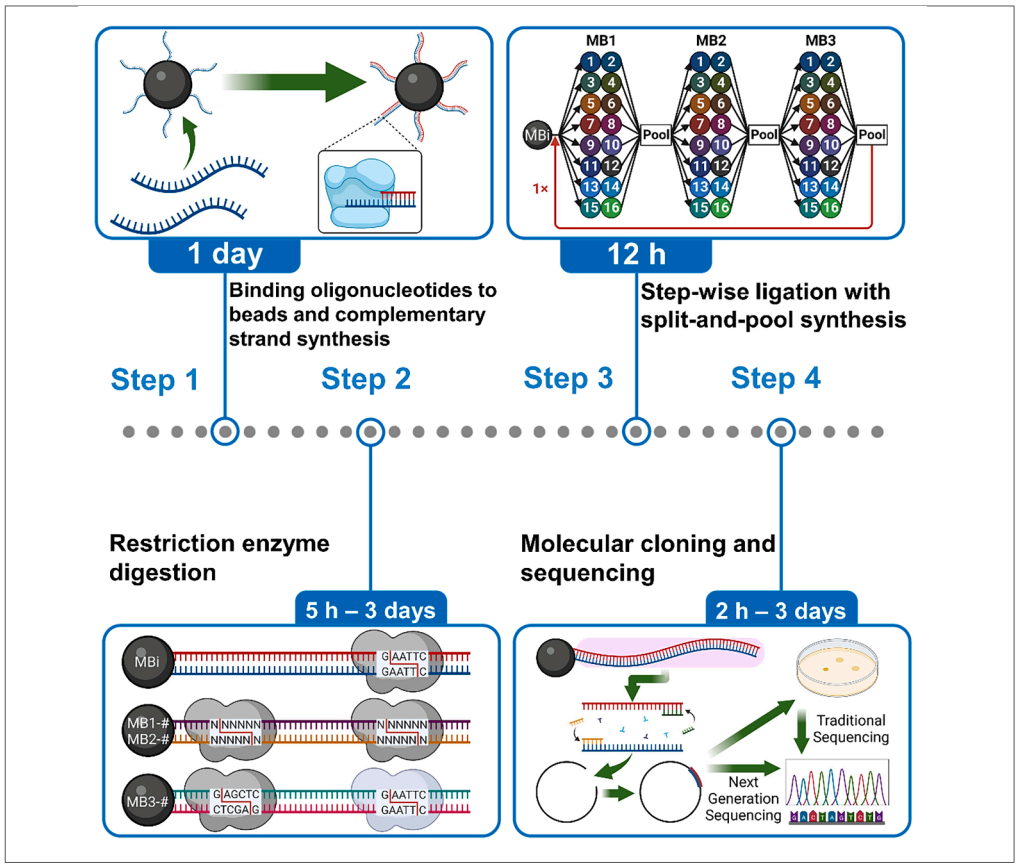
STAR Protocols

CellPress
OPEN ACCESS

Protocol

Split-and-pool synthesis to generate scalable combinatorial oligonucleotide libraries on magnetic nanoparticles



The generation of combinatorial oligonucleotide libraries is desirable for applications such as DNA aptamers, data storage, DNA origami, or synthetic genomes, but conventional libraries present challenges in detection and analysis. Synthesis of unique oligonucleotide sequences on magnetic nanoparticles would enhance the ability to manipulate, recover, and detect them. Here, we present a protocol for generating a scalable combinatorial oligonucleotide library on magnetic nanoparticles using split-and-pool synthesis. We then describe the process for preparing the library for conventional and next-generation sequencing (NGS) DNA sequencing.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

John V.L. Nguyen, Lidija Malic, Christina Nassif, Dillon Da Fonte, Maryam Tabrizian

lidija.malic@cnrc-nrc.gc.ca (L.M.)
maryam.tabrizian@mcgill.ca (M.T.)

Highlights

Combinatorial oligonucleotide library for DNA aptamers, DNA origami, or synthetic genomes

Synthesis of unique oligonucleotide sequences on magnetic nanoparticles

Enhanced manipulation, recovery, and detection of oligonucleotide

Split-and-pool synthesis to generate an ssDNA library for further analysis or sequencing

Nguyen et al., STAR Protocols 6, 103859
June 20, 2025
https://doi.org/10.1016/j.xpro.2025.103859

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Split-and-pool synthesis to generate scalable combinatorial oligonucleotide libraries on magnetic nanoparticles

John V.L. Nguyen,[1] Lidija Malic,[1,2,4,*] Christina Nassif,[2] Dillon Da Fonte,[2] and Maryam Tabrizian[1,3,5,*]

[1]Department of Biomedical Engineering, McGill University, 3775 University Street, Montreal, QC H3A 2B4, Canada

[2]Medical Devices, Life Sciences Division, National Research Council of Canada, 75 de Mortagne Boulevard, Boucherville, QC J4B 6Y4, Canada

[3]Faculty of Dental Medicine and Oral Health Sciences, McGill University, 2001 McGill College Avenue, Montreal, QC H3A 1G1, Canada

[4]Technical contact

[5]Lead contact

*Correspondence: lidija.malic@cnrc-nrc.gc.ca (L.M.), maryam.tabrizian@mcgill.ca (M.T.)
https://doi.org/10.1016/j.xpro.2025.103859

SUMMARY

The generation of combinatorial oligonucleotide libraries is desirable for applications such as DNA aptamers, data storage, DNA origami, or synthetic genomes, but conventional libraries present challenges in detection and analysis. Synthesis of unique oligonucleotide sequences on magnetic nanoparticles would enhance the ability to manipulate, recover, and detect them. Here, we present a protocol for generating a scalable combinatorial oligonucleotide library on magnetic nanoparticles using split-and-pool synthesis. We then describe the process for preparing the library for conventional and next-generation sequencing (NGS) DNA sequencing.

For complete details on the use and execution of this protocol, please refer to Nguyen et al.[1]

BEFORE YOU BEGIN

The protocol below describes the specific steps for synthesizing an ssDNA library on streptavidin nanoparticles (or beads) with a total theoretical library size of 16^6 unique sequences, with each bead containing unique sequences reaching 143 nucleotides (nt) in length. We have also tested the uptake of ssDNA used in this protocol with oligo$(dT)_{25}$ beads and noted a lower level of ssDNA binding. The library synthesis may be scaled to produce either fewer or greater unique sequences, with the length of individual sequences correlating to the size of the library generated. This protocol also includes a procedure for the preparation of samples for next-generation sequencing (NGS). The samples produced from this protocol will generate an ssDNA library directly ready for further physicochemical analysis or sequencing. To achieve this, we describe the protocol in four major steps, (1) binding oligonucleotides to beads and complementary strand synthesis, (2) restriction enzyme digestion, (3) step-wise ligation with split-and-pool synthesis, and (4) molecular cloning and sequencing. The library is synthesized in major steps 1-3, and the final steps of preparing the library synthesis products for NGS are described in step 4. Library synthesis is facilitated by coupling DNA to magnetic beads (MB), with the initial MBi fragment constant, and all subsequent fragments containing variable regions. Three sets of fragments, MB1, MB2, and MB3, each contain a pair of restriction sites, permitting sequential addition of additional fragments, with the final MB3t fragment containing an undigested restriction site, terminating the strand. With six ligation steps, the theoretical final strand should have a sequence of MBi+MB1+MB2+MB3+MB1+MB2+MB3t. The MBi fragment

remains coupled to the bead from start to finish, while all subsequent MB1, MB2, MB3, and MB3t fragments start coupled to the beads and are later digested and unbound from the beads to be ligated onto the growing strands.

