Amplification of complex gene libraries by emulsion PCR

Richard Williams[1], Sergio G Peisajovich[2], Oliver J Miller[1,3], Shlomo Magdassi[4], Dan S Tawfik[2] & Andrew D Griffiths[1,3]

[1]MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, UK. [2]Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel. [3]Institut de Science et d'Ingénierie Supramoléculaires (ISIS), 8 allée Gaspard Monge, BP 70028, 67083 Strasbourg Cedex, France. [4]Casali Institute of Applied Chemistry, The Hebrew University of Jerusalem, Givat Ram, 91904, Jerusalem, Israel. Correspondence should be addressed to A.D.G. (griffiths@isis.u-strasbg.fr) or D.S.T. (tawfik@weizmann.ac.il).

The efficient amplification of genomic libraries, cDNA libraries and other complex mixtures of genes by PCR is impeded by two phenomena: firstly, short fragments tend to be amplified in preference to larger ones; and, secondly, artifactual fragments are generated by recombination between homologous regions of DNA[1]. Recombination in this case occurs when a primer is partially extended on one template during one cycle of PCR and further extended on another template during a later cycle. Thus, chimeric molecules are generated, the short ones of which are then preferentially amplified as described in Figure 1. A variety of PCR protocols have been proposed to minimize these problems, most of which rely on high template concentrations and low numbers of PCR cycles[2,3]. Clearly, however, such an approach is not viable if little template DNA is available. Here we describe a protocol for amplifying complex DNA mixtures, based on the compartmentalization of genes in a water-in-oil (w/o) emulsion. Template fragments are segregated in the minute aqueous droplets of the emulsion and amplified by PCR in isolation (Fig. 1). This approach alleviates the problems described above while enabling the use of small amounts of template DNA and high numbers of PCR cycles. Box 1 describes an alternative method for generating very stable emulsions for emulsion PCR using the surfactant ABIL EM 90 (Fig. 2).

MATERIALS

REAGENTS

ABIL EM 90, a surfactant (Degussa)

Bovine serum albumin (BSA) (for molecular biology, powder; Sigma-Aldrich)

Deoxynucleoside triphosphate (dNTP) mix (Roche Diagnostics)

Diethyl ether (water-saturated; ≥99.5%; Riedel-de Haën)

Ethyl acetate (water-saturated; ≥99.7%; Riedel-de Haën)

Mineral oil (for molecular biology, light oil; Sigma-Aldrich)

Pfu Turbo DNA polymerase (2.5 U/μl; Stratagene)

Span 80, a surfactant (Fluka)

Triton X-100, a surfactant (general-purpose grade; Fisher Scientific)

Tween 80, a surfactant (Sigma-Aldrich)

EQUIPMENT

Centrifuge tubes (Falcon, 50 ml, polypropylene; BD Biosciences)

CryoTube vials (1.8 ml, round bottoms, star-feet; Nunc)

Magnetic stirrer with speed controller (Telesystem Stirring Drive 15 with Telemodul 40C; H+P Labortechnik AG)

Microcentrifuge (Minispin; Eppendorf)

Microcentrifuge tubes (MaxyClear, 1.7 ml; Axygen Scientific)

PCR plates (Thermowell, 96-well, polycarbonate; Corning Life Sciences)

Stir bars (3 × 8 mm, pivot rings, polytetrafluoroethylene; VWR International)

Thermal-cycler (DNA Engine Tetrad 2 Peltier Thermal Cycler; BioRad)

Vacuum centrifuge (Concentrator 5301; Eppendorf)

Wizard SV Gel and PCR Clean-Up System (Promega)

PUBLISHED ONLINE 21 JUNE 2006; DOI: 10.1038/NMETH896

PROCEDURE

Generating the emulsion

1| Prepare the oil-surfactant mixture by thoroughly mixing the following components in a 50-ml centrifuge tube at 25 °C:

Component	Amount	Final concentration
Span 80	2.25 ml	4.5% (vol/vol)
Tween 80	200 µl	0.4% (vol/vol)
Triton X-100	25 µl	0.05% (vol/vol)
Mineral oil	to 50 ml	

2| Transfer 400 µl of the oil-surfactant mixture to a CryoTube vial, and add a 3 × 8 mm stir bar. Begin stirring the mixture at 1,000 r.p.m. on the magnetic stirrer.

