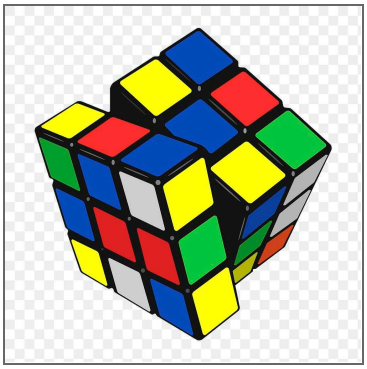

VERSION 2

OCT 03, 2023

OPEN ACCESS

DOI:
dx.doi.org/10.17504/protocols.io.ewov1q2p7gr2/v2

Protocol Citation: Mathew Chu 2023. Barcode Composition by Overlap-Extension PCR. **protocols.io** https://dx.doi.org/10.17504/protocols.io.ewov1q2p7gr2/v2 Version created by Mathew Chu

Protocol status: Working
We use this protocol and it's working

Barcode Composition by Overlap-Extension PCR V.2

Mathew
Chu[1]

[1]Walter and Eliza Hall Institute of Medical Research

Mathew Chu

ABSTRACT

Traditionally, DNA barcodes are synthesised as random oligonucleotides. However, this leads to uncertainty regarding the ground truth of barcode sequences in the experimental setting. Without reference sequences, it is impossible to determine with absolute confidence whether or not the observed diversity of barcodes is due to errors arising from sequencing itself or various cellular and molecular processes.

Here, we propose a modular way to assemble high-diversity libraries of DNA barcodes from units whose individual sequences are known a priori. Barcode units are amplified and then combinatorically spliced in vitro using overlap-extension PCR, yielding products that can be assembled into vectors for downstream applications.

IMAGE ATTRIBUTION

Image is under a Creative Commons licence.

MATERIALS

- Q5 DNA polymerase (NEB) or some other high-fidelity DNA polymerase
- NEBuilder HiFi DNA assembly cloning kit (NEB)

Created: Oct 03, 2023

Last Modified: Oct 03, 2023

PROTOCOL integer ID:
88717

Single Stranded DNA Pools for Combinatorial Assembly

1 ssDNA oligos for combinatorial assembly can be ordered as a pool (oPool). For a final barcode of *n* units, each with *m* diversity, order a set of *m* different barcode sequences for each unit:

```
unit_i = [Li + barcode_1 + Ri, Li + barcode_2 + Ri, …, Li +
barcode_m + Ri]
```

where unit *i* ($1 \leq i \leq n$) consists of *m* barcodes flanked by left (L) and right (R) homology sequences.

Homology sequences 1 to *n* should be 20 nt long and serve as both primer binding sites as well as junctions for the linear assembly of the barcode units. A set of orthogonal sequences for this purpose has been described and experimentally validated by Subramanian, Russ & Ranganathan (2018).

oPools should be divided into sets of 3-4 units to enable assembly of 3-4 oligo sets (barcode units) per reaction.

2 For each set of units, order primers for the left and rightmost homology sequences.

If the assembly will be performed in sets of 3, order primers for L1, R3, L3, R6, …

Stage 1 Assembly of Barcode Units

3 For each oPool, assemble 3-4 barcode units by performing 15 rounds of PCR using Q5 DNA polymerase (NEB). If relying on homology sequences from the literature cited above, use the following PCR parameters:

- Ta = 65 C
- extension time 10 s/kb (for the total length of the first stage assembly)
- reaction volume 100 uL
- 0.5 uM template

4 Remove the PCR reaction from the thermocycler and place immediately on ice.

5 For each assembly, add primers for the left and rightmost homology sequences at 0.5 uM directly to the reaction. Perform an additional 15 cycles of PCR.

6 Perform gel separation and extract the assembly product.

Hierarchical Assembly of Barcodes

7 Continue assembling barcodes in a hierarchical manner by combining 0.5 uM of products from the previous stage of assembly into a reaction with the above PCR parameters, adjusting the extension time at each stage.

8 In each stage, repeat the steps from the initial stage of assembly (15 initial rounds of PCR with template only, followed by 15 rounds of extension PCR) until n barcode units are assembled.

Sequencing and Cloning Barcode Libraries

9 Using overhang primers for the left and rightmost homology in the completed assembly (L1 and Rn), perform a final PCR with the same parameters to add adaptor sequences for next-generation sequencing. If deferring sequencing until after cloning is completed, skip this step.

10 To clone the assembled barcodes into a plasmid, design overhang primers with at least 20 nt homology to the linearized backbone at the desired cloning site. Using HiFi assembly (NEB), insert the barcodes as a single piece into the backbone. The resulting plasmid library can be sequenced on the Oxford Nanopore platform.

The theoretical diversity of the final barcode library is m^n.