The significant benefit of this protocol compared to conjugating a premade library onto beads is that each bead produced from this protocol will have multiple copies of one unique sequence on the bead surface. In contrast, a premade library that is later conjugated to beads will have multiple copies of different sequences on the bead surface. This may lead to difficulty in analyzing beads and oligonucleotides, due to an uncertainty in which particular oligonucleotide on a bead surface may be of interest.

Design of fragments

Timing: 7 days

1. Consider the nucleotide-bead coupling method, biotin-streptavidin coupling or complementary poly(A)–oligo$(dT)_{25}$ binding.

   Note: A 5′ biotin modification is needed if using biotin–streptavidin coupling and a 25 nt poly (A) sequence is needed for poly(A)–oligo$(dT)_{25}$ binding. Both of these are present in the Library ssDNA Fragments list presented in Table S1.

   Note: The final ssDNA bound to the nanoparticle will be the biotinylated strand if using biotin-streptavidin coupling or the oligo$(dT)_{25}$ strand if using poly(A)–oligo$(dT)_{25}$ binding.

2. Customize the nucleotide sequences or keep them as described in this protocol.
   a. The maximum number of variable regions in each fragment (changes the maximum diversity with each step). The variable regions are the group of nucleotide base positions that can be changed with the incorporation of each unique fragment. In our protocol, as described, we perform six ligations and thus have a theoretical number of six variable regions per sequence.
   b. Any other modifications (e.g., fluorescent labeling).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Deoxynucleotide (dNTP) solution mix	New England Biolabs	Cat# N0447L
EcoRI-HF	New England Biolabs	Cat# R3101L
EcoRV-HF	New England Biolabs	Cat# R3195L
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich	Cat# E6758-100G
HindIII-HF	New England Biolabs	Cat# R3104L
Hydrochloric acid	Supelco	Cat# HX0603-4
Q5 high-fidelity DNA polymerase	New England Biolabs	Cat# M0491L
Q5 reaction buffer pack	New England Biolabs	Cat# B9027S
SacI-HF	New England Biolabs	Cat# R3156L
Tris (base), ultra pure	BioShop Canada	Cat# TRS001.500
Critical commercial assays		
Monarch PCR & DNA cleanup kit	New England Biolabs	Cat# T1030L
Quick ligation kit	New England Biolabs	Cat# M2200L
Zero Blunt PCR cloning kit	Invitrogen	Cat# K2700-40
Oligonucleotides		
Library ssDNA fragments	See Table S1	N/A
Forward primer: CATCTGACCTCTGTGCTGCT	Integrated DNA Technologies; Oh et al.[2]	N/A

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Reverse primer: TTCACGGTAGCACGCATAGG	Integrated DNA Technologies; Oh et al.[2]	N/A
Software and algorithms		
Ligation Calculator	New England Biolabs	https://nebiocalculator.neb.com/#!/ligation
mfold	Zuker[3]	http://www.unafold.org/mfold/applications/dna-folding-form.php
Other		
CFX Opus 96 real-time PCR system	Bio-Rad Laboratories	Cat# 12011319
DNA LoBind microcentrifuge tubes 1.5 mL	Eppendorf	Cat# 022431021
DynaMag-2 magnet	Invitrogen	Cat# 12321D
myBlock II dual chamber, with 2× BSWCMB	Benchmark Scientific	Cat# BSH5002-2B
NanoDrop One^C	Thermo Fisher Scientific	Cat# ND-ONEC-W
SuperMag streptavidin beads, 100 nm	Ocean NanoTech	Cat# SV0100-05
Dynabeads Oligo(dT)$_{25}$	Invitrogen	Cat# 61005

MATERIALS AND EQUIPMENT

Wash Buffer		
Reagent	Final concentration	Amount
Tris base	20 mM	1.21 g
Ultrapure water	N/A	As required
EDTA	2 mM	0.29 g
HCl	N/A	As required
Total	N/A	**500 mL**

Store the solution at 4°C for a maximum of 12 months.

Mix 1.21 g tris base, 400 mL ultrapure water, and 0.29 g EDTA. Slowly add a bit of the HCl to the solution while mixing, as EDTA is most soluble between pH 7.5 to 8.0 and mix until all solid ingredients have dissolved. Adjust the final solution to pH 7.5 and add ultrapure water until the total solution volume is 500 mL. Filter the solution through a 0.2 μm filter.

⚠ CRITICAL: Concentrated HCl is volatile and toxic if inhaled. Work with HCl should be performed in a fume hood. Appropriate personal protective equipment should be worn.

STEP-BY-STEP METHOD DETAILS

Binding oligonucleotides to beads and complementary strand synthesis

Timing: 1 day

This major step describes the conjugation of ssDNA fragments to beads and the synthesis of a complementary strand to form dsDNA on beads. One tube of MBi, 32 tubes of MB1-#, 32 tubes of MB2-#, and 32 tubes of MB3-#. The conjugation of the ssDNA and the synthesis of the complementary strand should be as shown in Figure 1. For each MB1-#, MB2-#, and MB3-# oligonucleotide fragment, there are 16 different iterations: MB1-1 to MB1-16, MB2-1 to MB2-16, MB3-1 to MB3-16. For simplicity these are denoted as MB1-#, MB2-#, and MB3-# in the text, and the differences between MB1-#, MB2-#, MB3-#, and MB3t-# are as shown in Figure 2. There exists only one MBi, and it is constant among all sequences produced.

Note: At the end of this major step, a total of 97 tubes, containing 49 unique sequences bound to beads will be obtained.

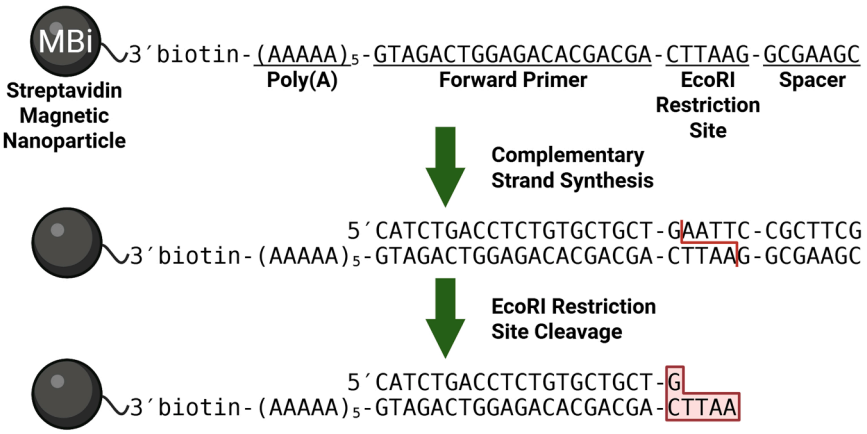


Figure 1. Synthesis of the complementary strand and cleavage by EcoRI of MBi

MBi sequence is coupled to the bead by a streptavidin-biotin interaction. The complementary strand is synthesized by Q5 polymerase (not shown) to facilitate the double-stranded digestion by EcoRI. Red highlights the EcoRI restriction site. The L-shaped box with solid lines surrounds the digested restriction site and the resulting sticky ends.