3| Prepare the aqueous phase for the emulsion by mixing the following components:

10× Cloned *Pfu* buffer	26 µl
BSA (100 g/l stock)	26 µl
Forward primer (10 µM stock)	7.8 µl
Reverse primer (10 µM stock)	7.8 µl
dNTPs (5 mM stock)	10.4 µl
Pfu Turbo DNA polymerase	5.2 µl
Template DNA	≤10^9 molecules (1.66 fmol)
Water	to 260 µl

▲**CRITICAL STEP**

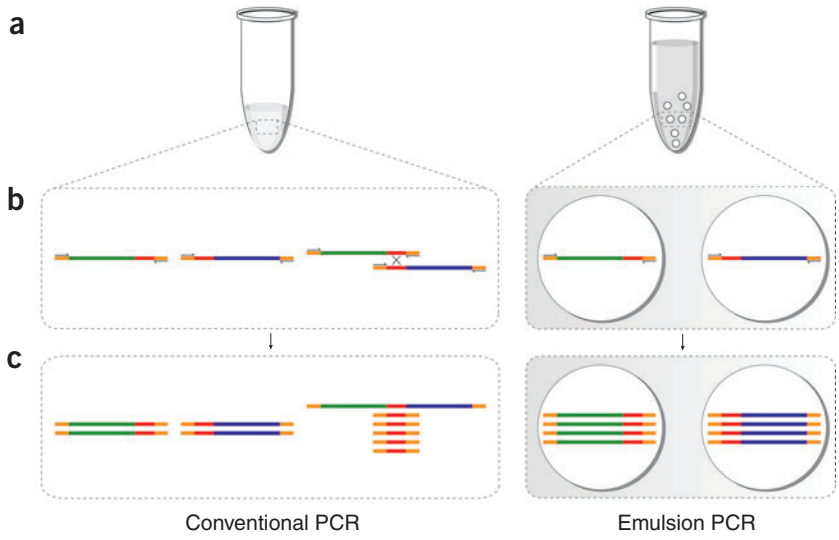

Figure 1 | Amplification of complex gene libraries by conventional PCR and emulsion PCR. (**a**) DNA fragments from a complex gene library are pooled together for amplification in a conventional PCR (left) or compartmentalized in the aqueous droplets of a w/o emulsion such that each droplet contains a single, or at most a few, template DNA molecules (right). (**b**) Each DNA molecule has an identical linker (orange) ligated at each end to allow amplification with a single PCR primer. If two template DNA molecules containing nonhomologous regions (green and blue) and homologous regions (red) are amplified in a conventional PCR, recombination events can occur, leading to the formation of chimeric products. In contrast, when the same fragments are amplified in an emulsion, the segregation of template DNA molecules prevents the formation of chimeric products. (**c**) In a conventional PCR, the short, chimeric products are amplified more efficiently than the template DNA molecules, leading to a buildup of artifactual DNA. The absence of chimeric products in the emulsion PCR prevents this from happening.

4| Add 200 µl of the aqueous phase to the oil-surfactant mixture in a dropwise manner over a period of 1.5 min. After the addition is complete, continue stirring for 5 min. A w/o emulsion is generated containing approximately 10^8–10^9 PCR-competent compartments per milliliter of emulsion.

➨ **TROUBLESHOOTING**

▲**CRITICAL STEP**

5| Pipet the emulsion into the wells of a PCR plate as 10 aliquots of 50 µl. Pipet 50 µl of the aqueous phase (Step 3) into a well as a nonemulsified control. Overlay the emulsified and nonemulsified reactions with mineral oil.

Temperature-cycling

6| Subject the PCR plate to the following program of temperature-cycling:

Cycle number	Denaturation	Annealing	Polymerization
1	2 min at 95 °C		
2–26	30 s at 95 °C	30 s at *A* °C	*B* min at 72 °C
27			10 min at 72 °C

A = *C* – 5 where *C* is the melting temperature (T_m) in °C of the primer with the lowest T_m. *B* is equal to the length of the amplification target in kbp.

