

Overlap Extension PCR: An Efficient Method for Transgene Construction

Matthew D. Nelson and David H.A. Fitch

Abstract

Combining genes or regulatory elements to make hybrid genes is a widely used methodology throughout the biological sciences. Here, we describe an optimized approach for hybrid gene construction called overlap extension PCR. In this method, the polymerase chain reaction (PCR) is employed for efficient and reliable construction of hybrid genes. A PCR-based approach does not rely on available restriction sites or other specific sequences, an advantage over more conventional cloning or recombineering methods. With the use of high-fidelity DNA polymerase, this method can be used for making even very large constructs (>20 kb) with minimal unwanted mutations. Finally, overlap extension-PCR can be used as a means for site-directed mutagenesis, introducing desired mutations to the final hybrid gene.

Key words: PCR, Cloning, Hybrid genes, Transgenics, Site-directed mutagenesis, Gene expression

1. Introduction

Hybrid genes are powerful tools for the study of gene expression and function. Temporal and spatial expression of genes and gene products can be determined by fusing *cis*-regulatory elements and coding sequences of genes to the coding sequences of reporters, such as the green fluorescent protein (1). Gene regulation can be studied by subsequent mutagenesis of stretches of sequence within these hybrid genes. For example, mutating a functional transcription factor binding site within the promoter of a gene can result in altered expression of a reporter relative to its nonmutated form. Also, fusing variations of the coding sequence of a protein to a reporter can provide information about the localization and function of gene products. In addition, for a more detailed study of where and how a protein functions *in vivo*, one can misexpress or

partially express a gene using the promoter of another gene. In the context of evolutionary investigations, fusion of genes can be used to create hypothetical ancestors, interspecific hybrids, or other such constructs. *Cis*-regulatory evolution can involve the reorganization and change in sequences of *cis*-regulatory elements (2), and specific hypotheses can be tested via the construction of reporters driven by various regulatory sequences. Also, directed evolution experiments to design thermostable enzymes (3) have used various gene fusion methods to study protein-coding sequence evolution.

For all of the above applications, one must combine multiple genes or portions of genes together to construct a hybrid gene. Hybrid gene construction has conventionally been accomplished by restriction enzyme-based cloning methods. The gene constructs can then be further manipulated via site-directed mutagenesis. Powerful as these methods are, conventional cloning usually requires multiple days to weeks of work and troubleshooting. Such cloning methods are dependent on particular sequences recognized by restriction enzymes. Recombineering methods, which utilize recombinase systems, are much faster and more versatile than traditional restriction enzyme-based methods but more costly and still require the incorporation of particular sequences.

Overlap-extension PCR (OE-PCR) provides a rapid and cost effective means for the creation of hybrid genes without the need for available enzyme recognition sequences (4–7). Furthermore, PCR-based methods have become increasingly reliable with the introduction of high-fidelity DNA polymerases, thus limiting the number of unwanted mutations in the final construct (8). This method has proven to be useful for the construction of even very large hybrid genes (>20 kb) (9). Not only does OE-PCR allow for combining genes; it can also be used to directly mutate sequences within the genes that are being fused together (10, 11). OE-PCR produces a linearized hybrid gene that can simply be cloned into a standard plasmid for bacterial propagation (or used directly, as in the case of *C. elegans* transgenics). This chapter describes the methodology of OE-PCR for the construction of hybrid genes and for the application of site-directed mutagenesis. The protocol described here has been adapted from previously published methods (9, 10) and optimized in our lab for making fluorescent protein fusions and gene expression reporters for *C. elegans* developmental genetics, as well as for creating hybrids from genes of different species. However, the method is generally applicable for genetic engineering.

2. Materials

2.1. PCR

1. DNA oligonucleotides (oligos) purified by desalting and manufactured at standard synthesis scale (0.025 μmol for oligos <50 bp, and 0.05 μmol for >50 bp). Oligos are dissolved in 1 \times

TE buffer (pH 8.0, 10 mM Tris-HCl, 1 mM EDTA) to 100 μ M stocks and stored at -20°C . Working stocks are diluted in ultrapure water to 25 μ M and stored at -20°C .