Note: Dynabeads Oligo(dT)$_{25}$ have a binding capacity of 3 μg per 300 μL beads. SuperMag Streptavidin Magnetic Beads have a binding capacity of 12 μg per 300 μL of beads. Aside from the starting mass of the beads in step 1, and the omission of a forward primer when PCR amplifying, all steps are identical between the Dynabeads Oligo(dT)$_{25}$ approach and the Streptavidin Bead approach.

Note: Reagents should be placed on ice as much as reasonably possible, except for steps in which other temperatures are specified.

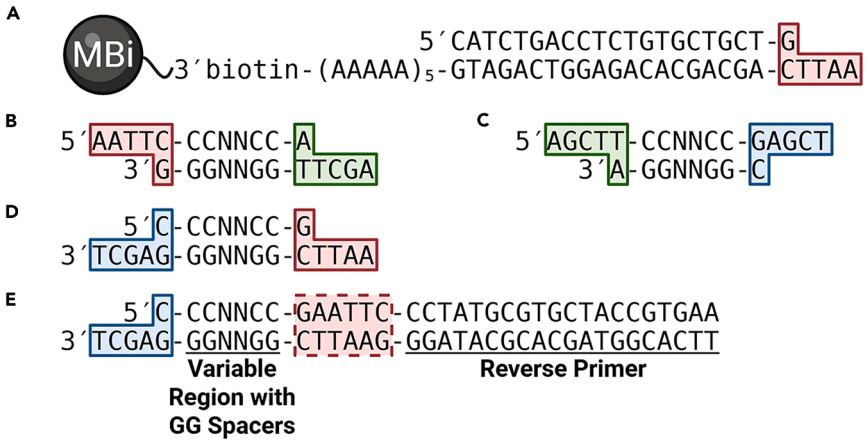


Figure 2. Fragments generated by the restriction enzyme digestion

Red highlights the EcoRI restriction site. Green highlights the HindIII restriction site. Blue highlights the SacI restriction site. The rectangular boxes with dashed lines surrounds the restriction site undigested during this method. The L-shaped boxes with solid lines surround the digested restriction sites and the resulting sticky ends.

(A) MBi after EcoRI digestion.

(B) MB1-# after HindIII and EcoRI digestion.

(C) MB2-# after SacI and HindIII digestion.

(D) MB3-# after EcoRI and SacI digestion.

(E) MB3t-# after SacI digestion.

1. Aliquot beads into 1.5 mL DNA LoBind Microcentrifuge Tubes and label the tubes in preparation for the addition of MBi, MB1-#, MB2-#, and MB3-# oligonucleotides (in step 3). For substeps a and b below, if using Streptavidin Beads, X is 75 μL and if using Dynabeads Oligo$(dT)_{25}$, X is 300 μL.
   a. MBi: One tube with X μL of beads at the stock concentration.
   b. MB1-#, MB2-#, and MB3-#: Two tubes each with X μL of beads at the stock concentration.
2. Wash beads in 500 μL of Wash Buffer. Place on a magnetic rack and remove the supernatant. Repeat for a total of two washes.

⚠ CRITICAL: The capture of the beads on the magnet should be allowed sufficient time to prevent the loss of beads during subsequent supernatant removal. This should take approximately between 2 to 5 min.

3. Dilute oligonucleotides in Wash Buffer and heat at 65°C for 2 min then place on ice.
   a. Dilute MBi oligonucleotides so there is 0.5 μg (500 ng) per 100 μL of Wash Buffer.
   b. Dilute MB1-#, MB2-#, and MB3-# oligonucleotides so there are 2 μg per 100 μL of Wash Buffer.
4. Determine the exact concentration of each oligonucleotide before incubation with the beads using a NanoDrop on ssDNA mode.
5. Add 100 μL of each oligonucleotide: MBi, MB1-#, MB2-#, and MB3-# to 300 μL of prewashed beads (supernatant removed).
6. Incubate with rotation at 20-25°C for at least 15 min.
7. Place on a magnetic rack and remove the supernatant.
8. Using a NanoDrop on ssDNA mode, measure the concentration of each oligonucleotide remaining in the supernatant after they have been incubated with the beads.
9. For MBi, calculate the exact quantity of oligonucleotides that were successfully bound to beads using the following equation. This is required for calculations to be completed during step 28. Troubleshooting 1.

$$ssDNA\ bound = \frac{ssDNA\ before\ incubation - ssDNA\ after\ incubation}{ssDNA\ before\ incubation} \times 500\ ng$$

10. Wash beads in 500 μL of Wash Buffer.
11. Place on a magnetic rack and remove the supernatant.
12. Re-suspend beads in 100 μL 1× Q5 Reaction Buffer.
13. For MB1-#, MB2-#, and MB3-#, split beads equally into two tubes (half of the bead suspension per tube).
14. Place on a magnetic rack and remove the supernatant.
15. Re-suspend beads in 50 μL of PCR reaction master mix and transfer to PCR tubes. Please prepare and use the appropriate mix, depending on which type of beads were used, according to the tables below.

SuperMag Streptavidin beads PCR reaction master mix

Reagent	Amount	Final concentration
5× Q5 Reaction Buffer	10 μL	1×
DNA Template	Template bound to 300 μL of beads	< 1,000 ng
10 μM Forward Primer	2.5 μL	0.5 μM
10 mM dNTPs	1 μL	200 μM
Nuclease-Free Water	36 μL	
Q5 High-Fidelity DNA Polymerase	0.5 μL	0.02 U/μL

Oligo(dT)$_{25}$ beads PCR reaction master mix		
Reagent	Amount	Final concentration
5× Q5 Reaction Buffer	10 μL	1×
DNA Template	Template bound to 300 μL of beads	< 1,000 ng
10 mM dNTPs	1 μL	200 μM
Nuclease-Free Water	38.5 μL	
Q5 High-Fidelity DNA Polymerase	0.5 μL	0.02 U/μl

⚠ CRITICAL: You must add the forward primer into the Streptavidin bead PCR reaction master mix, otherwise there will be no synthesis of the complementary strand.

Note: The forward primer is not required for the Oligo(dT)$_{25}$ bead PCR reaction master mix as the oligo(dT)$_{25}$ sequence on the beads acts as a primer. The reverse primer is not required, as only the complementary strand will be synthesized.

16. Transfer PCR tubes to a thermal cycler. Please prepare and use the appropriate thermal cycling conditions, depending on which type of beads were used, according to the tables below.

SuperMag Streptavidin beads PCR cycling conditions			
Steps	Temperature	Time	Cycles
Annealing	68°C	30 sec	1
Extension	72°C	2 min	1
Hold	4°C	forever	

Oligo(dT)$_{25}$ beads PCR cycling conditions			
Steps	Temperature	Time	Cycles
Annealing	50°C	30 sec	1
Extension	72°C	2 min	1
Hold	4°C	forever	

17. Transfer PCR reaction mixtures to 1.5 mL DNA LoBind microcentrifuge tubes.
18. Wash beads in 500 μL of Wash Buffer. Place on a magnetic rack and remove the supernatant.

