DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res.* 2002;30(10):e43. DOI 10.1093/nar/30.10.e43.
- Huang M.C., Ye H., Kuan Y.K., Li M.-H., Ying J.Y. Integrated two-step gene synthesis in a microfluidic device. *Lab. Chip*. 2009;9:276-285. DOI 10.1039/b807688j.
- Hutchison C.A., Chuang R.-Y., Noskov V.N., Nacyra A.-G., Deerinck T.J., Ellisman M.H., Gill J., Kannan K., Karas B.J., Ma L., Pelletier J.F., Qi Z.-Q., Richter R.A., Strychalski E.A., Sun L., Suzuki Y., Tsvetanova B., Wise K.S., Smith H.O., Glass J.I., Merryman C., Gibson D.G., Venter J.C. Design and synthesis of a minimal bacterial genome. *Science*. 2016;351(6280):aad6253. DOI 10.1126/science.aad6253.
- Knight T. Idempotent Vector Design for Standard Assembly of Bricks. MIT Artificial Intelligence Laboratory, 2003;1-11.
- Kodumal S.J., Patel K.G., Reid R., Menzella H.G., Welch M., Santi D.V. Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proc. Natl. Acad. Sci. USA*. 2004;101(44):15573-15578. DOI 10.1073/pnas.0406911101.
- Kosuri S., Church G. Large-scale *de novo* DNA synthesis: technologies and applications. *Nat. Methods*. 2014;11(5):499-507. DOI 10.1038/nmeth.2918.
- Kosuri S., Eroshenko N., Leproust E.M., Super M., Way J., Li B.L., Church G.M. Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nat. Biotechnol.* 2010;28(12):1295-1299. DOI 10.1038/nbt.1716.
- Lausted C., Dahl T., Warren C., King K., Smith K., Johnson M., Saleem R., Aitchison J., Hood L., Lasky S.R. POSaM: a fast, flexible, open-source, inkjet oligonucleotide synthesizer and microarrayer. *Genome Biol.* 2004;5(8):R58.1-R58.17. DOI 10.1186/gb-2004-5-8-r58.
- Lee C.C., Snyder T.M., Quake S.R. A microfluidic oligonucleotide synthesizer. *Nucleic Acids Res.* 2010;38(8):2514-2521. DOI 10.1093/nar/gkq092.
- LeProust E.M., Peck B.J., Spirin K., McCuen H.B., Moore B., Namsaraev E., Caruthers M.H. Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process. *Nucleic Acids Res.* 2010;38(8):2522-2540. DOI 10.1093/nar/gkq163.
- Li M.Z., Elledge S.J. Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nat. Methods*. 2007;4(3):251-256. DOI 10.1038/nmeth1010.
- Liang X., Peng L., Baik C.-H., Katzen F. Single step BP/LR combined Gateway reactions. *BioTechniques*. 2013;55(5):265-268. DOI 10.2144/000114101.
- Lim H., Cho N., Ahn J., Park S., Jang H., Kim H., Han H., Ji H.L., Bang D. Highly selective retrieval of accurate DNA utilizing a pool of *in situ*-replicated DNA from multiple next-generation sequencing platforms. *Nucleic Acids Res.* 2018;46(7):e40. DOI 10.1093/nar/gky016.
- Lorimer D., Raymond A., Walchli J., Mixon M., Barrow A., Wallace E., Grice R., Burgin A., Stewart L. Gene composer: database software for protein construct design, codon engineering, and gene synthesis. *BMC Biotechnol.* 2009;9:36. DOI 10.1186/1472-6750-9-36.
- Louw T.M., Whitney S.E., Termaat J.R., Pienaar E., Viljoen H.J. Oligonucleotide optimization for DNA synthesis. *AIChE J.* 2011;57(7):1912-1918. DOI 10.1002/aic.12410.
- Ma S., Saaem I., Tian J. Error correction in gene synthesis technology. *Trends Biotechnol.* 2012a;30(3):147-154. DOI 10.1016/j.tibtech.2011.10.002.
- Ma S., Tang N., Tian J. DNA synthesis, assembly and applications in synthetic biology. *Curr. Opin. Chem. Biol.* 2012b;16:260-267. DOI 10.1016/j.cbpa.2012.05.001.
- Matteucci M.D., Caruthers M.H. Synthesis of deoxyoligonucleotides on a polymer support. *J. Am. Chem. Soc.* 1981;103(11):3185-3191. DOI 10.1021/ja00401a041.
