Gene Assembly from Chip-Synthesized Oligonucleotides

Nikolai Eroshenko,[1,5] Sriram Kosuri,[2,3,5] Adam H. Marblestone,[3,4] Nicholas Conway,[3] and George M. Church[2,3]

[1]Harvard School of Engineering and Applied Sciences, Cambridge, Massachusetts
[2]Department of Genetics, Harvard Medical School, Boston, Massachusetts
[3]Wyss Institute for Biologically Inspired Engineering, Boston, Massachusetts
[4]Harvard Biophysics Program, Cambridge, Massachusetts
[5]These authors contributed equally to this work

ABSTRACT

De novo synthesis of long double-stranded DNA constructs has a myriad of applications in biology and biological engineering. However, its widespread adoption has been hindered by high costs. Cost can be significantly reduced by using oligonucleotides synthesized on high-density DNA chips. However, most methods for using off-chip DNA for gene synthesis have failed to scale due to the high error rates, low yields, and high chemical complexity of the chip-synthesized oligonucleotides. We have recently demonstrated that some commercial DNA chip manufacturers have improved error rates, and that the issues of chemical complexity and low yields can be solved by using barcoded primers to accurately and efficiently amplify subpools of oligonucleotides. This unit includes protocols for computationally designing the DNA chip, amplifying the oligonucleotide subpools, and assembling 500- to 800-bp constructs. *Curr. Protoc. Chem. Biol.* 4:1-17 © 2012 by John Wiley & Sons, Inc.

Keywords: oligonucleotide • gene synthesis • nucleic acids • synthetic biology

INTRODUCTION

Gene synthesis has been used in applications as diverse as decoding the genetic code (Nirenberg and Matthaei, 1961), screening industrially useful enzymes discovered through environmental metagenomic sequencing (Bayer et al., 2009), and building custom pathways and genomes (Tian et al., 2004; Ro et al., 2006; Hanai et al., 2007; Gibson et al., 2008). Unfortunately, despite the rapid decrease in the cost of synthesizing short single-stranded oligonucleotides, synthesis of double-stranded gene-sized fragments remains too expensive for ubiquitous adoption by academic laboratories (Carr and Church, 2009). To make large-scale synthesis cheaper, we have recently developed a set of techniques, which allow assembly of DNA from commercially available high-density oligonucleotide chips into gene-sized 500- to 800-bp fragments (Kosuri et al., 2010). Specifically, we showed that PCR could be used to separate the DNA synthesized on a chip into subsets (or, as we will refer to them from here on, subpools) consisting of only the oligonucleotide species necessary to build one particular gene-sized fragment. These methods use inexpensive sources of DNA and require no specialized instrumentation or expertise that cannot be found in a typical molecular biology laboratory.

The major stages of our synthesis pipeline are computational design, chip synthesis, serial PCRs that isolate the oligonucleotides necessary to build each construct, and assembly of the constructs. The key principle is that well-designed primers can amplify a desired subset of oligonucleotides and, thereby, dilute the undesired DNA to the point where it does not interfere with the downstream gene assembly reaction. An off-chip pool (see

Current Protocols in Chemical Biology 4: 1-17, March 2012
Published online March 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/9780470559277.ch110190

Table 1 Nomenclature Used to Describe the Synthesis Protocol

Term	Definition	Source
Off-chip pool	The pool of oligonucleotides cleaved from the DNA microchip	Synthesized on a DNA chip
Plate subpool	The subset of the off-chip pool necessary to build 96 assemblies	Amplified from off-chip pool
Plate subpool-specific primers	A pair of PCR primers that bind to sites shared by all members of a plate subpool	Traditional (not chip) synthesis
Assembly subpool	The subset of a plate subpool necessary to build 1 assembly	Amplified from plate subpool
Assembly subpool-specific primers	A pair of PCR primers that bind to sites shared by all members of an assembly subpool	Traditional (not chip) synthesis
Construction primers	A pair of PCR primers used to assemble one or more assembly subpools	Traditional (not chip) synthesis
Assembly	A 500-800 bp dsDNA construct built from an assembly subpool using a pair of construction primers	Built from assembly subpool

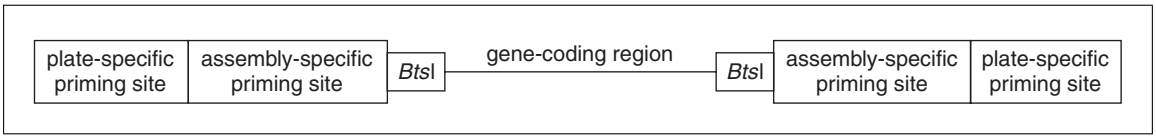

Figure 1 A schematic of the features present on each off-chip oligonucleotide. The gene-coding regions of the oligonucleotides within each assembly subpool partially overlap, allowing them to be assembled into the full-length construct using a high-fidelity polymerase. The gene-coding region is flanked by *Bts*I cut sites that permit enzymatic removal of the subpool-specific priming sites. The gene-coding region is also flanked by a pair of assembly-specific priming sites, which are shared by all the oligonucleotides within a particular assembly subpool. The assembly-specific priming sites are, in turn, flanked by a pair of plate-specific priming sites common to all the oligonucleotides within a particular plate-specific subpool.

Table 1 for nomenclature) consists of oligos with a gene-coding region flanked by nested subpool-specific sequences and a restriction enzyme recognition site for removing the priming sequences (Fig. 1). The subpool-specific sequences act as barcodes for selecting subsets of oligos. Each pair of assembly-specific priming sites is shared by all the oligonucleotides needed to build a particular construct; in turn, each pair of plate-specific subpool priming sites is shared by all the oligos necessary to build 96 constructs. Another pair of primers—the construction primers—are used to amplify the full-length construct.

We have developed a software tool called GASP (Gene Assembly by Subpool PCR) that deconstructs a given set of genes to generate the sequences of oligonucleotides for synthesis on a chip. Using GASP for chip design is described in Basic Protocol 1. Once the oligonucleotides have been designed, synthesized, and cleaved off the chip surface, the off-chip pool is divided into aliquots among plate subpool amplification reactions (Fig. 2). After the PCR, each plate subpool is divided into aliquots into a 96-well plate. Following the addition of assembly-specific primers, another amplification is performed, at which point a restriction enzyme is used to cleave off the priming sites. Basic Protocol 2 provides directions for amplifying the plate and assembly subpools, and Basic Protocol 3 describes the enzymatic removal of priming sites. A polymerase is used to assemble the oligos into the full-length construct, and the assemblies are amplified using the construction primers (which are not necessarily unique for the assembly subpool), as described in Basic Protocol 4.