Pause point: Beads with bound oligonucleotides can be stored at 4°C while resuspended in 500 μL of Wash Buffer. To avoid possible batch effects, samples should continue to be processed as soon as possible.

19. Wash beads in 100 μL 1× CutSmart buffer. Place on a magnetic rack and remove the supernatant.

Restriction enzyme digestion

Timing: 5 h–3 days

This major step describes single and double digestion of the oligonucleotide sequences. At the end of this major step, 80 tubes will contain double-digested sequences, and 17 tubes will contain single-digested sequences, all ready for subsequent ligation to form a full-length strand in the next major step. The double-digested sequences, including double-digested MB3-#, facilitate further addition of more fragments, whereas the single-digested MB3t-# will be the final fragment and

Table 1. Overview of restriction enzymes used, retained and discarded products, and restriction enzyme heat inactivation temperatures

Oligonucleotide	1st digestion	2nd digestion	Keep	Discard	Heat inactivation temperature
MBi	EcoRI-HF	N/A	Beads	Supernatant	N/A
MB1-#	HindIII-HF	EcoRI-HF	Supernatant	Beads	65°C for 20 min
MB2-#	SacI-HF	HindIII-HF	Supernatant	Beads	80°C for 20 min
MB3-# (double digested)	EcoRI-HF	SacI-HF	Supernatant	Beads	65°C for 20 min
MB3t-# (single digested)	SacI-HF	N/A	Supernatant	Beads	65°C for 20 min

therefore terminate the step-wise iteration of ligations. An overview of the treatment applied to each oligonucleotide group is detailed in Table 1.

20. For each tube, set up the first restriction enzyme digestion according to the following table:

Restriction enzyme reaction	
Reagent	Amount
Oligonucleotides (Bound to Beads)	Variable
10× rCutSmart Buffer	10 μL
Restriction Enzyme	2 μL (40 units)
Nuclease-Free Water	To 100 μL

21. Incubate for 1 h at 37°C or 8-12 h at 12°C.
22. Place tubes on a magnetic rack.
    a. For MBi, MB1-#, MB2-#, and for double digested MB3-# discard the supernatant and wash beads in 500 μL of Wash Buffer.
        i. Re-suspend beads with MBi 500 μL of Wash Buffer and aliquot 100 μL into four tubes to have 250 ng MBi per tube. Set aside for subsequent ligation.
    b. For single digested MB3t-#, collect supernatant and heat inactivate the restriction enzyme for 20 min at 65°C. Determine the exact concentration of each oligonucleotide in the supernatant using a NanoDrop on dsDNA mode and set aside for subsequent ligation. The concentration of MB3t-# will be required in step 38.

⚠ CRITICAL: Do not discard the supernatant for the single-digested MB3t-# in the step above, otherwise you will lose the digested oligonucleotides required for step 38.

23. Perform second digestion in MB1-#, MB2-#, and double digested MB3-# by repeating steps 13 and 14. This should complete digestion for all fragments, and the digested products should be as shown in Figure 2.
24. Place MB1-#, MB2-#, and double-digested MB3-# tubes on a magnetic rack and collect the supernatant.

⚠ CRITICAL: Do not discard the supernatant for MB1-#, MB2-# and the double-digested MB3-# in the step above, otherwise you will lose the digested oligonucleotides required for steps 28-38.

25. Heat inactivate the restriction enzymes in the supernatant of MB1-# and double digested MB3-# for 20 min at 65°C and MB2-# for 20 min at 80°C.
26. Determine the exact concentration of oligonucleotide in each supernatant using a NanoDrop on dsDNA mode to determine the volume of each oligonucleotide needed for step-wise ligation.

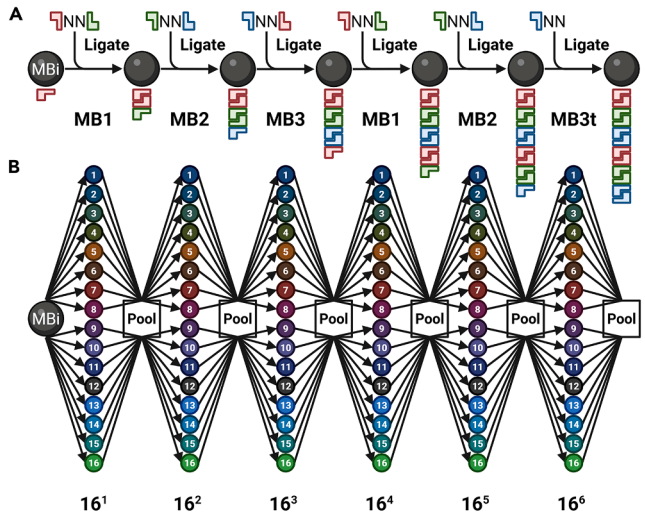


Figure 3. Schematic of the combinatorial oligonucleotide generation

Red represents the EcoRI restriction site. Green represents the HindIII restriction site. Blue represents the SacI restriction site. The L-shaped boxes with solid lines represent the digested restriction sites and the resulting sticky ends.

(A) Schematic of step-wise ligation to form a growing oligonucleotide onto the magnetic bead.

(B) Schematic of split-and-pool synthesis, showing multiple iterations of splitting of initial MBi into one of 16 groups for ligation with a subsequent fragment, and pooling back into a common group.

Stepwise ligation with split-and-pool synthesis

Timing: 12 h

This major step describes the ligation of each oligonucleotide fragment to form a growing oligonucleotide of 143 nt in length. Split-and-pool mixes during each ligation step facilitate a combinatorial synthesis of the oligonucleotides. An overview of the process is shown in Figure 3, and the resulting theoretical sequences that should be seen after each stage are as shown in Figure 4. While the protocol described stops at a total of six ligation steps, it is possible to increase or reduce the number of ligation steps, which will either increase or decrease the final strand length. As seen in Table 1, a portion of MB3-# is only single-digested. If the step-wise ligation does not terminate on an MB3-# fragment (i.e., instead terminating on an MB1-# or MB2-# fragment), then the restriction enzyme digestions described in the previous major step must be modified so that a portion of the terminating oligonucleotide is only single digested, so that it may be used as the final ligated fragment. If either MB1-# or MB2-# are single-digested, then a portion of MB3 will no longer need to be reserved for single digestion. At the end of the synthesis, the library is treated with NaOH, as 1 M NaOH has been shown to be very effective at denaturing DNA within 10 min.[4]

27. Place the tube containing immobilized MBi on a magnetic rack and discard the supernatant.
28. Remove the tube from the rack, divide both MBi and MB1-# supernatant into 16 different tubes and add the reagents for the ligation reaction as shown in the table below.