Place the nonemulsified reaction aside until Step 9.

■ **PAUSE POINT** Both the emulsified and nonemulsified PCR reactions can be stored at 4 °C overnight.

BOX 1 AN ALTERNATIVE PROTOCOL FOR EMULSION PCR

Whereas the main protocol can be used for many applications[4,5,7] the protocol described in this box may be especially useful when performing more than 35 PCR cycles, owing to the stability of the generated emulsion. As with the main protocol, it has been observed that emulsion PCRs generated by this method exhibit reduced competition between template molecules of different lengths (**Fig. 2a**) and produce fewer chimeric molecules (**Fig. 2b**).

To generate an ABIL EM 90-based emulsion, follow the main protocol, but perform all the emulsification steps on ice and make the following changes:

Step 1 Prepare an oil-surfactant mixture that is 2% (vol/vol) ABIL EM 90 and 0.05% (vol/vol) Triton X-100 in mineral oil.

Step 2 Use 900 µl of the oil-surfactant mixture and a stirring speed of 1,400 r.p.m.

Step 3 Prepare a 100-µl PCR as the aqueous phase, but supplement it with 0.5 g/l of BSA.

Step 4 Add the aqueous phase to the oil-surfactant mixture over a period of 2 min and continue stirring for 5 min.

Step 8 After the two extractions with diethyl ether, perform an extraction with 1 ml of water-saturated ethyl acetate, and then two more extractions with diethyl ether. The extraction with ethyl acetate is necessary to remove the 'gel-like' phase, which is formed when diethyl ether is mixed with ABIL EM 90.

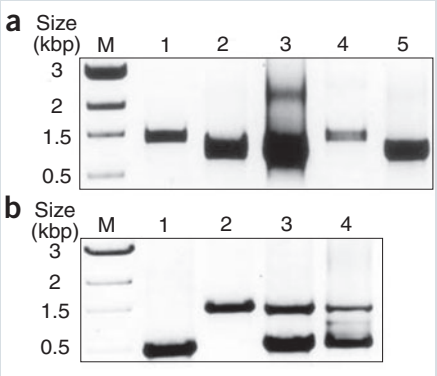

Figure 2 | Ethidium bromide–stained agarose gels reveal the absence of size competition and chimeric product–formation in ABIL EM 90–based emulsion PCRs. (**a**) Prevention of size competition. A long template and a short template were both amplified by emulsion PCR using the same pair of primers (lanes 1 and 2). When the primers were excluded from the short template-emulsion reaction and the two emulsions were mixed before temperature cycling, only the long template was amplified (lane 4). This demonstrated that the emulsion had been stable over 35 PCR cycles and that no exchange of primers had occurred between the emulsion droplets. Mixing the templates before emulsification (lane 5) or not emulsifying the reaction (lane 3) led to the short template being preferentially amplified. (**b**) Prevention of chimeric product–formation. Two similar templates were constructed with one template containing two internal deletions. Both templates could be amplified by emulsion PCR using the same pair of primers (lanes 1 and 2). Owing to the extensive homology, when the two templates were amplified together by nonemulsified PCR, recombination occurred and chimeric products of intermediate length were generated (lane 4). When the templates were amplified together in an emulsion PCR, no chimeric products were observed (lane 3).

Breaking the emulsion

7| Pool the emulsified PCR reactions in a 1.7-ml microcentrifuge tube and centrifuge at 13,000*g* for 5 min at 25 °C. Dispose of the upper (oil) phase.

8| Several extractions with an organic solvent will remove the remaining oil from the emulsion and cause it to break. Perform the following extraction twice: add 1 ml of water-saturated diethyl ether, vortex the tube, and dispose of the upper (solvent) phase.
▲**CRITICAL STEP**

9| Remove residual solvent from the broken emulsion by centrifuging under vacuum for 5 min at 25 °C.

10| Clean the emulsified (Step 9) and nonemulsified (Step 6) PCR products using the Wizard SV Gel and PCR Clean-Up System.