2. Herculanse Hotstart DNA polymerase (Agilent Technologies), stored at -20°C . Any *Pfu* DNA polymerase can be used. A 10 \times stock of Herculanse reaction buffer is provided with the enzyme. Importantly, a DNA polymerase such as that derived from *Pyrococcus furiosus* (*Pfu*) must be used, not an enzyme such as that derived from *Thermus aquaticus* (*Taq*). *Taq* enzymes leave A-overhangs at the 3'-end, which would disrupt the overlap steps (9).
3. Deoxyribonucleotide triphosphates (dNTPs) are diluted to a working solution of 10 mM each in ultrapure water.

2.2. DNA Purification and Agarose Gel Electrophoresis

1. Agarose is dissolved in 1 \times TAE buffer (40 mM Tris-acetate, 1 mM EDTA), heated to make an agarose gel, and supplemented with ethidium bromide (0.5 $\mu\text{g}/\text{mL}$). Electrophoresis is carried out in 1 \times TAE.
2. When a PCR reaction contains only a single fragment, visualized by gel electrophoresis as a single band, the PCR reaction can be purified directly over a column. We use the QIAquick PCR purification kit (Qiagen).
3. However, if the PCR product contains multiple fragments, the PCR products must be separated by gel electrophoresis and the appropriate bands cut out and purified. We use the Wizard[®] SV gel and PCR-purification system (Promega).
4. Concentration of PCR products is determined using a NanoDrop 2000c (Thermo Scientific).

3. Methods

The concept of OE-PCR is to combine fragments of DNA by using complementary 3' or 5' ends in a PCR reaction without primers (Fig. 1). The complementary sequence is introduced to one or both ends of a fragment through hybrid primers which match the template for approximately 20 nucleotides at the 3' end and have a 5' tail which matches the sequence of the fragment to be fused. This allows the complimentary ends of the fragments to act as priming sites for DNA polymerase so that the fragments are spliced together. Nested primers are added in a subsequent PCR reaction to amplify the desired hybrid gene.

3.1. Oligo Design

A crucial first step in OE-PCR is the design of oligonucleotides. Designing oligos that bind at the 5' and 3' ends of the desired hybrid gene is straightforward and does not require the addition of