- Mullis K., Faloona F., Scharf S., Saiki R., Horn G., Erlich H. Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* 1986;51:263-273. DOI 10.1101/SQB.1986.051.01.032.
- Nowak R.M., Wojtowicz-Krawiec A., Plucienniczak A. DNASynth: A computer program for assembly of artificial gene parts in decreasing temperature. *BioMed. Res. Int.* 2015; Article ID:413262. DOI 10.1155/2015/413262.
- Paddon C.J., Keasling J.D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* 2014;12:355-367. DOI 10.1038/nrmicro3240.

- Pengpumpiat S., Koesdjojo M., Rowley E.R., Mockler T.C., Remcho V.T. Rapid synthesis of a long double-stranded oligonucleotide from a single-stranded nucleotide using magnetic beads and an oligo library. *PLoS One*. 2016;11(3):e0149774. DOI 10.1371/journal.pone.0149774.
- Plesa C., Sidore A.M., Lubock N.B., Zhang D., Kosuri S. Multiplexed gene synthesis in emulsions for exploring protein functional landscapes. *Science*. 2018;369(6373):343-347. DOI 10.1126/science.aao5167.
- Prodromou C., Pearl L.H. Recursive PCR: A novel technique for total gene synthesis. *Protein Eng. Des. Sel.* 1992;5(8):827-829. DOI 10.1093/protein/5.8.827.
- Quan J., Tian J. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One*. 2009;4(7):e6441. DOI 10.1371/journal.pone.0006441.
- Richardson S.M., Nunley P., Yarrington R.M., Boeke J.D., Bader J.S. GeneDesign 3.0 is an updated synthetic biology toolkit. *Nucleic Acids Res.* 2010;38(8):2603-2606. DOI 10.1093/nar/gkq143.
- Richmond K.E., Li M.H., Rodesche M.J., Patel M., Lowe A.M., Kim C., Chu L.L., Venkataramaiah N., Flickinger S.F., Kaysen J., Belshaw P.J., Sussman M.R., Cerrina F. Amplification and assembly of chip-eluted DNA (AACED): A method for high-throughput gene synthesis. *Nucleic Acids Res.* 2004;32(17):5011-5018. DOI 10.1093/nar/gkh793.
- Rouillard J., Lee W., Truan G., Gao X., Zhou X., Gulari E. Gene2Oligo: Oligonucleotide design for *in vitro* gene synthesis. *Nucleic Acids Res.* 2004;32(Web Server issue):W176-W180. DOI 10.1093/nar/gkh401.
- Ryan J., Brown E.L., Sekiya T., Kiipper H., Khorana H.G. Total synthesis of a tyrosine suppressor tRNA gene. XVIII. Biological activity and transcription, *in vitro*, of the cloned gene. *J. Biol. Chem.* 1979;254(13):5817-5826.
- Saaem I., Ma S., Quan J., Tian J. Error correction of microchip synthesized genes using Surveyor nuclease. *Nucleic Acids Res.* 2012;40(3):e23. DOI 10.1093/nar/gkr887.
- Saiki R.K., Scharf S., Faloona F., Mullis K.B., Horn G.T., Erlich H.A., Arnheim N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*. 1985;230:1350-1354. DOI 10.1126/science.2999980.
- Sanger F., Coulson A.R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* 1975;94:441-448. DOI 10.1016/0022-2836(75)90213-2.
- Schmidt T.L., Beliveau B.J., Uca Y.O., Theilmann M., Cruz F.D., Wu C.T., Shih W.M. Scalable amplification of strand subsets from chip-synthesized oligonucleotide libraries. *Nat. Commun.* 2015;6:8634. DOI 10.1038/ncomms9634.
- Schmitz C., Reetz M.T. Solid-phase enzymatic synthesis of oligonucleotides. *Org. Lett.* 1999;1(11):1729-1731.
- Schuster S.C. Next-generation sequencing transforms today's biology. *Nat. Methods*. 2008;5(1):16-18. DOI 10.1038/nmeth1156.
- Schwartz J.J., Lee C., Shendure J. Accurate gene synthesis with tag-directed retrieval of sequence-verified DNA molecules. *Nat. Methods*. 2012;9(9):913-915. DOI 10.1038/nmeth.2137.