Step-wise ligation reaction

Reagent	Amount
2× Quick Ligation Reaction Buffer	5 μL
MBi	X ng on bead
MB1-#	X volume
Quick Ligase	2 μL
Nuclease-Free Water	To 50 μL

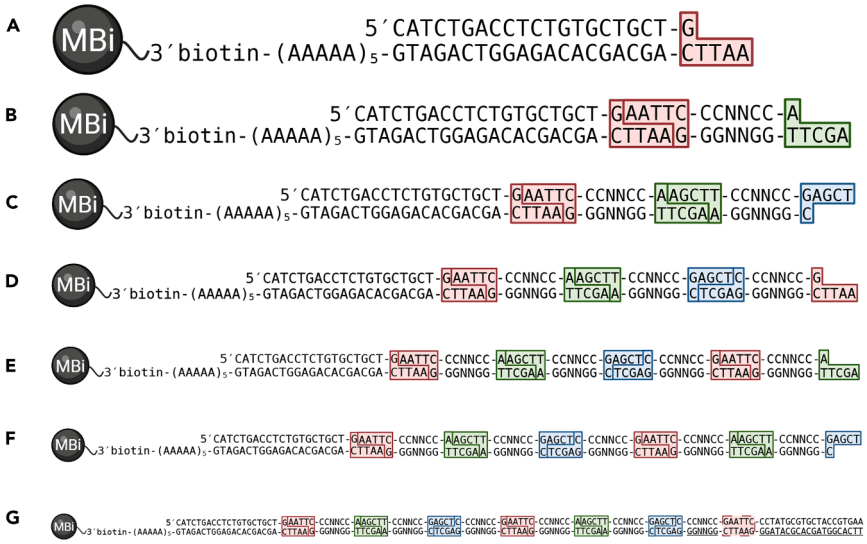


Figure 4. Theoretical products at each stage of the split-and-pool ligation

Red highlights the EcoRI restriction site. Green highlights the HindIII restriction site. Blue highlights the SacI restriction site. The rectangular box with dashed lines surrounds the restriction site undigested during this method. The L-shaped boxes with solid lines surround the digested restriction sites and the resulting sticky ends.

(A) MBi before ligation.

(B) MBi+MB1 after one iteration of ligation.

(C) MBi+MB1+MB2 after two iterations of ligation.

(D) MBi+MB1+MB2+MB3 after three iterations of ligation.

(E) MBi+MB1+MB2+MB3+MB1 after four iterations of ligation.

(F) MBi+MB1+MB2+MB3+MB1+MB2 after five iterations of ligation.

(G) MBi+MB1+MB2+MB3+MB1+MB2+MB3t after six iterations of ligation.

Note: Each of the 16 different tubes should only possess one unique MB1-# (i.e., MB1-1, or MB1-2, or MB1-3, etc.). When repeating with subsequent fragments for ligation (MB2-# and MB3-#), be sure to again put only one unique fragment in each of the 16 tubes. For the amount of MB1-# (and MB2-# and MB3-#) required at each ligation step, please consult the NEB Ligation Calculator, which can be found at the following link: https://nebiocalculator.neb.com/#!/ligation. Please note that the insert DNA length will be the size of the MB1-#, MB2-#, or MB3-# to be ligated onto the growing oligonucleotide, the vector DNA length is the cumulative length of the growing oligonucleotide up to that stage, and the vector DNA mass is the cumulative mass of the growing oligonucleotide up to that stage, with the first vector DNA mass (mass of MBi on the bead) being the result of the calculation performed in step 9 in the first major step. For reference, MBi is ~54 bp, each double-digested oligonucleotide is ~12 bp, and each single-digested oligonucleotide is ~35 bp. The amount of required insert DNA mass for every ligation (e.g., MB1-# to MBi) used for the reaction should be at the 2:1 ratio stated by the NEB Ligation Calculator, but for calculations, add only the value of the 1:1 insert DNA mass to the vector DNA mass into subsequent ligations. This is because higher ratios of the smaller fragment generally improve ligation efficiency, although only one fragment is likely to be ligated regardless of the higher ratio.

For example:

Determined ~205 ng MBi is bound to beads (step 9).

Ligation #1.

Insert DNA length = ~12 bp (MB1-#).

Vector DNA length = ~54 bp (MBi).

Vector DNA mass = 205 ng (MBi).

Required insert DNA mass:

1:1 = 45.56 ng (add onto subsequent Vector DNA mass calculations).

2:2 = 91.11 ng (use this mass for the MB1-# amount required for the ligation reaction, along with the concentration obtained in step 26 to determine the volume required).

Ligation #2.

Insert DNA length = ~12 bp (MB2-#).

Vector DNA length = ~66 bp = ~54 bp + ~12 bp (MBi + MB1-#).

Vector DNA mass = 250.56 ng = 205 ng + 45.56 ng (MBi + MB1-#).

Required insert DNA mass:

1:1 = 45.56 ng (add onto subsequent Vector DNA mass calculations).

2:2 = 91.11 ng (use this mass for the MB2-# amount required for the ligation reaction, along with the concentration obtained in step 26 to determine the volume required).

29. Gently mix the reaction by pipetting up and down.
30. Incubate at 20-25°C for 15-30 min.
31. Place the tubes on a magnetic rack and discard the supernatant.
32. Resuspend beads in 500 μL of Wash Buffer.
33. Pool all beads together into one 10 mL or 15 mL centrifuge tube and pipette up and down to mix, then split back into 16 individual 1.5 mL DNA LoBind microcentrifuge tubes.
34. Place on a magnetic rack and remove the supernatant.

Note: Do not proceed to steps 35-38 if ligating the terminal single-digested MB3t-# or any other single-digested terminal fragment. Proceed to step 39.

35. Wash beads in 100 μL 1× Quick Ligation Reaction Buffer. Place on a magnetic rack and remove the supernatant.
36. Prepare the second ligation reaction with MB2-# supernatant instead of MB1-# supernatant. Repeat steps 28-35.
37. Prepare the third ligation reaction with a double-digested MB3-# supernatant instead of MB1-# supernatant. Repeat steps 28-35.
38. Repeat steps 28-35, restarting at using MB1-# supernatant, then MB2-# supernatant, then single-digested MB3t-# supernatant instead of the double-digested MB3-# supernatant.
39. Incubate bead-bound dsDNA library with 500 μL 1 M NaOH for 10 min to denature the dsDNA.
40. Wash beads in 500 μL of Wash Buffer.
41. Place on a magnetic rack and remove the supernatant.

⚠ CRITICAL: The final ligation must be performed with a single-digested fragment, otherwise, the synthesized oligonucleotides will not possess the reverse primer site.

Note: Undigested MB1-#, MB2-#, and MB3-# all possess the reverse primer site, which will be retained even if single-digested, but cleaved off if double-digested.

Optional: Check that your synthesized oligonucleotides have reached the desired length. This can be done by first PCR amplifying your sample and using a gel electrophoresis or instrument such as an Agilent 2200 TapeStation System for sizing of oligonucleotides. Be sure to use only the PCR products suspended in the supernatant, and do not include the beads. Troubleshooting 2.

Optional: The 1.5 mL microcentrifuge tubes allow for mounting on the magnetic rack, which is required for removing supernatant from the sample and can be used for changing sample concentration. If desired, the final pool can be kept in a 10 mL or 15 mL centrifuge tube for easier storage.

Pause point: Beads with bound oligonucleotides can be stored at 4°C while resuspended in 500 μL of Wash Buffer. Do not allow beads to dry out.

Molecular cloning and sequencing

Timing: 2 h–3 days

It may be of interest to sequence either the entire oligonucleotide library or only a few select sequences generated from the library. This major step prepares the library for DNA sequencing, by inserting generated oligonucleotides into a vector, and optionally, transformation, amplification, isolation, and recovery of the oligonucleotides.