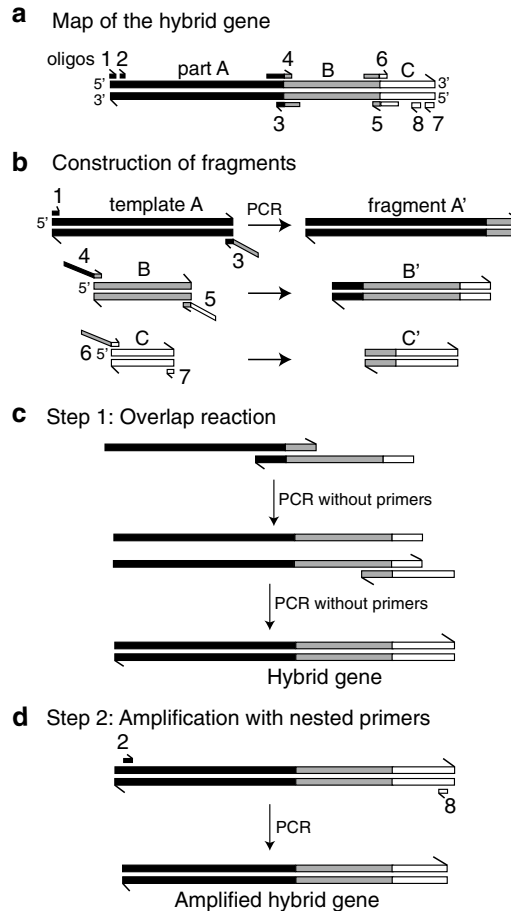


Fig. 1. The concept of OE-PCR. Bars represent single DNA strands; different shades of the bars represent the different sequences that will be fused. The 3' end of the strand is indicated by a small line like a half-arrowhead projecting from the end of each bar. The 5' end is sometimes indicated with a label. (a) Overall map of the oligo sequences to the final hybrid product. The hybrid consists of three parts: A, B, and C. In the example described in the text, A is the 5'-regulatory region of *C. elegans nhr-25*, B is the coding sequence for GFP, and C is the 3'-UTR of *nhr-25*. Sequences corresponding to the eight primers (thin bars) used to generate this hybrid gene are indicated adjacent to the strand from which the sequences come (sense on top, antisense on bottom). (b) Construction of the fragments. Oligo pairs (e.g., oligos 1 and 3) are shown that bind to a template sequence (e.g., template A) and are used to amplify a fragment (e.g., fragment A'). Only the 3' end of each hybrid primer (here, oligos 3–6) binds to the template. The 5' tails of the hybrid primers generate the portions of the fragment that overlap the sequence of a different template. (c) In step 1 of the overlap reaction (overlap extension), which occurs in the absence of primers, some of the reannealing strands will be from different fragments; they will anneal in the overlap regions created by the hybrid primers. Extension from the 3' ends of the overlapping fragments will result in “fused” products. This will actually occur in at least two steps. In the example shown, fragment A' and B' combine first and, in a subsequent step, an A' + B' strand anneals with a C' strand; extension will then result in the final tripartite hybrid product. (d) In step 2 of the overlap reaction (amplification), the hybrid gene is specifically amplified by using nested primers (oligos 2 and 8).

any noncomplementary sequences to the ends (Fig. 1, oligos 1, 2, 7, 8). However, the internal oligos that connect different stretches of DNA require the addition of nucleotides to the 5' end (Fig. 1, oligos 3–6). The additional sequence is not complementary to the template sequence but is complementary to the sequence to which the PCR product will be fused. The specifics for the design of oligos are as follows (see Fig. 1a, b):

1. The number of oligos needed for a given gene construct can be determined by the formula: $2N + 2$, where N is the number of unique PCR amplified sequences to be fused together. Figure 1 shows a hybrid gene composed of three parts and thus requires 8 oligos. Theoretically, however, this technique can be used to fuse as many fragments as desired.
2. Oligos 1, 2, 7, and 8 should be >21 bp in length to prevent nonspecific binding and thus unwanted products that will interfere with the overlap steps. In this example, oligos 1 and 2 are “forward” primers (contain sequence from the “sense” strand) and oligos 7 and 8 are “reverse” primers (i.e., contain sequence from the “antisense” strand). That is, oligos 7 and 8 must be antiparallel to the orientation of oligos 1 and 2. Also, it should be noted that oligos 2 and 8 are nested with respect to oligos 1 and 7. The distance between oligos 1 and 2 or between 2 and 8 is arbitrary (see Note 1).
3. Oligos 1 (sense) and 3 (antisense) are the primer pair used to produce and amplify fragment A'. The 3' portion of oligo 3 which binds to template A should be 21–30 bp in length. The 5' portion of oligo 3 that is not complementary to template A should be as long as possible (>20 bp) without significant secondary structure and should be complementary to template B. If fusing two coding sequences, make sure the sequences maintain the correct reading frame. Also, ensure that the orientation of the fusion is correct (i.e., for oligo 3, the 5' end of the antisense sequence from template A is linked to the 3' end of the antisense sequence from template B).
4. Oligos 4 (sense) and 5 (antisense) are the primer pair for amplifying fragment B'. Oligo 4 is essentially the reverse complement of oligo 3. However, the 3' part that binds to template B should be adjusted to include 21–30 bp at the end of template B. Also, the 5' tail of oligo 4 should be lengthened (relative to the corresponding portion of oligo 3) if possible to increase the amount of overlap. The 3'-end of oligo 5 is the antisense sequence of template B. The 5' portion of oligo 5 contains sequence complementary to the beginning of template C.
5. Oligos 6 (sense) and 7 (antisense) are used to produce and amplify fragment C'. Oligo 6 is similar to, but the reverse complement of oligo 5. The 3' portion of oligo 6 that binds to

template C should be 21–30 bp long and the 5' tail that overlaps with B sequence should be as long as possible.

A general principle in designing the overlap primers is that the 3' portion that is complementary to the template should range 21–30 bp in length and the 5' tail with the “overlap sequence” should be as long as possible so as to maximize the chances of overlap in subsequent steps. This requires the design of relatively long oligos (50–90 bp), and optimization of sequences to limit homo- and heterodimer formation is necessary. Tools useful for estimating the properties of the designed oligos are listed below. However, other online or downloadable tools are also available.