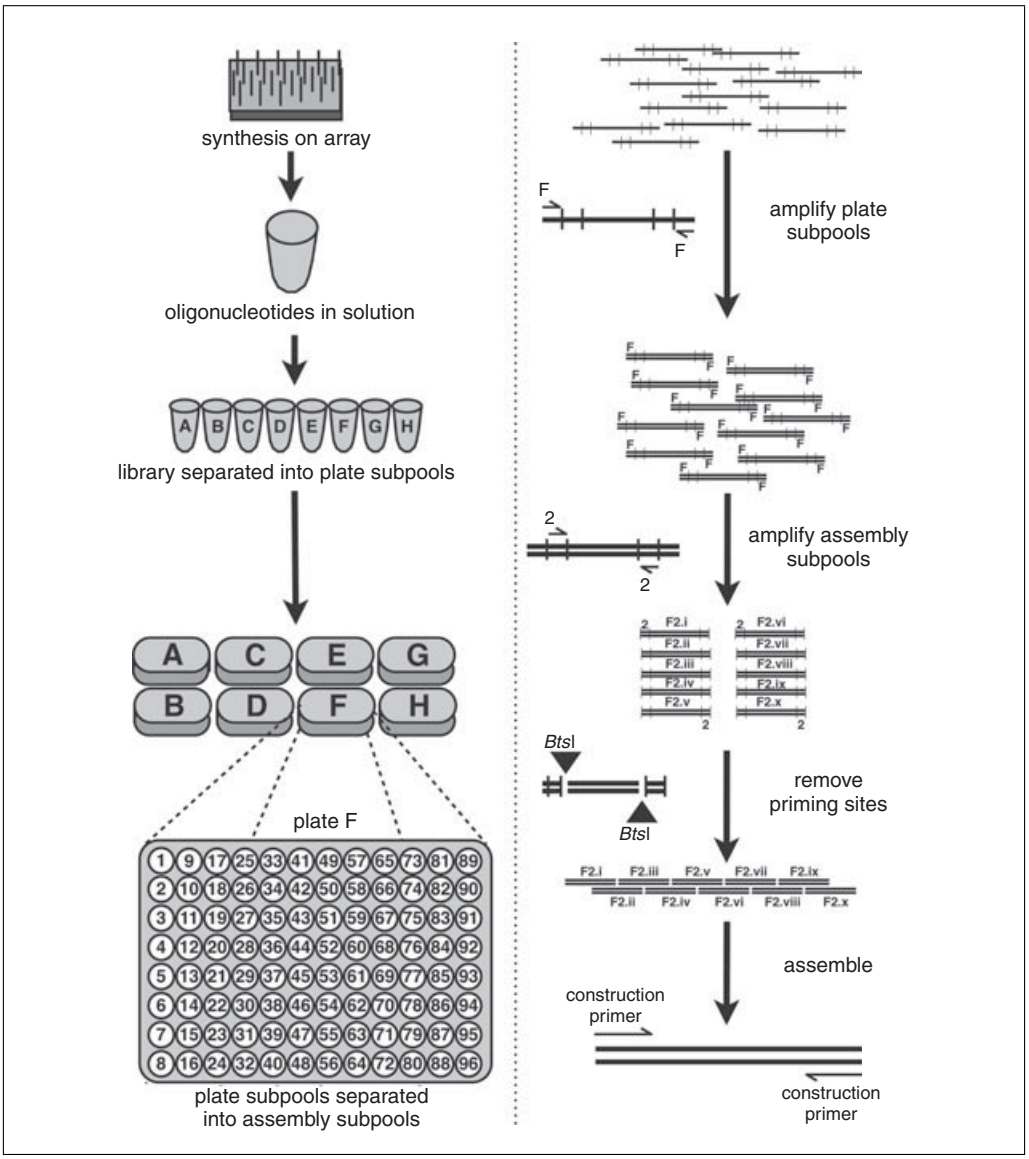

Figure 2 A schematic of the gene synthesis workflow's liquid handling steps (left) and the oligonucleotide processing steps (right). In the left half of the schematic, each off-chip oligonucleotide is drawn as a horizontal line with vertical lines used to indicate the plate and assembly subpool-specific priming sites. After the DNA chip has been synthesized and the oligonucleotides have been cleaved off the array surface, plate-subpool-specific primers are used to amplify only the oligonucleotides needed to build 96 assemblies. In the drawing, the primer pair "F" amplifies all the oligonucleotides that make up plate subpool "F." The plate subpool is then divided into aliquots among the wells of a 96-well plate. Assembly subpools are amplified with assembly-specific primers. In the figure, assembly-specific primers "2" amplify subpool F2, which consists of ten double-stranded fragments (F2.i-F2.x). The assembly subpool-specific priming sites are removed via a restriction digest, and a polymerase assembles the overlapping fragments into full-length constructs. Last, full-length constructs are amplified by a pair of construction primers.

OLIGONUCLEOTIDE DESIGN AND SYNTHESIS

BASIC PROTOCOL 1

Each construct to be built must be split up into short overlapping fragments. Each fragment must, in turn, be flanked by the assembly- and plate-specific subpool priming sequences, as well as restriction sites for removing the priming sequences (Fig. 1). We have automated these design steps with Biopython scripts (Cock et al., 2009). This software can be either run from a server (*http://synbiosis.med.harvard.edu:8080/gaspserver.rpy*) or as a script on any computer that has the Biopython package installed (current and all

future versions of GASP can be found at *https://bitbucket.org/skosuri/gasp/overview*). The protocol below describes using the server version.

Materials

Computer with a Web browser
List of sequences to be built (in FASTA format)

Setting up the parameters

1. Open http*://synbiosis.med.harvard.edu:8080/gaspserver.rpy* with a browser.
2. Enter your name, email address, and the location of your input file (the FASTA file with a list of genes you want built).
3. The parameter configuration text box will contain the following text:

```
"initialPlateNum": 2,
"RESpacing": [
2,
5,
4
],
"REVector": [
  "BtsI",
"BsmBI",
"BspQI"
],
"SearchForRE": "True",
"REToUse": "BtsI",
"avgoverlapsize": 20,
"deltaGThresholdForOverlaps": -3,
"selfDimersThreshold": 3,
"lengthleeway": 10,
"positionleeway": 10,
"oligoSizeMax": 200,
"seqsToAvoidInOverlapRegions":[],
"skip":[]
```

Change the SearchForRE parameter to False.

The parameters, which are described in detail below, may have to be further adjusted if the DNA will be processed using methods that deviate from the workflow described here.

InitialPlaneNum: 96-well plates of assemblies will be numbered sequentially initiating at this value. This should never be set to 1, as plate #1 is reserved for construction primers.

RESpacing: The distance between the end of the recognition site to the cut location for the enzymes listed in REVector setting. The values should be separated by a comma, and be in the same order as the enzymes in the REVector setting. For example, the RESpacing for BtsI should be set to 2 because the enzyme nicks 2 bp away from its binding site (such that the full recognition site is GCAGTGN_2). Similarly, the recognition site for BsmBI is CGTCTCN_5, so its RESpacing should be set to 5. Therefore, if "REVector": ["BtsI","BsmBI"], then "RESpacing": [2,5]. The site can be left blank if SearchForRE is False.

REVector: List of type IIS restriction enzymes to use for removing the subpool-specific amplification priming sites. The software will attempt to find an enzyme from this list that does not cut each particular construct, starting with the first enzyme listed. In other words, each assembly subpool will be processed using just one of the enzymes

from this list. The enzyme names should be properly capitalized, placed between double quotes, and separated by a comma (e.g., "REVector": ["BtsI","BsmBI"]). The parameter accepts enzymes defined by the Bio.Restriction module of Biopython (for the latest list, see http://www.biopython.org/DIST/docs/api/Bio.Restriction.Restriction-module.html). This can be left blank if SearchForRE is False.