The PacBio Sequel II is an NGS DNA sequencing platform capable of sequencing libraries. It utilizes circular consensus (or HiFi) sequencing, forming a consensus sequence from multiple passes of the template molecule,[5] making it very attractive for sequencing samples which have variability in a few key nucleotide positions, such as the oligonucleotide synthesized by this protocol. However, our synthesized oligonucleotides are too short to be used in such a system and will therefore be inserted into a plasmid to reach a suitable length. This protocol will describe how to address this.

If there are only a few oligonucleotide sequences of interest, it may be more cost-effective to perform DNA sequencing using traditional methods, such as in low-throughput platforms utilizing the chain-termination method. In this case, individual sequences will need to be inserted into a cloning vector, transformed into bacteria, amplified, isolated, and recovered. This process can be adapted to isolate individual oligonucleotide sequences from the library.

42. Place tubes containing beads on a magnetic rack and remove the supernatant.
43. Re-suspend beads to roughly the bead's stock concentration of 1 mg/mL.
44. Transfer up to 300 μL of beads at the stock concentration to 1.5 mL DNA LoBind microcentrifuge tubes.
45. Place on a magnetic rack and remove the supernatant.
46. Re-suspend beads in 100 μL 1× Q5 Reaction Buffer.
47. Place on a magnetic rack and remove the supernatant.
48. Re-suspend beads in 50 μL of PCR reaction master mix according to the tables below, and transfer to PCR tubes.

PCR reaction master mix

Reagent	Amount
5× Q5 Reaction Buffer	10 μL
DNA Template	Template bound to 300 μL of beads

(*Continued on next page*)

Continued	
Reagent	Amount
10 μM Forward Primer	2.5 μL
10 μM Reverse Primer	2.5 μL
10 mM dNTPs	1 μL
Nuclease-Free Water	33.5 μL
Q5 High-Fidelity DNA Polymerase	0.5 μL

Note: In this step, both forward and reverse primers should be used regardless of whether the beads are SuperMag Streptavidin Beads or Oligo(dT)$_{25}$ beads.

49. Transfer PCR tubes to a thermal cycler. Prepare the appropriate thermal cycling conditions according to Table #.

PCR cycling conditions			
Steps	Temperature	Time	Cycles
Initial Denaturation	98°C	30 sec	1
Denaturation	98°C	10 sec	35 cycles
Annealing	68°C	30 sec	
Extension	72°C	2 min	
Final extension	72°C	2 min	1
Hold	4°C	forever	

50. Recover supernatant from PCR tubes, making sure to leave behind all beads. It may be necessary to transfer the PCR reaction to 1.5 mL tubes to enable tube mounting on a magnetic rack.

Note: The beads should still have dsDNA bound to the bead surface. The beads may be kept or discarded but are not required in the following steps.

51. Insert the PCR products into a cloning vector with a Zero Blunt PCR Cloning Kit using the manufacturer's instructions, which can be found in the following link: https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2Fzeroblunt_man.pdf.

Note: In the Zero Blunt PCR Cloning Kit manufacturer's instructions, only complete up to the Clone into pCR-Blunt section to ligate your PCR product into the cloning vector, and do not proceed to the Transform Competent Cells section.

Optional: If performing sequencing of only a few samples and not of the entire library, in the Zero Blunt PCR Cloning Kit manufacturer's instructions, complete the Transform Competent Cells section and also complete the Analyze positive clones step of the Analyze Transformants section, including recovery of the plasmid by a plasmid purification kit. These additional steps amplify and isolate individual sequences, generating the DNA mass required for low-throughput sequencing methods. Troubleshooting 3.

52. Determine the exact concentration of each oligonucleotide using a NanoDrop on dsDNA mode to determine the volume needed for digestion by EcoRV-HF.

53. Set up the restriction enzyme digestion according to the table below.

EcoRV digestion reaction	
Reagent	Amount
Oligonucleotides	1 μg
10× rCutSmart Buffer	5 μL
EcoRV-HF	1 μL
Nuclease-Free Water	To 50 μL

Note: Multiples of the enzyme reaction shown above may be performed, as needed, given the number of oligonucleotides that you wish to sequence.

54. Incubate for 1 h at 37°C or 8-12 h at 12°C.
55. Heat inactivate the EcoRV for 20 min at 65°C.
56. Cleanup the PCR products with a Monarch PCR & DNA Cleanup Kit using the manufacturer's instructions, which can be found at the following link: https://www.neb.com/en/protocols/2015/11/23/monarch-pcr-and-dna-cleanup-kit-protocol.

Note: Your oligonucleotides should meet the minimum length and purity requirements for DNA sequencing as dictated by the DNA sequencing instrument that is used. You may use quality control methods before sequencing to verify that your samples meet these requirements (e.g., gel electrophoresis for sizing oligonucleotides and/or a PicoGreen assay for precise quantification).

57. Proceed to DNA sequencing.

Optional: You can predict the folding and hybridization of your sequence using in silico methods, such as the mfold web server.[3]

EXPECTED OUTCOMES

After this protocol, provided that each digestion-ligation step is performed with high efficiency, the bead surface is theoretically expected to contain multiple copies of the same oligonucleotide. One possible application where it is beneficial to have multiple copies of the same sequence on the bead is in capture-SELEX aptamer screening, where the inclusion of different sequences on the bead surface may make the characterization of the recovered DNA aptamer-containing beads ambiguous. Currently, a limitation of the capture-SELEX is the inclusion of non-target binding DNA sequences in the recovered beads also containing the target-binding aptamers.[6] However, if each bead had the same sequences on the surface, then DNA amplification using the recovered beads containing target-binding aptamers will have reduced levels of non-target binding DNA sequences.

A diagram of the theoretical product of this protocol, a bead with multiple DNA sequences on the bead surface, along with the sequence of the final desired synthesized oligonucleotide is shown in Figure 5. Completion of this protocol should make the samples immediately ready for further analysis with other laboratory techniques, including nanoparticle tracking analysis, surface plasmon resonance imaging, and zeta potential measurement. This protocol has apparent potential applications for research in synthetic biology, therapeutics, and diagnostics, and can also help elucidate new functional nucleic acids that can also be applied to fundamental cell research.

We sequenced the oligonucleotide library and obtained 4,109,874 reads, of which 2,479,459 contained library synthesis products. Following sequencing, we performed bioinformatics analysis to characterize the library synthesis products. Library synthesis products were identified by locating

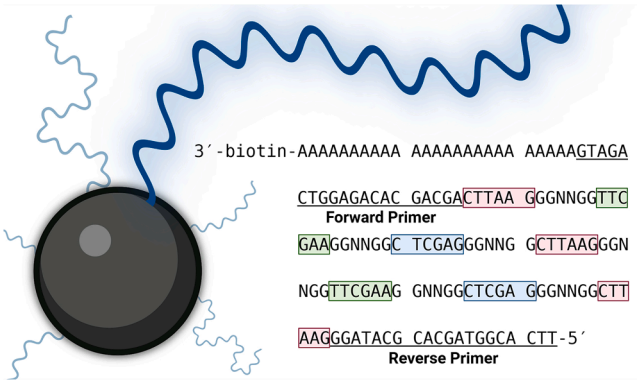


Figure 5. Diagram of the ideal final product produced by this protocol

The magnetic beads should possess multiple oligonucleotides bound to the bead surface. The ideal final synthesized sequence is as shown. Red highlights the EcoRI restriction site. Green highlights the HindIII restriction site. Blue highlights the SacI restriction site. The rectangular boxes with solid lines surround digested and ligated restriction sites. The magnetic nanoparticle and ssDNA are not to scale.

the forward and reverse primer sequences within the plasmid. While this method yields a theoretical full-length desired oligonucleotide of 143 nt on the nanoparticle, PCR amplification reduces the theoretical PCR products to 118 nt due to the loss of the downstream poly(A) sequence. From the characterization of the produced library, we observed that 77.4% of library synthesis products were within 40 to 58 nt, and 99.9% were within 40 to 118 nt (Figure 6A). The importance of 58 nt and 118 nt is that it is the length of the forward and reverse primers with integration of fragments from one or six iterations of split and pool, respectively. Unless a high level of synthesis efficiency is achieved, there may be a mix of lengths produced. However, it is possible that the level of heterogeneity of sequences on the bead surface is still reduced compared to if the library had been produced first and then subsequently coupled to beads.