To check for homodimers and to estimate the melting temperature (T_m), we use the Sigma DNA Calculator (<http://www.sigma-genosys.com/calc/DNACalc.asp>).

To determine the nature of interactions between oligo pairs (i.e. heterodimers), we use the Finnzyme Multiple primer analyzer (http://www.finnzymes.com/java_applets/multiple_primer_analyzer.html). The default settings are sufficiently stringent to design useful oligos, and the “value for sensitivity of dimer detection” has been adjusted as high as 5 (scale 1–10) with success.

3.2. Amplification of Fragments and Purification

While OE-PCR can be used to create a broad spectrum of different types of hybrid genes, one very useful application is the construction of GFP (green fluorescent protein) transgenes. We use this application as an example for the remainder of this chapter. In this example, the hybrid gene is composed of three fragments: A', B', and C', which are amplified from templates A, B, and C, respectively. Both the A' fragment and C' fragment are amplified from genomic DNA, while the B' fragment is amplified from a gfp-containing plasmid, pPD95.75 (Addgene) (12). Fragment A' is the 5'-*cis*-regulatory region and promoter of the *nhr-25* gene of *C. elegans*; fragment B' is the coding sequence, containing synthetic introns, of GFP; and fragment C' is the 3'-UTR (untranslated region) of *nhr-25* (Fig. 1a).

1. Reaction mix for fragment A':

38.5 µl Ultrapure water

5 µl 10× Herculanse Buffer

1 µl 10 mM dNTPs

2 µl oligo 1 (25 µM)

2 µl oligo 3 (25 µM)

1 µl purified genomic DNA at 50 ng/µl (see Notes 2 and 3)

0.5 µl (2.5 U) Herculanse Hotstart DNA Polymerase

In our example, fragment A' should be 9.1 kb in length, which is confirmed by the presence of a single band on a 1% agarose gel (Fig. 2a, lane 2).

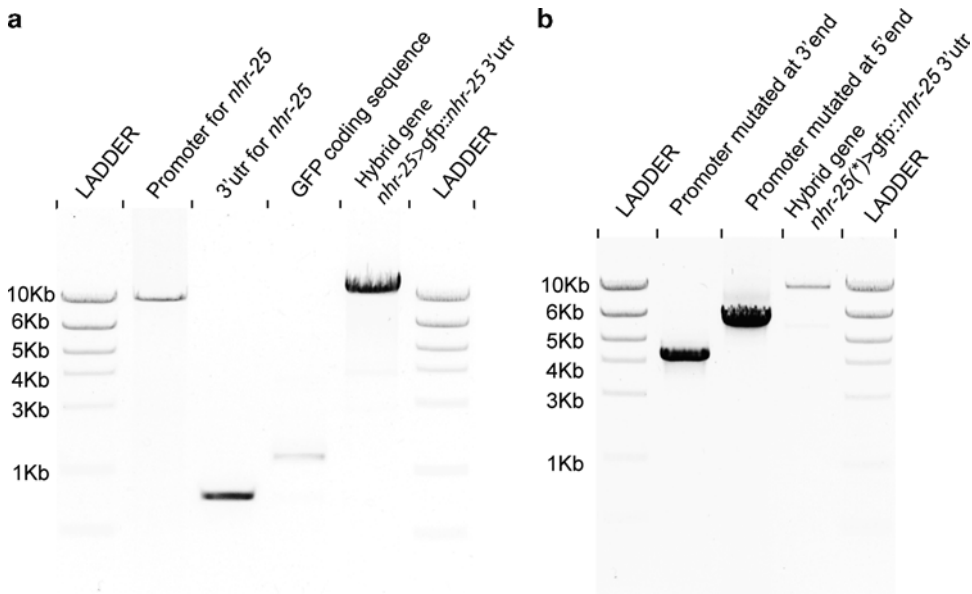


Fig. 2. 1% Agarose gels showing products of the PCR reactions to generate the fragments for OE-PCR and successful amplification of the hybrid genes *nhr-25>gfp::nhr-25 3'utr* (a) and *nhr-25(*)>gfp::nhr-25 3'utr*, which is mutated in the 5'-regulatory region (b). (a) Lanes 2–4 show the three fragments A', C', and B', and lane 5 contains the hybrid gene. (b) Lanes 2 and 3 show the two fragments for the mutated 5'-regulatory region of the gene. Lane 4 shows the final hybrid gene after amplification. The hybrid was made by combining the two mutated promoter fragments, the GFP-coding fragment and the *nhr-25 3'* UTR fragment. For this example, each fragment and the hybrid gene were column purified.