SearchForRE: Set to True if you wish to specify a list of restriction enzymes within the REVector setting. This should be done in applications in which it is impossible to eliminate a single type IIS enzyme cut site from all constructs. Set to False to use the enzyme listed in REToUse to process all the assembly subpools.

REToUse: Restriction enzyme to use if SearchForRE is set to False. This can be left blank if SearchForRE is True.

avgoverlapsize: Each construct will be broken up into assembly oligos that will be fused using a polymerase. The fusion reaction requires priming through overlaps between neighboring oligos. This setting specifies the mean length of the overlap region.

deltaGThresholdForOverlaps: Rejects any overlaps with a secondary structure that has a hybridization free energy less than the value specified (in units of kcal/mol).

selfDimerThreshold: Rejects assembly oligos that have any self-dimerization configurations with a hybridization free energy less than the value specified (arbitrary units).

lengthleeway: Sets allowable variation in the length of the overlap regions.

positionleeway: Sets allowable variation in the assembly oligo junction position. Increasing this value results in a less constrained search space, but increases the computation time and increases variation in synthesized oligonucleotides' lengths.

oligoSizeMax: The maximum oligo size that will be designed. This includes the full-length oligos that include the coding region, the restriction enzyme processing site, and the assembly-specific and plate-specific priming sites. This value should typically be constrained by the commercial synthesis platform used. Note that many of the oligos will be shorter than this maximum value.

seqsToAvoidInOverlapRegion: Specifies positions to be avoided in the overlap between neighboring assembly oligos. This should usually be left blank, but can be used in specialized applications, such as constructing proteins with known repeated regions.

skip: Specifies names of constructs that the algorithm should skip from the design process. This option is used in specialized cases and is normally left blank.

4. Click Submit. An e-mail will be sent to the provided email confirming initiation of the run. Once the run is complete, two more emails should arrive. The first one will contain a report that contains: (1) The sequences to be synthesized on the DNA chip in FASTA format; (2) The plate-specific, position-specific, and construction primers needed to build the set of assemblies; (3) The plate-specific, position-specific, and construction primers that correlate with each individual assembly. The second e-mail will contain a FASTA file that contains the sequences that should be synthesized on the DNA chip.

Ordering DNA

5. Synthesize a DNA chip using the chip oligonucleotide sequences designed by GASP.

   We have validated and extended this protocol using Agilent's Oligo Library Synthesis (OLS) platform (Leproust et al., 2010), though we have also been able to build genes using DNA from LC Sciences (Borovkov et al., 2010) and CustomArray (Liu et al., 2006).

6. Order the plate-specific, position-specific, and construction primers listed in the output email from a commercial vendor (such as Integrated DNA Technologies or Sigma-Aldrich). If the chip has been designed for more than 48 assemblies, it is usually convenient to synthesize the oligos in a 96-well format. If that is the case,

each century of oligos should be synthesized in a separate plate. For instance, skpp-101, skpp-102, skpp-103 . . . , should be synthesized in separate wells of a single plate.

The primers are named with the following format: skpp-#-F/R. The F and R indicate forward and reverse primers, respectively. The numbering scheme assigns the first numbers to plate-specific primers (skpp-1, skpp-2,. . .). If initialPlateNum was set to 2, then the first plate of assembly subpool-specific primers is numbered skpp-201, skpp-202. . ., the second plate is numbered skpp-301, skpp-302 .., and so forth. The oligos numbered in the 100s are the construction primers. Each pair of construction primers assembles one assembly subpool in each plate. Specifically, skpp-101 assembles assembly subpool generated with skpp-201, skpp-301, skpp-401, and so on.

BASIC PROTOCOL 2

PCR AMPLIFICATION OF OLIGONUCLEOTIDE SUBPOOLS

This set of protocols describes going from a pool of off-chip oligonucleotides to 96-well plates containing a different assembly subpool in each well. We have found that while the experimental procedures are technically simple, the logistics can sometimes present a challenge. We highly recommend that each researcher thoroughly understand the logic behind the subpool amplification scheme (as shown in Fig. 2) before starting the procedure. The quantities provided in the protocol are enough to go from an off-chip pool to one plate-subpool, and then from that one subpool to 96 assembly subpools. The volumes used should be scaled in accordance with the actual number of assemblies being built.

Subpool amplification and assembly steps are punctuated with numerous reaction cleanup steps. Their primary purpose is to remove enzymes, unused primers, dNTPs, and salts. At some points of the protocol they have the additional role of concentrating the DNA. Since the protocol necessitates handling 96-well plates of reactions, it practical to use vacuum manifold-driven 96-well column plates. Although not described in this unit, we are currently working integrating the less expensive magnetic bead-based cleanups into our workflow (Hawkins et al., 1994; Rudi et al., 1997; Wiley et al., 2009).

Materials

Array-synthesized library
100× TE buffer (Sigma-Aldirch, cat. no. T9285)
Plate subpool-specific primers (Basic Protocol 1)
Assembly subpool-specific primers (Basic Protocol 1)
UltraPure DNase/RNase-free distilled water (Invitrogen, cat. no. 10977-023)
Phusion high-fidelity DNA polymerase (New England Biolabs, cat. no. M0530L) containing:
  5× Phusion HF Reaction buffer
dNTP solution mix (Enzymatics, cat. no. N205L)
QIAquick 96 PCR purification kit (QIAGEN, cat. no. 28183)
EB buffer (10 mM Tris·Cl, pH 8.5)

Galaxy microcentrifuge with 1.5-ml, 2.0-ml, and 0.2-ml tube adapters (VWR, cat. no. 37000-700)
Vortex mixer (VWR International, cat. no. 58816-121)
0.2-ml PCR tubes with flat caps (BioRad, cat. no. TFI-0201)
Microseal 96-well skirted low-profile PCR plates (BioRad, cat. no. MSP-9601)
96-Reaction thermal cycler (BioRad, cat. no. 186-1096)
Microseal ‘F’ Sealing Foil (BioRad, cat. no. MSF-1001)
QIAvac 96 (QIAGEN, cat. no. 19504)

DNA resuspension and dilution

1. DNA synthesized on Agilent OLS arrays is shipped lyophilized in a microcentrifuge tube. Before opening the tube, centrifuge it for 1 min at maximum speed, room temperature, in a microcentrifuge.

   Using qPCR, we have estimated that an OLS chip with 13,000 130-mers yields 1 pmol of DNA (or approximately 0.1 fmol per oligo species).

2. Add 500 µl 1× TE buffer to the lyophilized DNA. Vortex thoroughly for 5 sec, then briefly centrifuge 5 sec at ∼2000 × *g*, room temperature, to spin down liquid.

3. Make mixes of primers by diluting the appropriate plate and assembly subpool-specific primer pairs to 5 µM in DNase/RNase-free water.

   We recommend first making primer storage stocks by diluting each primer to 100 µM in 1× TE buffer. Note that high concentrations of EDTA in the working stocks may inhibit PCR and other enzymatic reactions.

Plate subpool amplification

4. Working on ice mix the following reagents in 0.2-ml PCR tube for each plate subpool to be amplified:

   33.1 µl distilled water
   10 µl Phusion HF Reaction buffer (5×)
   0.4 µl dNTPs (25 mM each deoxynucleotide)
   5 µl plate subpool-specific primer mix (5 µM each primer)
   1 µl array-synthesized library (from step 2)
   0.5 µl Phusion polymerase (2 U/µl).
   Vortex thoroughly 5 sec, then briefly centrifuge 5 sec at ∼2000 × *g*, room temperature, to spin down liquid.