11| Analyze aliquots of the emulsified and nonemulsified PCR products by electrophoresis through an agarose gel.
➨ **TROUBLESHOOTING**

TROUBLESHOOTING TABLE

PROBLEM	SOLUTION
Step 4 A coarse emulsion is formed. Low-quality emulsions can be identified by their appearance (they are pale and opaque rather than 'creamy'-white) and their tendency to break down during centrifugation. Of course, light microscopy and/or dynamic light scattering techniques can be used to make more quantitative determinations of emulsion quality.	Fix the tube at the precise center of the stirring point using a drilled aluminum block or clamp and ensure that the stir bar is spinning freely at the base of the CryoVial tube before adding the aqueous phase. If any of the surfactants are old then consider replacing them.
Step 11 Little or no product is generated in the emulsified reaction.	If the nonemulsified reaction has generated a product then the problem lies with the emulsification process: ensure that the correct amount of BSA is added to the aqueous phase (Step 3) and that the correct equipment (stir bar, CryoTube vial, etc.) is being used in Steps 2 and 4. If both the emulsified and nonemulsified reactions have failed then the PCR strategy may be flawed: consult the instructions provided by the manufacturer of the DNA polymerase.

CRITICAL STEPS

Step 3 Failure to add BSA or another 'bulk' protein to the aqueous phase will result in little or no amplification. It is speculated that the presence of a bulk protein is necessary to prevent the DNA polymerase from becoming trapped and denatured in the oil/water interface of the emulsion droplets. In the case of this particular emulsification strategy, the presence of a bulk protein is also critical for the formation of a stable emulsion.

Step 4 The temperature at which an emulsion is generated affects the size distribution of its droplets. When using this protocol, emulsification should be performed at 25 °C using 25 °C components for the main protocol and on ice using ice-cold components for the alternative protocol (Box 1).

Step 8 Ensure that the diethyl ether is water-saturated: shake 100 ml of diethyl ether with 100 ml of water in a bottle for 30s. Allow the phases to settle before use (the top phase is diethyl ether).

COMMENTS

Emulsion PCR has been used for the directed evolution of DNA polymerases[4,5], single-molecule reverse-transcription PCR[6] and haplotyping[7]. 'Solid-phase' emulsion PCR, using microbeads, has been used for the detection and enumeration of rare genetic mutations[8] and the high-throughput screening of transcription-factor targets[9]. Solid-phase emulsion PCR has also been used to generate the template DNA for two new DNA-sequencing technologies[10,11]. These technologies allow sequencing at one-ninth

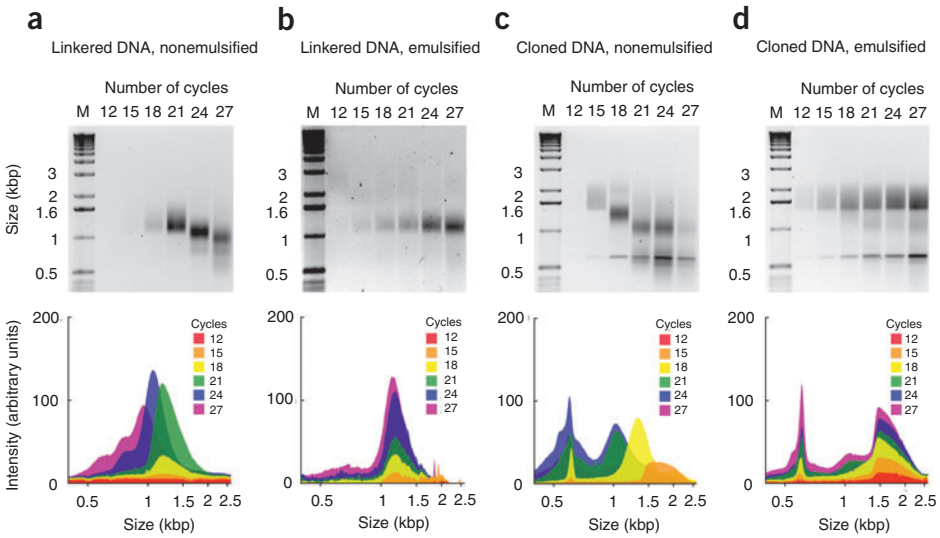

Figure 3 | Amplification of genomic DNA libraries by conventional PCR and emulsion PCR. (**a**) Linkered genomic DNA as the template, nonemulsified. (**b**) Linkered genomic DNA as the template, emulsified. (**c**) Cloned genomic DNA as the template, nonemulsified. (**d**) Cloned genomic DNA as the template, emulsified. Images are of ethidium bromide–stained agarose gels. M, marker. Each lane contains the DNA product from a 160-μl emulsified aqueous phase or a 16-μl nonemulsified aqueous phase. The intensity profiles of the lanes in each gel are shown below the gel.

the cost per base and with 100-fold greater throughput than the conventional technology, thus enabling the sequencing of a bacterial genome in a single four-hour run.