2. Reaction mix for fragment B':

39.4 μ l Ultrapure water

5 μ l 10 \times Herculase Buffer

1 μ l 10 mM dNTPs

2 μ l oligo 4 (25 μ M)

2 μ l oligo 5 (25 μ M)

0.1 μ l pPD95.75 at 200 ng/ μ l

0.5 μ l Herculase Hotstart DNA Polymerase

Fragment B', in our example, is the coding sequence for GFP, and is approx. 1.1 kb (Fig. 2a, lane 4).

3. Reaction mix for fragment C':

38.5 μ l Ultrapure water

5 μ l 10 \times Herculase Buffer

1 μ l 10 mM dNTPs

2 μ l oligo 6 (25 μ M)

2 μ l oligo 7 (25 μ M)

1 μ l purified genomic DNA

0.5 μ l Herculase Hotstart DNA Polymerase

The 3'UTR, which is '420 bp, is fragment C', and is verified on the gel (Fig. 2a, lane 3).

4. Thermocycler settings for all reactions:

Initial denaturation	92°C, 1 min
Amplification (37 cycles)	92°C, 10 s (denaturation) T_m -5°C, 30 s (annealing) (Note: for our example, we used an annealing temperature of 59°C) 68°C, 1 min/kb of fragment (extension)
Final extension	72°C, 1 min/kb of fragment + 5 min

- 2 µl of the PCR product should be run on a 1% agarose gel to check whether more than one fragment was made.
- Fragments can be purified using the QIAquick PCR purification kit and the clean products eluted in 50 µl of elution buffer. Go to Subheading 3.3.
- If there is evidence of contaminating bands, the fragment should be gel purified (see Note 4).

**3.3. Step 1 of the
Overlap PCR Reaction:
Overlap-Extension**

The next step of the protocol is to combine the purified fragments in a PCR reaction that does not contain oligos for priming the reaction. The priming in this reaction will be facilitated by the overlapping ends of the fragments (Fig. 1c). In order for OE-PCR to work optimally, it is crucial that the fragments are combined at equimolar ratios (9).

- Concentrations of each fragment must be determined. We use the Nanodrop 2000c, which determines concentrations in ng/µl. For our example, the concentrations of fragments A', B', and C' are 65, 43, and 81 ng/µl, respectively.
- The concentration of each fragment must then be converted to pmol/µl. Conversion tools are available online. Our lab uses the eBioInfogen Biotools micrograms to picomole calculator, www.ebioinfogen.com/biotools/micrograms-picomoles.htm. In our example, the concentrations of fragments A', B', and C' are 0.011, 0.059, and 0.292 pmol/µl, respectively.
- An equimolar amount of each fragment is combined in the reaction mix for a total concentration of 0.1 pmol for each fragment (the volume of each fragment solution required is calculated as 0.1/µl). The reaction mix (50 µl total volume) is as follows:
to 50 µl with ultrapure water
5 µl 10× Herculanse Buffer
1 µl 10 mM dNTPs
0.1 pmol/µl of each fragment
0.5 µl Herculanse Hotstart DNA polymerase

- Using our example, we would need the following reaction mix:
- 32.36 µl ultrapure water
 - 5 µl 10× Herculase Buffer
 - 1 µl 10 mM dNTPs
 - 9.1 µl fragment A
 - 1.7 µl fragment B
 - 0.34 µl fragment C
 - 0.5 µl Herculase Hotstart DNA polymerase
4. Thermocycler settings for this reaction is as follows:

Initial denaturation	92°C, 1 min
Amplification (13 cycles)	92°C, 10 s (denaturation) 60°C, 1 min (annealing of fragments) 68°C, 1 min/kb of final hybrid gene size (extension)
Final elongation	72°C, 1 min/kb of final hybrid gene size + 5 min

- In our example, the hybrid gene should be 9.1 kb + 1.1 kb + 420 bp, or approximately 10.6 kb in length. Thus, we used extension times of 11 min.
5. The products of this PCR reaction should not be visualized on a gel but should be purified directly using the QIAquick PCR purification kit, eluted in 30 µL of elution buffer.