   If more then one plate subpool will be amplified, then it is convenient to make a common master mix that contains everything except for the subpool-specific primers. The primers should be added once the master mix has been split among the appropriate number of 45-µl aliquots. The volume of the master mix should be in slight excess of how much is needed for the amplifications. For example, if eight plate subpools are to be amplified, an 8.5 reaction 382.5 µl reaction mix should be made. Be sure to vortex the mix thoroughly prior to dividing into aliquots.

5. Place the samples in a thermal cycler and run the following program:

Initial cycle:	30 sec	98°C	(initial denaturation)
25 cycles:	5 sec	98°C	(denaturation)
	10 sec	65°C	(annealing)
	10 sec	72°C	(extension)
1 cycle:	5 min	72°C	(final extension)
Final step:	indefinite	4°C	(hold).

   To minimize sample loss and increase reproducibility either use a thermal cycler with a heated lid or add a few drops of mineral oil (Sigma-Aldrich, cat. no. M8662) to the top of each reaction. Amplified DNA can be stored at 4°C for less than a month, or indefinitely at −20°C.

   At this stage, it is prudent to quantify yield and ensure proper amplification by running 1 to 10 µl of each of the amplified products on an agarose or acrylamide gel. A double-stranded DNA stain such as ethidium bromide should reveal a band for each oligo of a distinct size in the plate subpool.

Assembly subpool amplification

6. Working on ice make the following master mix for each plate subpool amplified:

   11,636 μl distilled water
   4 μl plate subpool amplification reaction products (from step 5)
   4000 μl Phusion HF reaction buffer (5×)
   160 μl dNTPs (25 mM each)
   200 μl Phusion polymerase (2 U/μl).
   Vortex thoroughly 5 sec, then briefly centrifuge 5 sec at ~2000 × *g*, room temperature, to collect liquid.

7. Working on ice, add 160 μl of the mix to each well of a 96-well plate.

   While an ice bucket will suffice for keeping the 96-well plates cold, we have found 96-well coolers (Eppendorf, cat. no. 022510525) to be a convenient alternative. Also, depending on the thermal cycler and plates used, we routinely split the 200-μl reaction into 50 to 100 μl across multiple plates.

8. Add 40 μl of the appropriate primer mix (forward and reverse assembly-specific primers at 5 μM each) to each well of the 96-well plate that contains the master mix aliquots. Cover the wells with foil plate sealer.

9. Place the plate in a thermal cycler and run the following program:

Initial cycle:	30 sec	98°C	(initial denaturation)
30 cycles:	5 sec	98°C	(denaturation)
	10 sec	65°C	(annealing)
	10 sec	72°C	(extension)
1 cycle:	5 min	72°C	(final extension)
Final step:	indefinite	4°C	(hold).

Reaction cleanup

10. Following the manufacturer's instructions, use the QIAquick 96 PCR Purification kit and the QIAvac 96 to bind the DNA from a 96-well assembly amplification to the kit columns. Elute the DNA on each column into 60 μl EB buffer (10 mM Tris·Cl, pH 8.5).

BASIC PROTOCOL 3

ENZYMATIC REMOVAL OF PRIMING SITES

After the second amplification, the subpools still have their subpool-specific priming sites. These must be removed before assemblies can be built. It is for this purpose that each oligo contains a *Bts*I recognition sequence (GCAGTG) between the assembly subpool priming site and the coding region. *Bts*I, like other member of the type IIs family of restriction enzymes, cuts both strands completely outside of its recognition site (Pingoud and Jeltsch, 2001). Consequently, processing with *Bts*I places no sequence restriction on the coding portion of the oligos. Other type IIs restriction enzyme sites may be used by adjusting the design parameters set in Basic Protocol 1.

Materials

UltraPure DNase/RNase-free distilled water (Invitrogen, cat. no. 10977-023)
*Bts*I (New England Biolabs, cat. no. R0614L) containing:
  NEBuffer 4
  BSA
96-well plate containing the assembly subpools (see Basic Protocol 2)
MinElute 96 UF PCR purification kit (QIAGEN, cat. no. 28051)
EB buffer (10 mM Tris·Cl, pH 8.5)

Seal-Rite 1.5-ml microcentrifuge tubes (USA Scientific, cat. no. 1615-5500)
Vortex Mixer (VWR International, cat. no. 58816-121)
Galaxy mini centrifuge with 1.5/2.0 ml and 0.2 ml tube adapters (VWR, cat. no. 37000-700)
QIAvac 96 (QIAGEN, cat. no. 19504)
96-reaction thermal cycler (BioRad, cat. no. 186-1096)

1. For each 96-well plate of cleaned-up assembly subpools to be processed, prepare the following master mix in a 1.5-ml tube:

   145 μl distilled water
   700 μl NEBuffer 4 (10×)
   70 μl BSA (10 μg/μl)
   85 μl *Bts*I (10 U/μl).

   Vortex thoroughly for 5 sec, and then briefly centrifuge 5 sec ~2000 × *g*, room temperature, to spin down liquid.

2. Add 10 μl of the master mix to each well of the 96-well plate containing the assembly subpools. Cover the wells with plate sealer.

3. Heat the plate to 55°C for 2 hr.

4. Following the manufacturer's instructions, use the MinElute 96 UF PCR Purification kit and the QIAvac 96 to clean up each 96-well of *Bts*I-digested assembly subpools. Elute the DNA bound to each column into 15 μl EB buffer

   We highly recommend quantifying the DNA concentration with a spectrophotometer and/or running the samples out on a gel. The total DNA concentration should be 30 to 300 ng/μl. If all of the oligos in a subpool are the same length, then a gel with sufficient resolution should reveal four bands: (1) the properly processed oligos lacking the two priming sites; (2) the oligos cut on only one side; (3) the uncut oligos; and (4) the cut-off pieces containing the priming sites.

ASSEMBLY OF PROCESSED SUBPOOLS INTO FULL-LENGTH CONSTRUCTS

BASIC PROTOCOL 4

Full-length assemblies are built from assembly subpools in two steps—assembly and amplification. The assembly stage consists of 15 thermal cycles with an annealing step consisting of a thermal ramp from 70°C to 50°C. The full-length assemblies are then selected by adding the appropriate construction primers and performing a 25-cycle PCR. In our experience, the more assembly subpool DNA in the assembly reaction, the more likely a full-length product is formed.

Materials

Construction primers
Assembly subpool-specific primers (Basic Protocol 1)
Ice
UltraPure DNase/RNase-free distilled water (Invitrogen, cat. no. 10977-023)
Phusion high-fidelity DNA polymerase (New England Biolabs, cat. no. M0530L) containing:
   HF reaction buffer
dNTP solution mix (Enzymatics, cat. no. N205L)
QIAquick 96 PCR purification kit (QIAGEN, cat. no. 28183)
QIAvac 96 (QIAGEN, cat. no. 19504)

Seal-Rite 1.5-ml microcentrifuge tubes (USA Scientific, cat. no. 1615-5500)
Vortex mixer (VWR International, cat. no. 58816-121)
Galaxy mini centrifuge with 1.5/2.0 ml and 0.2 ml tube adapters (VWR, cat. no. 37000-700)
Microseal 96-well skirted low-profile PCR plates (BioRad, cat. no. MSP-9601)
96-Reaction thermal cycler (BioRad, cat. no. 186-1096)
5-ml tubes

Assembly

1. Mix the forward and reverse construction primer pairs to 5 μM of each primer in water.

   As with the subpool-specific primers, it is a good practice to first make primer storage stocks by diluting each primer to 100 μM in 1× TE buffer.