- Sequeira A.F., Guerreiro C.I., Vincentelli R., Fontes C.M. T7 endonuclease I mediates error correction in artificial gene synthesis. *Mol. Biotechnol.* 2016;58(8-9):573-584. DOI 10.1007/s12033-016-9957-7.
- Shendure J., Ji H. Next generation DNA sequencing. *Eng. Life Sci.* 2008;26(10):1135-1345. DOI 10.1002/elsc.201600121.
- Shetty R.P., Drew E., Knight T.F. Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng.* 2008;2:5. DOI 10.1186/1754-1611-2-5.
- Sleight S.C., Bartley B.A., Lieviant J.A., Sauro H.M. In-Fusion BioBrick assembly and re-engineering. *Nucleic Acids Res.* 2010;38(8):2624-2636. DOI 10.1093/nar/gkq179.
- Smith J., Modrich P. Removal of polymerase-produced mutant sequences from PCR products. *Proc. Natl. Acad. Sci. USA.* 1997;94:6847-6850.
- Smith L.M., Sanders J.Z., Kaiser R.J., Hughes P., Dodd C., Connell C.R., Heiner C., Kent S.B.H., Hood L.E. Fluorescence detection in automated DNA sequence analysis. *Nature*. 1986;321:674-679. DOI 10.1038/321674a0.
- Stemmer W.P.C., 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*. 1995;164:49-53. DOI 10.1016/0378-1119(95)00511-4.
- Szybalski W., Kim S.C., Hasan N., Podhajski A.J. Class-II restriction enzymes – a review. *Gene*. 1991;100:13-26. DOI 10.1016/0378-1119(91)90345-C.
- 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. Meth.* 2009;79(3):295-300. DOI 10.1016/j.mimet.2009.09.015.
- Tucker T., Marra M., Friedman J.M. Massively parallel sequencing: the next big thing in genetic medicine. *Am. J. Hum. Genet.* 2009;85:142-154. DOI 10.1016/j.ajhg.2009.06.022.
- Villalobos A., Ness J.E., Gustafsson C., Minshull J., Govindarajan S. Gene Designer: A synthetic biology tool for constructing artificial DNA segments. *BMC Bioinformatics*. 2006;7:285. DOI 10.1186/1471-2105-7-285.
- Wang H.H., Isaacs F.J., Carr P.A., Sun Z.Z., Xu G., Forest C.R., Church G.M. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*. 2009;460:894-898. DOI 10.1038/nature08187.
- Wasilkowski D., Swedziol Z., Mroziak A. The applicability of genetically modified microorganisms in bioremediation of contaminated environments. *Chemik*. 2012;8(66):817-826.
- Wu G., Wolf J.B., Ibrahim A.F., Vadasz S., Gunasinghe M., Freeland S.J. Simplified gene synthesis: A one-step approach to PCR-based gene construction. *J. Biotechnol.* 2006;124:496-503. DOI 10.1016/j.jbiotec.2006.01.015.
- Xiong A.-S., Yao Q.-H., Peng R.-H., Duan H., Li X., Fan H.-Q., Cheng Z.-M., Li Y. PCR-based accurate synthesis of long DNA sequences. *Nat. Protoc.* 2006;1(2):791-797. DOI 10.1038/nprot.2006.103.
- Yarimizu T., Nakamura M., Hoshida H., Akada R. Screening of accurate clones for gene synthesis in yeast. *J. Biosci. Bioeng.* 2015;119(3):251-259. DOI 10.1016/j.jbiosc.2014.08.006.
- Ye H., Huang M.C., Li M.H., Ying J.Y. Experimental analysis of gene assembly with TopDown one-step real-time gene synthesis. *Nucleic Acids Res.* 2009;37(7):e51. DOI 10.1093/nar/gkp118.
- Young L., Dong Q. Two-step total gene synthesis method. *Nucleic Acids Res.* 2004;32(7):e59. DOI 10.1093/nar/gnh058.
- Zhou X., Cai S., Hong A., You Q., Yu P., Sheng N., Srivannavit O., Muranjan S., Rouillard J.M., Xia Y., Zhang X., Xiang Q., Ganesh R., Zhu Q., Matejko A., Gulari E., Gao X. Microfluidic PicoArray synthesis of oligodeoxynucleotides and simultaneous assembling of multiple DNA sequences. *Nucleic Acids Res.* 2004;32(18):5409-5417. DOI 10.1093/nar/gkh879.