

BioTechniques 30th Anniversary Gem

Gene Splicing by Overlap Extension: Tailor-Made Genes Using the Polymerase Chain Reaction

Robert M. Horton, Zeling Cai, Steffan N. Ho, and Larry R. Pease
Mayo Clinic

Reprinted from BioTechniques 8(5):528-535 (November 1990)

Gene Splicing by Overlap Extension or “gene SOEing” is a PCR-based method of recombining DNA sequences without reliance on restriction sites and of directly generating mutated DNA fragments in vitro. By modifying the sequences incorporated into the 5'-ends of the primers, any pair of polymerase chain reaction products can be made to share a common sequence at one end. Under polymerase chain reaction conditions, the common sequence allows strands from two different fragments to hybridize to one another, forming an overlap. Extension of this overlap by DNA polymerase yields a recombinant molecule. This powerful and technically simple approach offers many advantages over conventional approaches for manipulating gene sequences.

The polymerase chain reaction (PCR) (10) has quickly become a fundamental analytical tool in molecular biology. Recently, several laboratories, including our own, have begun to use this technology as a synthetic tool for recombining DNA sequences without relying on restriction sites. Here we describe the technique and its uses and review briefly some related PCR applications.

Background

In PCR, primer extension by DNA polymerase is used to make a copy of a DNA strand which can then serve as a template for extension from a second primer in the opposite orientation. Multiple rounds of this process lead to exponential accumulation of the sequence of interest. The result of a PCR is a DNA segment of defined length which has incorporated synthetic oligonucleotide primers into its ends. The 3'-ends of these primers must match the sequence of the template gene well enough to act as primers for DNA polymerase, but the 5'-ends can include sequences unrelated to the template gene. This capability, called “mispriming” by Mullis et al. (10), has been used to perform site-directed mutagenesis (2,8) and to add sequences to the end of a PCR generated fragment (11,14). The limitation of this technique is that the mutation or added sequence must be in the primer, and thus must be within the length of an oligonucleotide from the end of the PCR fragment. If the fragment is to be cloned,

restriction sites would normally have to be included in the primers as well. This means that this method can only be used to make changes at positions close enough to the restriction sites used for cloning to be included in the same oligonucleotide.

Overlap Extension

Higuchi et al. (5), inspired by Mullis et al. (11), were the first to report the use of a technique to introduce mutations into the center of a PCR fragment making PCR mutagenesis much more flexible. Ho et al. (6), who independently developed the technique and refer to it as “overlap extension,” also demonstrated that the error frequency intrinsic to this method is sufficiently low to make it practical for widespread use. A modification of this method (7) allows segments from two different genes to be recombined or “spliced” together by overlap extension, a process we refer to as “gene Splicing by Overlap Extension” (SOE) or “gene SOEing.”

The mechanism is illustrated in Figure 1. PCR is used to generate two fragments (AB and CD) which have their ends modified by mispriming so that they share a region of homology. When these two fragments are mixed, denatured and reannealed, the 3'-end of the top strand of fragment AB anneals onto the 3'-end of the bottom strand of fragment CD, and this overlap can be extended to form the recombinant product. The overlap region is determined by primers ‘b’ and ‘c’

and can contain any sequence, as long as the oligomers are complementary. This region is where base changes are incorporated when the technique is used for site-directed mutagenesis. Alternatively, the overlap can be designed to make a “neat” joint between two fragments, with no new sequences included, as in the present report.

In this report we describe the construction of a recombinant gene encoding a chimeric protein in which parts of a class I major histocompatibility complex (MHC) antigen are replaced by corresponding portions of a distantly related class II MHC antigen.

Materials and methods

Primers

Oligonucleotide primers were synthesized on an Applied Biosystems Model 380A DNA synthesizer (Foster City, CA) and desalting on a Sephadex G-50 column (Pharmacia LKB Biotechnology, Piscataway, NJ).