2. Working on ice, prepare the following master mix for each 96-well plate of assembly subpools in a 1.5-ml tube:

   164 μl distilled water
   400 μl Phusion HF reaction buffer (5×)
   16 μl dNTPs (25 mM each)
   20 μl Phusion polymerase (2 U/μl).

   Vortex thoroughly for 5 sec, then briefly centrifuge 5 sec to spin down liquid.

   We have also successfully used KOD Polymerase (EMD, cat. no. 71085-3) in place of Phusion.

3. Transfer 6 μl of the master mix into each well of a new 96-well PCR plate.

4. Add 14 μl of each assembly subpool to each well of the PCR plate containing the master mix. Cover the wells with plate sealer.

   A total of 14 μl corresponds to 420 ng of assembly subpool DNA at the lowest expected concentration (30 ng/μl). If the concentration has been measured using a spectrophotometer or gel electrophoresis then all subpools can be diluted to 30 ng/μl for maximum reproducibility, although we have not observed higher concentrations inhibiting the assembly reaction.

5. Place the 96-well plate in a thermal cycler and run the following program:

Initial step:	2 min	95°C	(initial denaturation)
15 cycles:	20 sec	95°C	(denaturation)
	1 sec	70°C	
	30 sec	50°C	(cool to 50°C at 0.5°C/sec; annealing)
	20 sec	72°C	(extension)
1 cycle:	5 min	72°C	(final extension)
Final step:	indefinite	4°C	(hold).

   The annealing occurs as the temperature ramps from 70° to 50°C at 0.5°C/sec. The 1-sec step provides the starting point for the ramp.

Assembly amplification

6. Working on ice, prepare the following master mix for each 96-well plate of assembly subpools in a 5-ml tube:

   3.31 ml distilled water
   1 ml Phusion HF reaction buffer (5×)
   40 μl dNTPs (25 mM each)
   50 μl Phusion polymerase (2 U/μl).

Vortex thoroughly for 5 sec, and then briefly centrifuge 5 sec to spin down liquid.

7. Transfer 44 µl of the master mix into each well of a new 96-well plate.
8. For each plate of assemblies, add 198 µl distilled water to a fresh 96-well plate.
9. Dilute the assembly reactions 1:100 by adding 2 µl of the assembly reaction to each well of the PCR plate with water.
10. To each well of the PCR plate containing the amplification master mix, add 5 µl of the appropriate construction primer mix (5 µM each forward and reverse primers) and 1 µl of the 1:100 dilution of the appropriate assembly reaction. Cover the wells with plate sealer.
11. Place the samples in a thermal cycler and run the following program:

1 cycle:	30 sec	98°C	(initial denaturation)
25 cycles:	5 sec	98°C	(denaturation)
	10 sec	62°C	(annealing)
	20 sec	72°C	(extension)
1 cycle:	5 min	72°C	(final extension)
Final step:	indefinite	4°C	(hold).

Reaction cleanup

12. Following the manufacturer's instructions, use the QIAquick 96 PCR Purification kit and the QIAvac 96 to clean up the DNA from the 96-well assembly amplification to QIAquick 96 PCR purification kit column.

GEL-STAB PCR

SUPPORT PROTOCOL

Assembly amplification, particularly that of GC-rich or repetitive constructs, sometimes results in side-products of the wrong size (typically smaller than the construct). If there is substantial yield of the full-length assembly, then it can be purified using agarose gel isolation. However, in rare cases the full-length assembly forms only a faint band on an agarose gel. This protocol describes a simple technique for selectively re-amplifying the correct assembly product from a large background of incorrect assemblies.

Materials

Assembly amplification product (see Basic Protocol 4)
2% E-Gel EX gel (Invitrogen, cat. no. G4020-02)
1-Kb Plus DNA ladder (Invitrogen, cat. no. 10787-018)
UltraPure DNase/RNase-free distilled water (Invitrogen, cat. no. 10977-023)
Appropriate construction primers
Phusion high-fidelity DNA polymerase (New England Biolabs, cat. no. M0530L) containing:
HF reaction buffer
dNTP solution mix (Enzymatics, cat. no. N205L)
QIAquick PCR purification kit (QIAGEN, cat. no. 28104)

E-Gel iBase power system (Invitrogen, cat. no. G6400)
Thin metal edge or E-Gel opener (Invitrogen, cat. no. G5300-01)
Safe Imager 2.0 blue light transilluminator (Invitrogen, cat. no. G6600)
Seal-Rite 1.5-ml microcentrifuge tubes (USA Scientific, cat. no. 1615-5500)
10- or 20-µl pipet tips
0.2-ml PCR tubes with flat caps (BioRad, cat. no. TFI-0201)
Galaxy mini centrifuge with 1.5/2.0-ml and 0.2-ml tube adapters (VWR, cat. no. 37000-700)

Vortex mixer (VWR International, cat. no. 58816-121)
96-Reaction thermal cycler (BioRad, cat. no. 186-1096)

1. Following the manufacturer's instructions, load 1 to 5 μl of the assembly amplification product and 5 μl of the 1-Kb Plus ladder in separate wells of a 2% E-Gel EX and run on the E-Gel base.

   Traditional agarose gels work, as well as the precast gels, but the amount of DNA loaded may have to be adjusted to account for the differences in gel geometry and the dye used.

2. Using a thin metal edge or an E-Gel opener carefully remove the top plastic cover of the gel cassette.

3. Place the gel on a blue light transilluminator. Add 20 μl water into a 1.5-ml tube. Stab the band that corresponds to the desired assembly product with a sterile 10- or 20-μl pipet tip. Look at the tip to make sure a small slice of the gel is stuck in the tip. Pipet up and down in the water using the same pipet tip. Look at the tip again to make sure that the gel that was stuck there is gone.

4. Heat the water containing the gel slice to 65°C for 15 to 30 min.

5. Combine the following in a 0.2-ml PCR tube:

   4.6 μl distilled water
   1 μl water containing the gel isolate
   2 μl appropriate construction primers (5 μM each primer)
   4 μl Phusion HF Reaction Buffer (5×)
   0.2 μl dNTPs (25 mM each)
   0.2 μl Phusion polymerase (2 U/μl).

   Vortex thoroughly for 5 sec, and then briefly centrifuge for 5 sec at room temperature, to spin down liquid.

6. Place the samples in a thermal cycler and run the following program:

1 cycle:	30 sec	98°C	(initial denaturation)
30 cycles:	5 sec	98°C	(denaturation)
	10 sec	62°C	(annealing)
	20 sec	72°C	(extension)
1 cycle:	5 min	72°C	(final extension)
Final step:	indefinite	4°C	(hold).

