ETH Zurich	EPP-SOP-ETH01
	Version 1.1

EPP - Standard Operating Procedure

Title: Preparation of gene libraries by error-prone PCR

Describes the steps required in order to run an error-prone PCR (epPCR) for the creation of DNA libraries of a given template gene.

distribution list EPP consortium			
changes to prior version: Section ‘6. Biosafety’ added.			
	name	signature	date
experimenter 1	Maximilian Bahls		25th February 2018

Instruction

Preparation of gene libraries by error-prone PCR.

1 Introduction/Purpose

In order to create new variants of a gene - and thus of the protein it is encoding (e.g. an enzyme or regulator) - epPCR can be run with the wildtype gene as a template. This library can subsequently be transformed into microbial cells and screened for variants with improved function (e.g. changed substrate specificity of an enzyme or improved catalytic efficiency).

2 Equipment and chemicals

Please note, this lists the equipment and chemicals required for epPCR only, i.e. it does not include general equipment for cultivation, transformation, or sequencing.

2.1 Equipment

PCR cycler (Eppendorf Mastercycler)
Plastic tubes, 200 µL (Axygen)
Plastic tubes, 1.5 mL (Greiner Bio-One)

2.2 Chemicals

Pure water (autoclaved, nuclease free)
Taq DNA polymerase (NEB)
10x Taq reaction buffer (NEB)
dNTPs (Sigma)
50 mM $MgCl_2$ (NEB)
10 mM $MnCl_2$ (AppliChem)
10 µM forward primer (Sigma)
10 µM reverse primer (Sigma)
Template DNA
DpnI (NEB)
Gel purification kit (Zymo)
Plasmid purification kit (Zymo)
10 µM sequencing primer (Sigma)

2.3 Bacterial strains

Escherichia coli, e.g. strain 10β (NEB, StrepR)

3 Media and buffers

3.1 LB

BD Difco LB broth, Miller (Luria-Bertani)

for LB agar, add 15 g L^{-1} BD Bacto agar

4 Procedures

4.1 Optimizing epPCR conditions

First, the region for epPCR has to be defined. Whenever possible, this target region should exclude regulatory elements (e.g. promoters and ribosome binding regions, terminators), start- and stop-codons and domains that should not be changed (e.g. when they are known to kill the functionality of the protein of interest). Depending on the cloning strategy, the primers have to include the functionalities required for inserting the PCR product into a plasmid backbone, e.g. the appropriate restriction sites for subsequent ligation or homology overlaps for isothermal assembly (ITA,"Gibson assembly"). Second, the error-rate of the polymerase can be adjusted by adding varied amounts of $MnCl_2$ to the reaction mix. This can be optimized by testing varied concentrations (e.g. 0 mM, 0.01 mM, 0.05 mM, 0.1 mM, 0.5 mM, 1 mM) of $MnCl_2$. The resulting amount of PCR product has to be verified on an agarose gel (too high Mn^{2+} concentrations do not allow the PCR to run). Subsequently the DNA should be gel-purified (and inserted in to a plasmid backbone), transformed into cells and plated on LB agar containing the required antibiotics. Then, single clones are checked for protein functionality and the amount of mutations they carry (per gene of interest, by Sanger sequencing). If the number of mutations is too high in order to find any functional proteins, the amount of Mn^{2+} has to be lowered, otherwise the highest concentration that allows the PCR to run should be used (as a rule of thumb). This optimization is specific for every gene/protein and should be carried out at the beginning of every protein engineering project, in order to obtain the targeted amount of changes per gene variant as compared to the wildtype.

A typical error-prone PCR mixture should look like the following example:

	Volume [µL]
Nuclease-free water	33.3
10x Taq reaction buffer	5
"error-prone" dNTPs	5
50 mM $MgCl_2$	2.95
10 mM $MnCl_2$	0.25
10 µM forward primer	1
10 µM reverse primer	1
Template DNA	1 (about 50 ng)
Taq DNA polymerase	0.5
Total volume	50

This gives approx. 2.95 mM Mg^{2+} and 0.05 mM Mn^{2+} in the reaction mix. The "error-prone" dNTPs include A: 3.5 mM, C: 4 mM, G: 2 mM, T: 13.5 mM, in order to partially

counterbalance the biases of Taq polymerase (see also [1] and [2] for details about “optimal” Taq polymerase error-rates and possible nucleotide biases).

A typical temperature cycling profile is shown below (please note that one temperature step is primer specific and the elongation time is dependent on the fragment length):

Step	Temp. [°C]	Time
Initial denaturation	95	1 min
25x thermocycle	95	30 seconds
	x (primer specific)	45 seconds
	68	1 min per kb (length specific)
Final Extension	68	5 minutes
Hold	4	for ever (storage)

4.2 Library creation and storage

Once the optimal epPCR conditions are determined, a larger volume of PCR mix should be aliquoted into 50 µL reactions, e.g. 200 µL into 4x 50 µL, for thermocycling. Subsequently, the wildtype template can be digested with DpnI (if it is methylated, i.e. plasmids prepared from most *E. coli* cloning strains). DpnI can be added directly to the PCR product (1 µL per 50 µL of PCR mix), either for 1 h at 37°C or overnight at room temperature. Subsequently, the PCR product is gel purified, introduced into the plasmid backbone (e.g. by restriction/ligation or ITA), transformed into an appropriate *E. coli* strain and plated on LB agar containing the appropriate antibiotics. The cells are grown only until the cfu become barely visible and are then harvested by scraping with a Drigalski spatula using about 5 mL of LB medium per agar plate (⌀ 15 cm). In order to estimate the library size, also plate 50 µL of a dilution series and incubate until cfus are properly countable. The plasmid-based library can then be isolated (“mini-prep”) and stored in pure water at -20°C until further use.

5 Remarks/Troubleshooting

For fine-tuning the polymerase’s error-rate during PCR, also the amount of template DNA or the amount of thermocycles can be varied (i.e. more template and/or less PCR cycles lead to a reduced accumulation of mutations per variant, and vice versa). If a relatively large library size is required (>10^4 variants), electroporation should be used for transformations. For long-term storage, the DNA library should be aliquoted into multiple tubes in order to thaw single tubes for each transformation.

6 Biosafety

No biosafety issues are associated with this protocol and the microbiological strains used (BSL1). Please note that the transfer of virulence factor gene(s) could increase the biosafety level to the corresponding BSL of the original pathogen.

7 Literature

1. Drummond, D.A., et al., *Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins.* J Mol Biol, 2005. **350**(4): p. 806-16.
2. Fromant, M., S. Blanquet, and P. Plateau, *Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction.* Anal Biochem, 1995. **224**(1): p. 347-53.