Figure 6B shows the relative proportions of all two-nucleotide variations among the variable regions, looking separately at the unique and all (both unique and repeated) sequences. In this analysis, CCTACC and its reverse complement GGTAGG were the least represented in both groups, accounting for 1.76% and 0.34% in unique and all sequences, respectively. In contrast, CCGTCC and its reverse complement GGACGG were the most prevalent, accounting for 15.16% and 56.17% in unique and all sequences, respectively. Although some sequences were more strongly represented, this does not necessarily indicate that our method is biased toward producing any specific sequence. Following oligonucleotide library synthesis, an excess of dsDNA unbound from

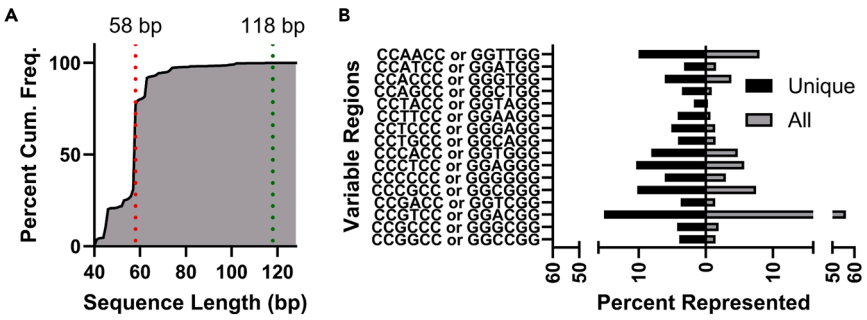


Figure 6. Bioinformatic analysis characterizing library diversity

(A) Percent cumulative frequency of lengths of DNA sequences after sequencing.

(B) Proportion of the variable regions present in unique sequence only and all (both unique and repeated) sequences.

(A and B) Data presented from 2,479,459 sequenced library synthesis products.

nanoparticles is produced by PCR amplification, and then the dsDNA is inserted into a plasmid. Both of these steps can randomly exaggerate or limit the representation of some sequences.

While we decided to make a few key decisions in the design of the final strand and as described in this protocol, the modular design of the fragments allows for either longer or shorter oligonucleotides, or potentially, a pool of mixed-length oligonucleotides terminated after a variable number of ligations. Longer oligonucleotides will yield a larger library, and shorter oligonucleotides will yield a smaller library. Additionally, the size of the library can also be changed by altering the number of variable nucleotides per fragment and can be calculated by the equation $theoretical\ library\ size = b^{n \times m}$, where b is the number of possible bases (b is four for natural DNA and RNA), n is the number of nucleotide positions available for variation per fragment, and m is the number of iterations of ligation. In our design, we are working with natural DNA with four possible bases (A, T, G, and C; $b = 4$), two variable nucleotide positions per fragment ($n = 2$), and six iterations of ligation ($m = 6$). This gives us $theoretical\ library\ size = 4^{2 \times 6} = 16^6$. Practically, the theoretical limit of such large libraries is never reached,[7] though this is a limitation that is not limited to our technique, but also to other randomized library generation methods. Python code simulating our split-and-pool ligation process for six iterations for one sequence is shown below.

```
import random

MBi = ("MBi") # MBi is constant

# MB1, MB2, and MB3 each contain 16 unique variations

MB1 = ("MB1-1", "MB1-2", "MB1-3", "MB1-4", "MB1-5", "MB1-6", "MB1-7", "MB1-8", "MB1-9", "MB1-
10", "MB1-11", "MB1-12", "MB1-13", "MB1-14", "MB1-15", "MB1-16")

MB2 = ("MB2-1", "MB2-2", "MB2-3", "MB2-4", "MB2-5", "MB2-6", "MB2-7", "MB2-8", "MB2-9", "MB2-
10", "MB2-11", "MB2-12", "MB2-13", "MB2-14", "MB2-15", "MB2-16")

MB3 = ("MB3-1", "MB3-2", "MB3-3", "MB3-4", "MB3-5", "MB3-6", "MB3-7", "MB3-8", "MB3-9", "MB3-
10", "MB3-11", "MB3-12", "MB3-13", "MB3-14", "MB3-15", "MB3-16")

m = 6 # Let m represent the number of iterations of ligation

ssDNA = MBi # MBi will appear in all sequences

for x in range(m): # Loop for the number of iterations of ligation

  if x%3 == 0:

    ssDNA += " + " + random.choice(MB1) # Split-and-pool to ligate an MB1-# fragment

  elif x%3 == 1:

    ssDNA += " + " + random.choice(MB2) # Split-and-pool to ligate an MB2-# fragment

  elif x%3 == 2:

    ssDNA += " + " + random.choice(MB3) # Split-and-pool to ligate an MB3-# fragment

print(ssDNA)
```

As we envisioned this library potentially used for aptamer screening, the number of ligations was to balance the length of the theoretical final sequence (143 nt on the bead and 118 nt after PCR amplification due to loss of the downstream poly(A) region), and the library size for typical aptamer screening. Aptamers are typically length-constrained and need a sufficiently large starting library to screen through to find an aptamer. While a larger number of variable bases in the variable region would increase theoretical diversity, two variable bases balance theoretical diversity with the practicality of performing this protocol (i.e., two variable bases lead to starting with 97 tubes, whereas three variable bases would lead to starting with 385 tubes). However, given the modular nature of

the fragments and the ability to tailor the DNA sequence to different applications, a different number of ligations or number of variable bases per variable region could be selected by the user for other potential applications.

LIMITATIONS

While the use of an Agilent 2200 TapeStation System can give an indication of sequence length, and whether the library synthesis product was successfully inserted into the plasmid, its sizing accuracy range is too wide to definitively distinguish between plasmid with and without the insert.

Unbound oligonucleotide can also be generated by PCR amplification, and this can be used to obtain either dsDNA or ssDNA by conventional or asymmetric PCR, respectively. This method is non-destructive, as the templates are not removed or destroyed during the PCR process. However, while the sequences on the nanoparticle are anticipated to be up to 143 nt long, products from PCR amplification are anticipated to be up to 118 long, as there is a loss of a downstream poly(A) region that isn't amplified.

Single-stranded oligonucleotides can be obtained by other methods, such as heating, NaOH, or formamide washes, although with poly(A)–oligo$(dT)_{25}$ coupling, the poly(A)-containing strand will become unbound from the bead, and with biotin-streptavidin coupling, the strand not containing biotin will become unbound from the bead (i.e., different oligonucleotide strands will be unbound from the bead, depending on the coupling method). This will yield one strand of oligonucleotide remaining coupled to the bead, and another strand in solution.