7. Cleanup the reaction using a QIAquick kit following the manufacturer's instructions.

COMMENTARY

Background Information

There is a large and fairly mature set of techniques for building genetic constructs without relying on a pre-existing template, including methods for chemically synthesizing short oligonucleotides (Michelson and Todd, 1955; Letsinger and Mahadevan, 1965; Brown, 1993), for using enzymes to fuse the oligonucleotides into long double-stranded fragments (Agarwal et al., 1970; Rossi et al., 1982; Li and Elledge, 2007; Gibson et al., 2009), and for performing in vivo assembly of genome-scale fragments in recombinogenic organisms such as *S. cerevisiae* (Gibson, 2009; Shao et al., 2009). Unfortunately, high costs hinder the widespread adoption of de novo synthesis, primary among them the cost of chemically synthesizing single-stranded oligonucleotides. The maturation of a number of technologies that allow parallel synthesis of thousands or millions of oligonucleotide on solid surfaces (DNA chips) has reduced oligonucleotide cost more than 100-fold over the past decade (Carlson, 2003; Carr and

Church, 2009). Contrary to the expectations of many, and despite a number of proof-of-concept publications (Tian et al., 2004; Zhou et al., 2004; Binkowski et al., 2005), the advent of commercially available DNA chips did not catalyze a widespread drop in the cost of synthesizing double-stranded gene-length fragments.

The three factors that until recently inhibited the adoption of DNA off chips for gene synthesis were the high chemical complexity of the resulting DNA pool, the oligonucleotide error rates, and the low synthesis yields. Chemical complexity refers to the fact that while tens of thousands of oligonucleotide species are synthesized in parallel on the same surface, only 5 to 50 are needed to build any particular gene-sized double-stranded fragment. At typical DNA chip synthesis scales, performing enzymatic assembly of a subset of oligonucletides in the presence of the unrelated species is difficult or impossible (Borovkov et al., 2010). Physical separation of DNA off a chip into subsets of oligonucleotides has recently been achieved using microfluidics (Zhou et al., 2004; Lee et al., 2010; Quan et al., 2011), beads in picotiter plates (Matzas et al., 2010), and barcode PCR (Kosuri et al., 2010). The second problem, high error rates, is important because having more errors raises costs by increasing the amount of screening needed to isolate error-free constructs. Off-chip oligonucleotides typically have an error rate of 1/25 to 1/50 errors/bases, while low-throughput synthesis on controlled pore glass beads results in an error rate of 1/200 to 1/500. Here, too, there has been much recent progress, including improvements in the chemical synthesis process (Leproust et al., 2010), advances in enzyme-mediated error depletion (Carr et al., 2004; Binkowski et al., 2005; Fuhrmann et al., 2005; Bang and Church, 2007; Kosuri et al., 2010; Quan et al., 2011), and the integration of high-throughput sequencing and synthesis for direct error screening (Matzas et al., 2010). The last challenge has been the low yields achieved with on-chip synthesis, which make it difficult to assemble the synthesized oligonucleotides into gene-length fragments. The microfluidic-based methods described above solve this problem by increasing the DNA concentration through low assembly reaction volumes. In contrast, our subpool-based technique achieves high concentrations through selective PCR amplification.

The workflow we presented in this set of protocols relies on the ability to PCR amplify a small subset of oligonucleotides out of a large background of DNA. To prevent hybridization of primers to the wrong priming sites we designed a set of 3000 orthogonal primer pairs (Kosuri et al., 2010). The method, in brief, was as follows: first, we searched the set of 240,000 orthogonal 25-mers designed by Xu et al. (2009) for 20-base windows that lack commonly used restriction enzyme sites and have a melting temperature of 60° to 64°C (SantaLucia, 1998; SantaLucia and Hicks, 2004). The resulting sequences were culled of any pairs that cross-hybridize using a modified version of the AutoDimer tool, a library-on-library BLAT search, and, finally, a BLAST-based network elimination algorithm (Kent, 2002; Vallone and Butler, 2004; Xu et al., 2009). Next, we used UNAFold to select only the sequences with a secondary structure with a $\Delta G \leq -2$ (Markham and Zuker, 2008). The selected sequences were clustered using ClustalW2, and 6000 sequences were selected such that they were maximally spread out on the resulting phylogenetic tree (Larkin et al., 2007). Finally, 3000 primer pairs were generated by matching the sequences on their melting temperature and propensity to form primer-dimers.

Critical Parameters

DNA synthesis

A number of manufacturers sell custom DNA microarrays, and differences between their synthesis methodologies result in differences in the oligo error rates. The error rate affects two key aspects of the gene synthesis process. First, a high error rate means a larger fraction of the synthesized genes will have an error. This directly increases the cost, as constructs must be screened by low-throughput sequencing until a perfect clone is found. A high error rate also makes it more difficult to synthesize longer oligos. Longer oligos are advantageous, as they reduce the complexity of the synthesis reaction by reducing the number of individual DNA pieces that must be put together in order to build a gene of a particular size. Most of our experience with the set of protocols provided in this unit comes from working with Agilent Technologies' Oligo Library Synthesis (OLS) platform, which can synthesize oligonucleotide 100- to 200-bp in length with an error rate on the order of 1/500 errors/bp (Kosuri et al., 2010; Leproust et al., 2010). Because the OLS platform is still under development, OLS pools are not yet widely sold. However, laboratories can gain access to them by signing a material transfer agreement

with Agilent Technologies. The protocols we describe should be applicable to all other array platforms, although if the DNA chip has a high error, then error correction may be necessary. Similarly, if the off-chip oligonucleotides are shorter than 120 bases, then different DNA cleanup steps may have to be used, as the QIAGEN columns used in the protocols do not recover fragments shorter than 40 bp.

It is useful to keep in mind that DNA from microarrays is one of the cheapest reagents in the synthesis process. Therefore, in the interest of rapid prototyping, it often makes sense to waste synthesis capacity rather than trying to design enough constructs to fill up all features on the chip. Unfortunately, subpool-specific and construction primers are expensive because they must be synthesized using low-throughput methods. However, primer synthesis yields are typically in excess of what is needed to process a single set of assemblies, so the primers can be used to process multiple DNA chips.

Construct design

If the constructs being built encode proteins, then there are many possible ways to reverse-translate the amino acid sequence (Welch et al., 2009b). One consideration when choosing among synonymous codons is the natural codon bias within the genome of organism that will express the protein (Gustafsson et al., 2004). Similarly, studies in *Escherichia coli* revealed that using codons that are most highly charged during amino acid starvation leads to significantly higher levels of protein expression (Welch et al., 2009a). The second variable to take into account when designing constructs is the propensity of the resulting DNA to form secondary structures that interfere with polymerase-based assembly methods. Despite our efforts to optimize the robustness of our assembly protocol, the reaction will fail often if the construct being built is highly GC-rich or repetitive. For designing protein-coding DNA sequences, we recommend GeneDesign, a Web-based sequence manipulation environment that allows facile manipulation of codon usage, GC-bias, restriction site placement, and other design parameters (Richardson et al., 2010).