Templates

The class I sequences were derived from a plasmid containing K^b (13), and the class II sequences were derived from plasmids containing A _{α} ^k (1) and A _{β} ^k (3).

Primer Design

The sequences of the eight primers used for this recombination are given in Table 1. Primers ‘a’ and ‘h’ are the flanking or “outside” primers, which serve to PCR amplify the final

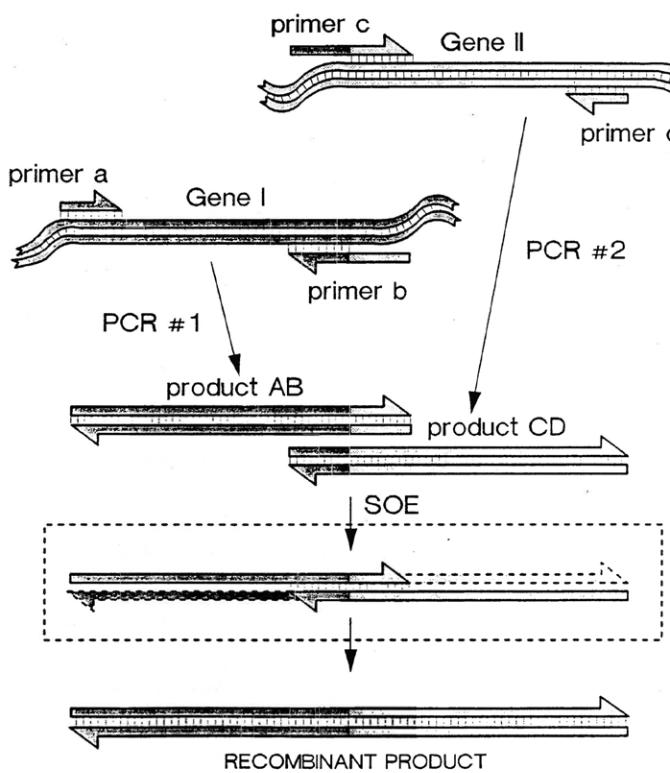


Figure 1. An illustration of the concept of SOEing genes. Product AB is PCR amplified from gene I, and fragment CD is amplified from gene II. Primers 'b' and 'c' are the SOEing primers, and they are used to modify the end of the two PCR products so that they have the same sequence. When these PCR products are mixed, denatured and reannealed under PCR conditions, the top strand of AB and the bottom strand of CD overlap and act as primers on one another, as shown inside the broken rectangle. The recombinant product is formed when this overlap is extended by polymerase. Inclusion of the outside primers 'a' and 'd' in the SOE reaction causes the recombinant product to be PCR amplified right after it is formed.

recombinant product. They do not contribute to the sequences added at the overlapping ends. Oligomers 'b' and 'c', 'd' and 'e', and 'f' and 'g' are the SOEing primers. The members of each pair are related because bases have been added to their 5'-ends to make them complementary to one another. In each case, the overlap region between the primers, and the priming region by which each primer recognizes its template, was designed to have an estimated T_d of approximately 50°C according to the formula $T_d = 4(C+G) + 2(A+T)$ in degrees Celsius (15). In practice, we have found that simply making these regions 15 to 16 nucleotides long generally works well. We have not made a careful examination of the minimum length of the oligomers.

In the primers shown in Table 1, all of the complementary bases have been added to one of the two primers (primers 'b', 'e' and 'g'), rather than adding some sequence to each primer. This way, the other primers ('c', 'd' and 'f') can potentially be used with new SOEing primers (analogous to 'b', 'd' and 'g') to join these fragments to other genes. Since the two templates share three nucleotides in primers 'f' and 'g', these nucleotides contribute both to the overlap and to the priming portion of oligomer 'g' (they are both underlined and asterisked in Table 1). The portion of oligomer 'e' in parentheses is not related to either template, and does not contribute to the overlap. This is an example of insertional mutagenesis (6) being carried out simultaneously with recombination.

As an example of the SOEing process, the complementary regions 'd' and 'e' containing the sequences which lead to the PCR products AD and EH having overlapping ends are shown in Figure 2.