Note that we have decided to use a Q5 polymerase as it produces blunt-end PCR products. If you elect to use another polymerase, it may require a PCR polishing step to remove overhangs. This is the case with *Taq* polymerase, which lacks 3′ to 5′ proofreading activity, and will leave an A on the 3′ end of the amplified product.

While our protocol theoretically produces multiple copies of the same sequence on a given bead, this is only if there is a very high efficiency of synthesis. Our characterization of the library we produced showed that there may be a mix of sizes, as some sequences are terminated before the intended final number of ligations.[1] It is proposed that this is due to failed digestions in intermediate MB1-#, MB2-#, and MB3-#, resulting in fragments not having a sticky end to facilitate further ligation of subsequent fragments.

TROUBLESHOOTING

Problem 1: Low ssDNA uptake onto the beads

After incubation of the beads with the ssDNA, the measured unbound ssDNA concentration shows only a small decrease in unbound ssDNA concentration compared to before incubation of the beads with ssDNA. (Binding oligonucleotides to beads and complementary strand synthesis, step 9).

Potential solution

We had an average ssDNA capture of 38.3% and 59.0% onto oligo$(dT)_{25}$ beads and streptavidin nanoparticles, respectively. If you are using oligo$(dT)_{25}$ beads, you may switch to SuperMag Streptavidin Beads, as the biotin-streptavidin coupling shares a higher affinity for each binding partner compared to complementary poly(A)–oligo$(dT)_{25}$ binding.

Problem 2: Poor ligation of MB1-#, MB2-#, or MB3-# onto the growing oligonucleotide

After a quality control method or sequencing, the synthesized oligonucleotides appear smaller than expected. This may be due to poor ligation efficiency (Step-wise ligation with split-and-pool synthesis, step 41 optional step).

Potential solution

- Try performing the ligation at a lower temperature. We recommend first trying to lower the reaction temperature to 16°C and incubating 8-12 h.
- If possible, design the library to increase the size of the variable region, as the ligase supplier states that the ligase may ineffectively ligate strands smaller than 10 bp on each side of the break.

Problem 3: Failed ligation of insert into cloning vector/no colonies after plating

The pCR-Blunt vector contains a lethal gene from *Escherichia coli* called ccdB.[8] If a blunt PCR product is inserted into the cloning vector, it disrupts the expression of the ccdB gene, permitting the growth of positive recombinants that have been transformed. The absence of colonies after plating cells may indicate failed ligation. (Molecular cloning and sequencing, step 51 optional step).

Potential solution

- Plate competent *E. coli* cells that have not been transformed and will therefore not possess the lethal ccdB gene. If they grow, then your cells are still viable, and this suggests they can be transformed. In this case, your synthesized oligonucleotides may not have been inserted into the cloning vector and did not disrupt the ccdB gene.
- Perform gel electrophoresis on an instrument such as an Agilent 2100 Bioanalyzer System to confirm that your synthesized oligonucleotides have reached the expected length. This confirms successful oligonucleotide synthesis and synthesized oligonucleotides are available for insertion into the cloning vector.
- Perform PCR polishing to remove any overhangs that may prevent your synthesized oligonucleotides from being ligated to the cloning vector.
- If PCR polishing does not work, try a different type of blunt-end DNA polymerase to synthesize the oligonucleotides.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Maryam Tabrizian (maryam.tabrizian@mcgill.ca).

Technical contact

Technical questions on executing this protocol should be directed to and will be answered by the technical contact, Lidija Malic (lidija.malic@cnrc-nrc.gc.ca).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The datasets supporting the current study have not been deposited in a public repository but are available from the corresponding author on request.

ACKNOWLEDGMENTS

This research was undertaken, in part, thanks to funding from the Canada Research Chair in Regenerative Medicine and Nanomedicine. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants and Research Tools and Instruments Grants, the National Research Council Canada (NRC) Ideation Fund – New Beginnings Initiative, and the Fonds de recherche du Québec – Santé (FRQS) Doctoral Training Scholarship. We would like to thank the McGill Genome Centre for assistance with DNA sequencing and the Canadian Centre for Computational Genomics (C3G) for assistance with bioinformatics analysis. The graphical abstract and Figures 1, 2, 3, 4, and 5 were created with BioRender.com.

AUTHOR CONTRIBUTIONS

Conceptualization, L.M. and M.T.; methodology, J.V.L.N., L.M., C.N., and D.D.F.; validation, J.V.L.N., L.M., C.N., D.D.F., and M.T.; formal analysis, J.V.L.N.; investigation, J.V.L.N., C.N., and D.D.F.; resources, L.M. and M.T.; data curation, J.V.L.N., L.M., C.N., and D.D.F.; writing – original draft, J.V.L.N.; writing – review and editing, J.V.L.N., L.M., C.N., D.D.F.,

and M.T.; visualization, J.V.L.N.; supervision, L.M. and M.T.; project administration, L.M. and M.T.; funding acquisition, L. M. and M.T.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xpro.2025.103859.

REFERENCES

1. Nguyen, J.V.L., Meziadi, A., Nassif, C., Da Fonte, D., Malic, L., and Tabrizian, M. (2025). Combinatorial Nanoparticle-Bound ssDNA Oligonucleotide Library Synthesized by Split-and-Pool Synthesis. ACS Appl. Bio Mater. *8*, 844–853. https://doi.org/10.1021/acsabm.4c01681.

2. Oh, S.S., Qian, J., Lou, X., Zhang, Y., Xiao, Y., and Soh, H.T. (2009). Generation of highly specific aptamers via micromagnetic selection. Anal. Chem. *81*, 5490–5495. https://doi.org/10.1021/ac900759k.

3. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415. https://doi.org/10.1093/nar/gkg595.

4. Wang, X., Lim, H.J., and Son, A. (2014). Characterization of denaturation and renaturation of DNA for DNA hybridization. Environ. Health Toxicol. *29*, e2014007. https://doi.org/10.5620/eht.2014.29.e2014007.

5. Wenger, A.M., Peluso, P., Rowell, W.J., Chang, P.-C., Hall, R.J., Concepcion, G.T., Ebler, J., Fungtammasan, A., Kolesnikov, A., Olson, N.D., et al. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. Nat. Biotechnol. *37*, 1155–1162. https://doi.org/10.1038/s41587-019-0217-9.

6. Lyu, C., Khan, I.M., and Wang, Z. (2021). Capture-SELEX for aptamer selection: A short review. Talanta *229*, 122274. https://doi.org/10.1016/j.talanta.2021.122274.

7. Glanville, J., D'Angelo, S., Khan, T.A., Reddy, S.T., Naranjo, L., Ferrara, F., and Bradbury, A. R.M. (2015). Deep sequencing in library selection projects: what insight does it bring? Curr. Opin. Struct. Biol. *33*, 146–160. https://doi.org/10.1016/j.sbi.2015.09.001.

8. Bernard, P., Gabant, P., Bahassi, E.M., and Couturier, M. (1994). Positive-selection vectors using the F plasmid ccdB killer gene. Gene (Amst.) *148*, 71–74. https://doi.org/10.1016/0378-1119(94)90235-6.