Downstream processing

An assembly reaction produces a heterogeneous population of molecules, some of which contain point mutations or structural rearrangements, such as large deletions or insertions. If error-free constructs are needed, then the assemblies must be cloned and screened by sequencing. If the constructs are to be used in a screen or selection, then it may be possible to use the heterogeneous assembly products directly. Error rates can be reduced by creating mismatches by melting and reannealing the assemblies and depleting heteroduplexes by a MutS pulldown (Carr et al., 2004; Binkowski et al., 2005) or cleavage with resolvases (Fuhrmann et al., 2005; Bang and Church, 2007). Assemblies with frame-shift mutations can be screened for by cloning them in-frame to a downstream *lacZ* fragment or fluorescent protein (Cronan et al., 1988; Kim et al., 2010). Error rates can be estimated by building control fluorescent protein assemblies, which allow the error rate to be estimated by counting colonies or performing flow cytometry. Such protein function assays give an accurate error rate estimate if most errors are deletions and insertions. However, due to the degeneracy of the genetic code and the stability of protein function in the face of some residue substitutions, function-based screens underestimate the error rate if there are many transitions and transversions.

Troubleshooting

No assembly products or products that are too short

Assembly of GC-rich or repetitive sequences sometimes proves challenging. We have found that most of the time this problem can be solved by increasing the amount of processed assembly subpool DNA added to the assembly reaction. At 5 to 20 ng/μl, the highest assembly subpool concentrations tested, we were able to get a correctly sized band for ~95% of the difficult templates tested (Kosuri et al., 2010). Assembly reactions can be further optimized by introducing PCR additives that increase specificity and reduce DNA secondary structure, such as betaine and dimethyl sulfoxide (DMSO; Winship, 1989; Henke et al., 1997). Similarly, varying annealing temperature or the concentrations of magnesium ions and deoxynucleotides may alter PCR performance (Cobb and Clarkson, 1994; Roux, 1995).

Assembly results in short side-products

Gel-isolate the band of the correct size and re-amplify using the appropriate construction primers (see Support Protocol).

Assembly results in high-molecular-weight smears

This is typically caused by "over-amplifying" the template, which happens

when PCR primers run out, causing the amplicon molecules to generate long fusions by mispriming off each other (Bell and DeMarini, 1991). This can be avoided by decreasing the number of amplification cycles, increasing the concentration of construction primers, decreasing the amount of post-assembly DNA added to the assembly amplification reaction, shortening extension times, or optimizing PCR specificity as described in the section on troubleshooting a lack of assembly products.

Synthesized genes have high error rates

Errors can be introduced during the chemical synthesis of DNA on a chip or at any of the subsequent enzymatic steps. Sequencing at various points of the synthesis workflow may help detect the processes that introduce the most error. Error detection or depletion techniques can be introduced after the assembly amplification step, some of which are described in greater detail within the *Downstream processing* section of Critical Parameters.

Loss of DNA during reaction cleanups

Our protocols rely on using QIAGEN's column-based PCR cleanups that have reduced recovery of double-stranded DNA shorter than 100 bp, and no recovery of DNA shorter than 40 bp. It is important to keep track of how the DNA length changes at each step of the protocol, as it decreases after both the assembly subpool amplification and the *Bts*I digest. If DNA is lost during cleanups in a size-dependent manner, then it is best to switch to either a different column-based cleanup kit or to use alternative methods such as magnetic beads (Hawkins et al., 1994; Rudi et al., 1997; Wiley et al., 2009).

Anticipated Results

Basic Protocol 1 will generate a FASTA file with the sequences of the oligonucleotides that need to be synthesized on the DNA chip. After the oligonucleotides have been synthesized and cleaved off the array surface, they will be amplified first into plate subpool and then into assembly subpools, as described in Basic Protocol 2. Running each sample out on an agarose or a polyacrylamide gel should result only bands that correspond to the expected lengths of the DNA species in each subpool. Note assembly subpools will be shorter than the plate subpools because plate-specific priming sites flank the assembly-specific priming sites. Basic Protocol 3 will result in the removal of the assembly-specific priming sites. Gel electrophoresis will reveal bands corresponding to the cleaved-off priming sites, the members of the assembly subpool with the priming sites removed, and partially digested DNA cut on just one of the two sides. Basic Protocol 4 should result in full-length assemblies, some of which may have to be gel-isolated or, if the assembly amplification yields were low, re-amplified from a gel-isolated sample. The constructs are ready to be used for downstream applications once they produce a single band of the expected size when resolved by gel electrophoresis.

Time Considerations

Going from a list of sequences to be built to a chip design should take a couple of hours. Once both the DNA chip and the amplification primers have been synthesized, the synthesis process should take 1 to 2 days, though assembling a large number of constructs will increase the amount of time spent setting up reactions (and, conversely, automation will reduce it significantly). Post-assembly gel-isolation and re-amplification typically add an additional day of work.

Acknowledgements

This work was supported by the U.S. Office of Naval Research (N000141010144), National Human Genome Research Institute Center for Excellence in Genomics Science (P50 HG003170), and the Department of Energy Genomes to Life grant (DE-FG02-02ER63445) (all to G.M.C.). Also, we thank Jeff Sampson from Agilent Technologies for helpful discussions related to process development and oligonucleotide libraries.

Literature Cited

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., Sgarawella, H., and Yamada, T. 1970. Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast. *Nature* 227:27-34.

Bang, D. and Church, G.M. 2007. Gene synthesis by circular assembly amplification. *Nat. Methods* 5:37-39.

Bayer, T.S., Widmaier, D.M., Temme, K., Mirsky, E.A., Santi, D.V., and Voigt, C.A. 2009. Synthesis of methyl halides from biomass using engineered microbes. *J. Am. Chem. Soc.* 131:6508-6515.

Bell, D.A. and DeMarini, D.M. 1991. Excessive cycling converts PCR products to random-length higher molecular weight fragments. *Nucleic Acids Res.* 19:5079.

Binkowski, B.F., Richmond, K.E., Kaysen, J., Sussman,, M.R., and Belshaw, P.J. 2005.

Correcting errors in synthetic DNA through consensus shuffling. *Nucleic Acids Res.* 33:e55.

Borovkov, A.Y., Loskutov, A.V., Robida, M.D., Day, K.M., Cano, J.A., Le Olson, T., Patel, H., Brown, K., Hunter, P.D., and Sykes, K.F. 2010. High-quality gene assembly directly from unpurified mixtures of microarray-synthesized oligonucleotides. *Nucleic Acids Res.* 38:e180.

Brown, D.M. 1993. A brief history of oligonucleotide synthesis. *In* Methods in Molecular Biology, Vol. 20: Protocols for Oligonucleotides and Analogs (S. Agrawal, ed.) pp. 1-20. Humana Press, Totowa, N.J.

Carlson, R. 2003. The pace and proliferation of biological technologies. *Biosecur. Bioterror.* 1:203-214.

Carr, P. A. and Church, G.M. 2009. Genome engineering. *Nat. Biotechnol.* 27:1151-1162.

Carr, P.A., Park, J.S., Lee, Y.-J., Yu, T., Zhang, S., and Jacobson, J.M. 2004. Protein-mediated error correction for de novo DNA synthesis. *Nucleic Acids Res.* 32:e162.

Cobb, B.D. and Clarkson, J.M. 1994. A simple procedure for optimizing the polymerase chain reaction (PCR) using modified Taguchi methods. *Nucleic Acids Res.* 22:3801-3805.

Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., and de Hood, M.J.L2009. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25:1422-1423.

Cronan, J.E. Jr., Narasimhan, M.L., and Rawlings, M. 1988. Insertional restoration of β-galactosidase α-complementation (white-to-blue colony screening) facilitates assembly of synthetic genes. *Gene* 15:161-170.