Detail of Overlapping Ends Being SOEn

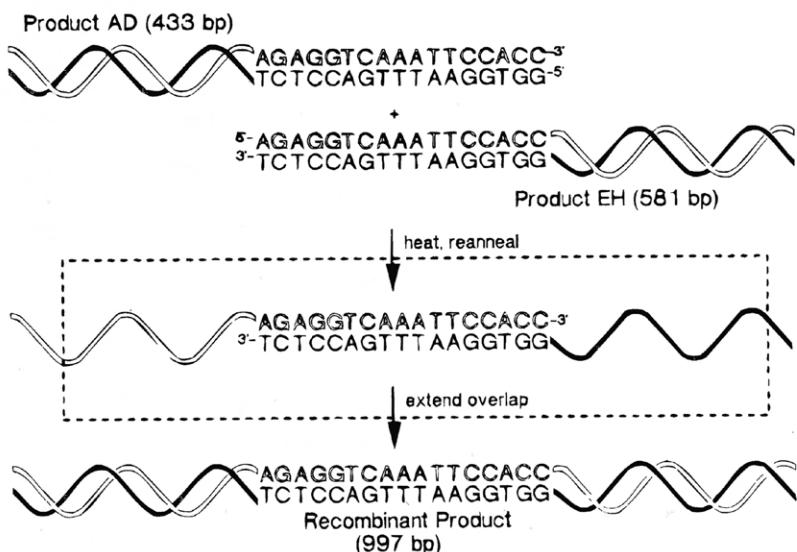


Figure 2. Detail of the ends of two fragments being SOEn. Only sequence from the overlap region shared by fragments AD and EH is shown.

Reaction Conditions

PCR and SOE reactions were carried out in a thermocycler (Perkin-Elmer Cetus, Norwalk, CT) for 25 cycles, each consisting of 1 min at 94°C, 2 min at 50°C, and 3 min at 72°C. (The reaction probably produces all of the product in fewer than 25 cycles, but we have not examined the minimum number of cycles required. Extra cycles do not appear to cause any problems.) *Taq* polymerase was from Perkin Elmer Cetus, and the reaction buffer was as recommended by the supplier (50 mM KCl, 10 mM Tris-Cl, pH 8.3, 1.5 mM MgCl₂, 0.01% (wt/vol) gelatin). Deoxyribonucleotides were used at a final concentration of 200 μM. The buffer and deoxynucleoside triphosphates were each made as a 10x stock, and 10 μl were used per 100 μl reaction. One-half μl of polymerase (2.5 U) was used per reaction. Reactions were covered with mineral oil before thermal cycling.

Table 1. Primer Design

'a' = 5'-AGGGAAACGGCCTCTGA-3'
'b' = 3'- <u>GGGTCCGGGTGTCCTCTGCTGTA</u> ACTCCG-5'
'c' = 5'- <u>GAA</u> GACGACATTGAGGC-3'
'd' = 3'- <u>TCTCCAGTTAAGGTGG</u> -5'
'e' = 5'-AGAGGTCAAATTCCACC (AAGGGCGGCTCTCAC) <u>GTGCACCAGTTCCAGC</u> -3'
'f' = 3'- <u>AGGGACGCCGCCGA</u> -5'
'g' = 5'-TCCCTGCGGCCGCTGCGCACAGGTGC-3'
'h' = 3'-TCGGTACCGGAGAGGGT-5'

Notes: 1. Primers 'b', 'd', 'f' and 'h' are written with their 3' -ends at the left, because they prime at the 3' -ends of their fragments.
2. The regions of the oligos which allow them to act as PCR primers on the appropriate templates are underlined.
3. The complementary sequences between pairs of primers are shown with asterisks between them.