**3.4. Step 2 of the
Overlap PCR:
Amplification and
Purification of the
Final Product**

The final step of the OE-PCR reaction is to amplify the desired hybrid genes. This is accomplished by using the purified product from step 1 as a template and the nested primers, oligos 2 and 8 (Fig. 1c). Using the nested primers instead of primers 1 and 7 ensures that only the desired fusion product is amplified.

- Reaction Mix for Step 2:
- 29.5 µl ultrapure water
 - 5 µl 10× polymerase buffer
 - 1 µl 10 mM dNTPs
 - 2 µl oligo 2 (25 µM)
 - 2 µl oligo 8 (25 µM)
 - 10 µl purified reaction products from step 1 above
 - 0.5 µl DNA polymerase (We usually use Herculase for this amplification, as in the other PCR reactions described above. However, if the final product is to be cloned into a TA-cloning vector, an enzyme that adds an adenosine residue should be used.)
- The final hybrid gene in our example is 10.6 kb in length (Fig. 2a, lane 5).

3.5. Using OE-PCR for Site-Directed Mutagenesis

OE-PCR can be used as a quick and efficient means to introduce mutations to specific positions of the hybrid gene (10, 11). The protocol for this application is the same as described above, except for modifications to the oligo design (Fig. 3a). As an example, we use the same hybrid gene described above, but we incorporate a mutation into the 5' regulatory region of *nhr-25*. This requires the initial A' fragment to be amplified as two fragments, A* and A**.

1. Mutations in the hybrid gene product are introduced using the primers. In our example, the 5' tails of oligos 3a (antisense) and 3b (sense) carry the mutation (Fig. 3a, b). Oligo 3a and oligo 1 are the primer pair used to amplify fragment A*. The 3' portion of oligo 3a that binds to the template should be directly adjacent to the sequence to be mutated and 21–30 bp in length. The 5' tail of oligo 3a carries the mutation (but in this case is only noncomplementary to the template in the region of the mutation but not the surrounding sequences; Fig. 3b). The 5' tail of oligo 3a should be as long as possible to increase the chances of overlap.
2. Oligo 3b (sense) and oligo 3 (antisense) are used to amplify fragment A**. Oligo 3b is essentially the reverse complement

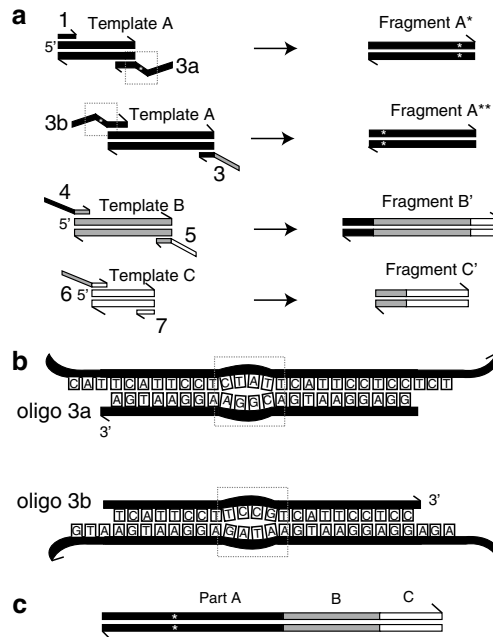


Fig. 3. Site-directed mutagenesis using OE-PCR. (a) Fragment A* is constructed using oligos 1 and 3a and fragment A** is made with oligos 3b and 3. Fragments B' and C' are made as described in Fig. 1. (b) Oligos 3a and 3b are complementary to the template but introduce a mutation (indicated as an asterisk). In this example, primers 3a and 3b are reverse complements of each other. (c) The result of the overlap reaction is a hybrid gene with a mutation in part A.

of oligo 3a (Fig. 27.3b). However, the 3' portion of the oligo that binds to the template should be 21–30 bp, and the 5' tail should be as long as possible.