Fuhrmann, M., Oertel, W., Berthold, P., and Hegemann, P. 2005. Removal of mismatched bases from synthetic genes by enzymatic mismatch cleavage. *Nucleic Acids Res.* 33:e58.

Gibson, D.G. 2009. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic Acids Res.* 37:6984-6990.

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, C.A. III, and Smith, H.O. 2008. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 319:1215-1220.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6:343-345.

Gustafsson, C., Govindarajan, S., and Minshull, J. 2004. Codon bias and heterologous protein expression. *Trends Biotechnol.* 22:346-353.

Hanai, T., Atsumi, S., and Liao, J.C. 2007. Engineered synthetic pathway for isopropanol production in *Escherichia coli*. *Appl. Environ. Microb.* 73:7814-7818.

Hawkins, T.L., O'Connor-Morin, T., Roy, A., and Santillan, C. 1994. DNA purification and isolation using a solid-phase. *Nucleic Acids Res.* 22:4543-4544.

Henke, W., Herdel, K., Jung, K., Schnorr, D., and Loening, S.A. 1997. Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res.* 25:3957-3958.

Kent, W.J. 2002. BLAT—The BLAST-like alignment tool. *Genome Res.* 12:656-664.

Kim, H., Han, H., Shin, D., and Bang, D. 2010. A fluorescence selection method for accurate large-gene synthesis. *Chembiochem* 11:2448-2452.

Kosuri, S., Eroshenko, N., Leproust, E.M., Super, M., Way, J., Li, J.B., and Church, G.M. 2010. Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nat. Biotechnol.* 28:1295-1299.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and Higgins, D.G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947-2948.

Lee, C.C., Snyder, T.M., and Quake, S.R. 2010. A microfluidic oligonucleotide synthesizer. *Nucleic Acids Res.* 38:2514-2521.

Leproust, E.M., Peck, B.J., Spirin, K., McCuen, H.B., Moore, B., Namsaraev, E., and Caruthers, M.H. 2010. Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process. *Nucleic Acids Res.* 38:2522-2540.

Letsinger, R.L. and Mahadevan, V. 1965. Oligonucleotide synthesis on a polymer support. *J. Am. Chem. Soc.* 87:3526-3527.

Li, M.Z. and Elledge, S.J. 2007. Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nat. Methods* 4:251-256.

Liu, R.H., Munro, S.B., Nguyen, T., Siuda, T., Suciu, D., Bizak, M., Slota, M., Fuji, H.S., Danley, D., and McShea, A. 2006. Integrated microfluidic CustomArray device for bacterial genotyping and identification. *J. Assoc. Lab. Autom.* 11:360-367.

Markham, N.R. and Zuker, M. 2008. UNAFold: Software for nucleic acid folding and hybridization. *In* Bioinformatics, Vol. 2: Structure, Function and Applications, Vol. 453 (J.M. Keith, ed.) pp. 3-31. Humana Press, Totowa, N.J.

Matzas, M., Stähler, P.F., Kefer, N., Siebelt, N., Boisguérin, V., Leonard, J.T., Keller, A., Stähler, C.F., Häberle, P., Gharizadeh, B., Babrzadeh, F., and Church, G.M. 2010. High-fidelity gene synthesis by retrieval of sequence-verified DNA identified using high-throughput pyrosequencing. *Nat. Biotechnol.* 28:1291-1294.

Michelson, A.M. and Todd, A.R. 1955. Nucleotides part XXXII. Synthesis of a dithymidine dinucleotide containing a 3′: 5′-internucleotidic linkage. *J. Chem. Soc.* 2632-2638.

Nirenberg, M.W. and Matthaei, J.H. 1961. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 47:1588-1602.

Pingoud, A. and Jeltsch, A. 2001. Structure and function of type II restriction endonucleases. *Nucleic Acids Res.* 29:3705-3727.

Quan, J., Saaem, I., Tang, N., Ma, S., Negre, N., Gong, H., White, K.P., and Tian, J. 2011. Parallel on-chip gene synthesis and application to optimization of protein expression. *Nat. Biotechnol.* 29:449-452.

Richardson, S.M., Nunley, P.W., Yarrington, R.M., Boeke, J.D., and Bader, J.S. 2010. GeneDesign 3.0 is an updated synthetic biology toolkit. *Nucleic Acids Res.* 38:2603-2606.

Ro, D.-K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C.Y, Withers, S.T., Shiba, Y., Sarpong, R., and Keasling, J.D. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440:940-943.

Rossi, J.J., Kierzek, R., Huang, T., Walker, P.A., and Itakura, K. 1982. An alternate method for synthesis of double-stranded DNA segments. *J. Biol.Chem.* 257:9226-9229.

Roux, K.H. 1995. Optimization and troubleshooting in PCR. *Genome Res.* 4:S185-S194.

Rudi, K., Kroken, M., Dahlberg, O.J., Deggerdal, A., Jakobsen, K.S., and Larsen, F. 1997. Rapid, universal method to isolate PCR-ready DNA using magnetic beads. *BioTechniques* 22:506-511.

SantaLucia, J. Jr. 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.* 95:1460-1465.

SantaLucia, J. Jr. and Hicks, D. 2004. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* 33:415-440.

Shao, Z., Zhao, H., and Zhao, H. 2009. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* 37:e16.

Tian, J., Gong, H., Sheng, N., Zhou, X., Gulari, E., Gao, X., and Church, G. 2004. Accurate multiplex gene synthesis from programmable DNA microchips. *Nature* 432:1050-1054.

Vallone, P.M. and Butler, J.M. 2004. AutoDimer: A screening tool for primer-dimer and hairpin structures. *BioTechniques* 37:226-231.

Welch, M., Govindarajan, S., Ness, J.E., Villalobos, A., Gurney, A., Minshull, J., and Gustafsson, C. 2009a. Design parameters to control synthetic gene expression in *Escherichia coli*. *PLoS One* 4:e7002.

Welch, M., Villabos, A., Gustafsson, C., and Minshull, J. 2009b. You're one in a googol: Optimizing genes for protein expression. *J. R. Soc. Interface* 6:S467-S476.

Wiley, G., Macmil, S., Qu, C., Wang, P., Xing, Y., White, D., Li, J., White, J.D., Domingo, A., and Roe, B.A. 2009. Methods for generating shotgun and mixed shotgun/paired-end libraries for the 454 DNA sequencer. *Curr Protoc. Hum. Genet.* 61:18.1.1-18.1.21.

Winship, P.R. 1989. An improved method for directly sequencing PCR amplified material using dimethyl sulfoxide. *Nucleic Acids Res.* 17:1266.

Xu, Q., Schlabach, M.R., Hannon, G.J., and Elledge, S.J. 2009. Design of 240,000 orthogonal 25mer DNA barcode probes. *Proc. Natl. Acad. Sci. U.S.A.* 106:2289-2294.

Zhou, X., Cai, S., Hong, A., You, Q., Yu, P., Sheng, N., Srivannavit, O., Maranjan, S., Rouillard, J.M., Xia, Y., Zhang, X., Xiang, Q., Ganesh, R., Zhu, Q., Matejko, A., Gulari, E., and Gao, X. 2004. Microfluidic PicoArray synthesis of oligodeoxynucleotides and simultaneous assembling of multiple DNA sequences. *Nucleic Acids Res.* 32:5409-5417.