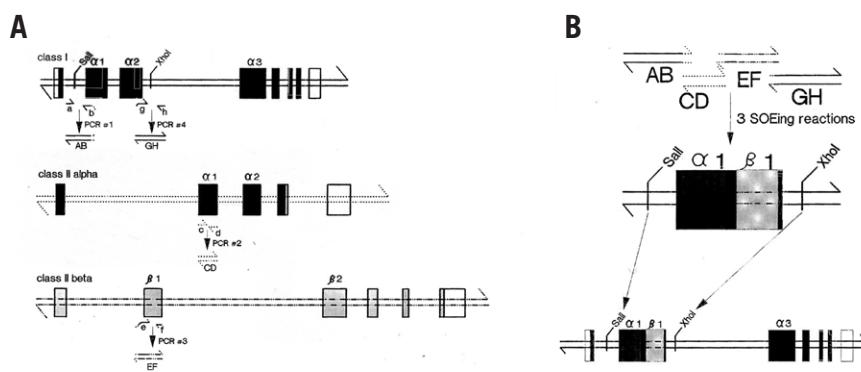


Figure 3. Strategy for creation of a chimeric class I/class II MHC gene. Part A shows the three template genes which provide portions for the chimeric molecule and the products of the first four PCRs. The DNA strands of the class I gene are shown as solid lines, with the coding exons shown as dark gray boxes. Noncoding exons are depicted as clear boxes. The class II α -chain is represented by dotted lines and medium gray boxes, and the β -chain is represented by dashes and dots with light gray boxes. Positions of oligomers (not to scale) are shown beneath the templates on which they prime and the PCR products they generate below them. Part B shows the PCR products with their overlap regions lined up and the final recombinant product. The intermediate SOEn products AD and EH are not shown. Note that the short class I sequence between $\alpha 1$ and $\beta 1$ in the recombinant was not amplified from the class I template by PCR, but was added by insertional mutagenesis. Also, the intron was deleted precisely.

Purification of Fragments

PCR and SOEn products which were to be used as templates in further reactions were purified by electrophoresis through agarose (1% SeaKem LE agarose + 2% NuSieve GTG agarose, FMC BioProducts, Rockland, ME) in TAE buffer (0.04 M Tris-acetate, 0.001 M EDTA) with 0.5 μ g/ml ethidium bromide in the gel. DNA from the appropriate bands was recovered from the gel fragment by GeneClean (Bio 101, La Jolla, CA). The final recombinant product was similarly gel-purified before cloning.

Cloning of Fragments

The SOEn products were cut with restriction enzymes *Sall* and *Xba*I and ligated into the corresponding position of a pUC-derived

plasmid which has been designed to act as an expression vector for class I MHC antigen binding regions, as described elsewhere (13).

Analysis of Products

The cloned product was sequenced from the double-stranded template using a Sequenase kit (United States Biochemical, Cleveland, OH) with a modified protocol (9).

Results

Production of Fragments

Figure 3 shows the strategy for construction of the hybrid class I/class II gene. Part A shows the original genes and the locations of the primers, while part B shows the strategy for SOEing these PCR products together.

Figure 4 is a photograph of an ethidium bromide stained gel of the products of these reactions. The PCR products AB, CD, EF and GH, as well as the intermediate SOEn products AD and EH, are the major products of the reactions, though some byproduct bands are visible. These other products were eliminated in the gel purification step. Also, some additional products are visible in the final product lane (AH), but the band of the expected size is the major product.

Analysis of Product

The sequence of the cloned product is shown in Figure 5. Of approximately 1700 bases sequenced (coding regions from two clones), there was one unplanned mutation, presumably caused by misincorporation by the polymerase. This is compatible with the error frequency of 0.06% reported earlier (7).

Discussion

Quite complicated constructions are possible with this approach. For example, we have used it to create and express a gene for a fusion protein in which the α helices are replaced by the corresponding segments of a similar gene (7). The construction used to illustrate the technique in the present report is similarly complicated. In this construction, four gene segments from three different genes were SOEn together, two at a time, to produce the chimeric product. Simultaneously, a 15-base pair (bp) segment encoding a portion of the class I gene which was in a different exon from the one that was amplified was added at the recombination joint. We have thus accomplished recombination and insertional mutagenesis in one step.