EXAMPLE OF APPLICATION

We sheared genomic DNA from *Haemophilus parahaemolyticus* to produce fragments with an average length of ~1.3 kbp, blunt-end repaired the fragments, phosphorylated them and size-selected fragments of 1.0–1.6 kbp by agarose gel electrophoresis[12]. We ligated short oligonucleotide pairs, containing a primer binding site, to the ends of DNA in one aliquot of this preparation ('linkered genomic DNA') and ligated another aliquot into a plasmid vector ('cloned genomic DNA'). We amplified 5×10^9 DNA molecules from each library by conventional PCR and by emulsion PCR. Analysis of the PCR products on an agarose gel (**Fig. 3**) revealed that, for both the linkered DNA (**Fig. 3a**) and cloned DNA (**Fig. 3c**), the average size of the amplified DNA in the nonemulsified PCR decreased with the number of PCR cycles. When we amplified the libraries by emulsion PCR, the size distributions of the amplified DNA remained essentially constant, irrespective of the number of PCR cycles (**Fig. 3b**,**d**).

The absence of artifactual PCR products when using emulsion PCR can be explained by the presence of a single, or at most a few, template DNA molecules in each droplet of the emulsion. This segregation of template DNA molecules prevents recombination between homologous or partially homologous gene fragments during PCR, thus eliminating the synthesis of short, chimeric products and other artifacts (**Fig. 1**). The compartmentalization of template DNA molecules also reduces the competition between fragments of different lengths, thus diminishing the bias for amplifying smaller fragments.

COMPETING INTERESTS STATEMENT
The authors declare competing financial interests (see the *Nature Methods* website for details).

Published online at http://www.nature.com/naturemethods/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Meyerhans, A., Vartanian, J.P. & Wain, H.S. DNA recombination during PCR. *Nucleic Acids Res.* **18**, 1687–1691 (1990).
2. Polz, M.F. & Cavanaugh, C.M. Bias in template-to-product ratios in multitemplate PCR. *Appl. Environ. Microbiol.* **64**, 3724–3730 (1998).
3. Qiu, X. *et al.* Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. *Appl. Environ. Microbiol.* **67**, 880–887 (2001).
4. Ghadessy, F.J., Ong, J.L. & Holliger, P. Directed evolution of polymerase function by compartmentalized self-replication. *Proc. Natl. Acad. Sci. USA* **98**, 4552–4557 (2001).
5. Ghadessy, F.J. *et al.* Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. *Nat. Biotechnol.* **22**, 755–759 (2004).
6. Nakano, M. *et al.* Single-molecule reverse transcription polymerase chain reaction using water-in-oil emulsion.

J. Biosci. Bioeng. **99**, 293–295 (2005).
7. Wetmur, J.G. *et al.* Molecular haplotyping by linking emulsion PCR: analysis of paraoxonase 1 haplotypes and phenotypes. *Nucleic Acids Res.* **33**, 2615–2619 (2005).
8. Dressman, D., Yan, H., Traverso, G., Kinzler, K.W. & Vogelstein, B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc. Natl. Acad. Sci. USA* **100**, 8817–8822 (2003).
9. Kojima, T. *et al.* PCR amplification from single DNA molecules on magnetic beads in emulsion: application for high-throughput screening of transcription factor targets. *Nucleic Acids Res.* **33**, e150 (2005).
10. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380 (2005).
11. Shendure, J. *et al*. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728–1732 (2005).
12. Frohlich, M.W. & Parker, D.S. Running gels backwards to select DNA molecules larger than a minimum size. *Biotechniques* **30**, 264–266 (2001).