3. Reaction mix for amplification of fragment A*:

38.5 µl Ultrapure water
5 µl 10× Herculase Buffer
1 µl 10 mM dNTPs
2 µl oligo 1 (25 µM)
2 µl oligo 3a (25 µM)
1 µl purified genomic DNA
0.5 µl Herculase Hotstart DNA Polymerase

In our example, fragment A* is 3.6 kb in length (Fig. 2b, lane 2).

4. Reaction mix for amplification of fragment A**:

38.5 µl Ultrapure water
5 µl 10× Herculase Buffer
1 µl 10 mM dNTPs
2 µl oligo 3b (25 µM)
2 µl oligo 3 (25 µM)
1 µl purified genomic DNA
0.5 µl Herculase Hotstart DNA Polymerase

This results in the amplification of fragment A**, which in our example is 5.5 kb (Fig. 2b, lane 3).

5. See Subheadings 3.2, steps 2 and 3 for the amplification of fragments B' and C', respectively.

6. Thermocycler settings are the same as above, Subheading 3.2, step 4.

7. For purification of fragments, see Subheading 3.2 steps 5–7.

8. OE-PCR is carried out in the same way as described above, Subheadings 3.3 and 3.4.

In our example, we combine four fragments, A*, A**, B', and C' to make a 10.6-kb hybrid gene which carries a mutation in a specific region of the 5'-regulatory region (Fig. 3c).

4. Notes

1. Oligos 2 and 8 can also be designed with noncomplementary tails. Any restriction enzyme site can be incorporated for cloning into a circular plasmid for propagation of the hybrid gene in a bacterial strain. If such a site is to be engineered, it is often important to include additional nucleotides beyond the

restriction recognition site to allow the enzyme to cut the DNA. Manufacturers of restriction enzymes (e.g., New England Biolabs) generally provide information about how many or what additional nucleotides are required.

2. Cell lysates can be used in place of isolated genomic DNA as a template in the amplification of the initial fragments.
3. It is best to not use the same plasmid for generating the fragments for both ends of the hybrid gene. In such a case, the nested primers for amplifying the overlap product will also bind to the plasmid. Even small amounts of plasmid carried over from the initial PCR reactions will preferentially act as template in the amplification reaction instead of the hybrid gene. To eliminate this problem, gel purification can be used at all steps of the protocol.
4. An alternative to gel purifying the fragments, which we have found to work equally well, is to skip gel purification and purify the PCR product directly with the QIAquick PCR purification kit (Qiagen, Valencia, CA), leaving the contaminating fragments in with the desired fragments. The OE-PCR steps can be carried out with the contaminating fragments, resulting in a final desired product with multiple undesirable byproducts. Only one gel purification step is then required at the end.

References

1. Chalfie M, Tu Y, Euskirchen G et al (1994) Green fluorescent protein as a marker for gene expression. *Science* **263**:802–805
2. Crocker J, Tamori Y, Erives A (2008) Evolution acts on enhancer organization to fine-tune gradient threshold readouts. *PLoS Biol* **6**:e263
3. Hamamatsu N, Aita T, Nomiya Y et al (2005) Biased mutation-assembling: an efficient method for rapid directed evolution through simultaneous mutation accumulation. *Protein Eng Des Sel* **18**:265–271
4. Mullis K, Faloona F, Scharf S et al (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* **51**:263–273
5. Horton RM, Hunt HD, Ho SN et al (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**:61–68
6. Yon J, Fried M (1989) Precise gene fusion by PCR. *Nucleic Acids Res* **17**:4895
7. Yelov AA, Shabarova ZA (1990) Constructing DNA by polymerase recombination. *Nucleic Acids Res* **18**:3983–3986
8. Lundberg KS, Shoemaker DD, Adams MW et al (1991) High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* **108**:1–6
9. Shevchuk NA, Bryksin AV, Nusinovich YA et al (2004) Construction of long DNA molecules using long PCR-based fusion of several fragments simultaneously. *Nucleic Acids Res* **32**:e19
10. Heckman KL, Pease LR (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc* **2**:924–932
11. Ho SN, Hunt HD, Horton RM et al (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**:51–59
12. Fire A, Harrison SW, Dixon D (1990) A modular set of lacZ fusion vectors for studying gene expression in *Caenorhabditis elegans*. *Gene* **93**:189–198