Synthetic Uses of PCR

Several modifications to the basic SOE concept are possible. A "mega-primer" approach which uses fewer primers has recently been used by several groups (14,16). Recently, Sandhu, Precup and Kline (personal communication) have demonstrated that fragments can be SOEn directly to a vector, and then recircularized by blunt-end ligation. Thus it may be possible to entirely avoid the use of restriction enzymes for the construction of recombinant genes in appropriate vectors. Different approaches to using PCR as a synthetic tool are also being developed. For example, asymmetric PCR has been used to generate what is in effect a huge mutagenic oligonucleotide for use in an otherwise standard m13-based mutagenesis strategy, in which regions as large as an exon can be replaced (4).

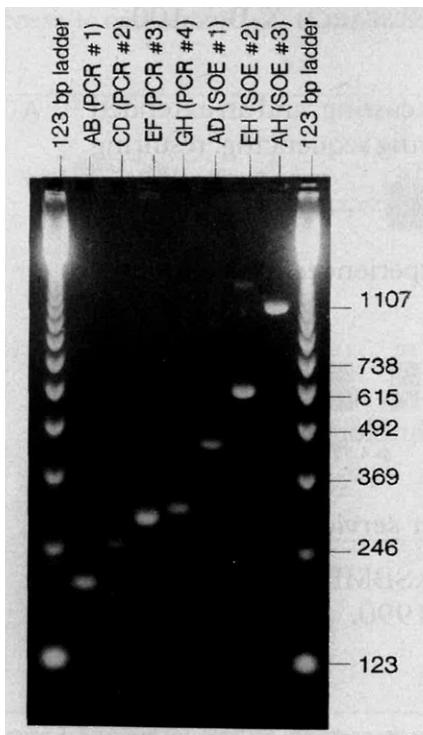


Figure 4. Ethidium bromide stained gel showing DNA fragments used in this construction. The expected sizes of the fragments are as follows: AB = 198 bp, CD = 252 bp, EF = 289 bp, GH = 306 bp, AD = 433 bp, EH = 581 bp and AH = 997 bp.

Limitations of SOE

The major drawback of the gene SOEing technique is that, even though the frequency of polymerase errors is low, it is necessary to sequence the cloned products in order to be certain that you have what you want. This limits the usefulness of gene SOEing to engineering problems which are not conveniently solved by more conventional methods. One caveat to bear in mind is that, in addition to mediating directed recombination, PCR can lead to random recombinations between related genes present in the same reaction (12). A partially elongated fragment of one gene can act as a primer on a different gene, producing a recombinant product. This can be viewed as either a hazard in attempting to amplify one gene out of a multi-gene family or a possible means of generating a set of recombinants between related molecules.

Conclusion

PCR is much more than a sensitive tool for detecting DNA sequences. Splicing by Overlap Extension, or "gene SOEing," is a novel, PCR-mediated recombinant DNA technology. It represents a significant advancement over the standard restriction enzyme based methods for recombin-

```

Primer 'a' ->
aggaaaacggcctcta-3'
AGGGAAACGGCTCTGAGGGGAGGGCGCACCGGGAAGCCGCTCTCGGTGACCAACGGGACCCCTC 70
TCCCTTGCGGAGACTCCCCCTCCCGCGTGGCCCTTCGGCGCAGGAGCCCAGCTGGTGGCTGGAG

CGCCCTTCTCACCCGAGTCCCGGCCCTGCTCCCTCTAGCCCGCGCAGCCGCCGGGTCTGGTGA 140
GCCGGGAAGAGGTGGCTCAGGGCGGGACGAGGGAGAGTCGGCGCTGGCGGGCCCAAGACACT
Primer 'c' -> 5'-gaagacacattgaggc-3'
GGTGGTCCGGGCTCACCGCGGCCCCCCCAGGCCAACACGAAAGACACATTGAGGGCACCACGTAGG 210
CCACCAAGGCCAGATGGCAGCCGGGGGTCTGGCTCTGTAACTCCGGCTGGTGCATCC
<- Primer 'b' 3'-gggtccgggtgtctctgtctgtaaactccg-5'
g p h e d d i e a d h v g

CTCCATGGTATAACTGTATATCAGTCTCTGGAGACATTGCCAGTACACATTGAATTGATGGTGT 280
GAGGATACCATATTGACATATAGTCAGAGGACCTCTGAAACCGGTATGTGTAACCTAAACTACCACTA
s y g i t v y q s p g d i g q y t f e f d g d

GAGTTGTTCTATGGACTTGGATAAGAAGGAGACTGCTGGATGCTCCCTGAGTTGCTCAACTGAGAA 350
CTCAACAAGATAACACTGAACCTATTCTCCCTGACAGACCTACGAAGGACTCAAACGAGTTGACTCTT
e l f y v d l d k k e t v w m l p e f a q l r r
Primer 'e' -> 5'-agag
GATTTGAGCCCCAAGGGACTGCAAAACATAGCTACAGGAAAACACAACACTGGAAATCTGACTAAGAG 420
CTAAACTCGGGGTTCCACCTGACGTTGTATCGATGCTTTGTGTTGAAACCTTAGAACTGATTCTC
<- Primer 'd' 3'-tctc
f e p q g g l q n i a t g k h n l e i l t k r

gtcaaattccaccaaggcggtctcacgtcaccatgttcagc-3'
GTCAAATTCCACCAAGGGCGGCTCACGTGCAACCAGTCCAGGCTCTGCTACTTACCAACGGGACCG 490
CAGTTAAAGGTGGTCCCGCCAGAGTCACGTGGTCAAGGTGGAGACGATGAAGTGGTGGCTCG
cagttaaagggtgg-5'
s n s t k g g s h v h q f q p f c y f t n g t

CAGCGCATACGGCTTGTGATCAGATACATCTACAACGGGAGGGATACGTGCGCTTCACAGCGACGTGG 560
GTCGCGTATGCCAACACTAGTCTAGTGTAGATGTTGGCCCTCTCATGCACGCGAAGCTGTCGCTGCACC
q r i r l v i r y i y n r e e y v r f d s d v g

GCGAGTACCGCGCGGTGACCGAGCTGGCGGAGACGCCAGACTGGATAAGCAGTACCTGGAGGG 630
CGCTCATGGCGCGGACTGGCTGACCCCGCCGTCAGCGCTCATGACCTTATCGTCATGGACCTCGC
e y r a v t e l g r p d a e y w n k q y l e r
Primer 'g' -> 5'-tccctgcgg
AACGCGGCCGAGCTGGACACGAGACACAACACTCGAGAAGACGGAGACCCCAACCTCCCTCGGG 700
TTGCGCCGGCTCGACCTGTGCCCCACCGTGTGTTGATGCTCTCTGCCCTGGGGCTGGAGGGACGCC
<- Primer 'f' 3'-aggacgcc
t r a e l d t v c r h n y e k t e t p t s l r

cgctcgacacagggtc-3'
CGGCTGCGCACAGGTGCAAGGGCCGGCAGCTCTCCCTCTGCCCTGGGCTGGGCTAGTCTGG 770
GCCGACCGTGTCCACGTCAGCTCCCGCCGCGCCGTCAGGAGGGAGACGGAGCCGACCCGAGATCAGGACC
gccga-5'
r l r t g

GGAAGAAGAAACCTCAGCTGGGTGATGCCCTGTCAGAGGGAGAGGTGTCGCTGGCTCCCTGAT 840
CCTCTCTTGGAGTCGACCCACTACGGGACAGAGTCCTCCCTCTCACAGCACCAGAGGACTA

CCCTCATCACAGTGAUTGACTGACTCTCCAGGGCTCACCTCTCCCTGAGAGTGCCTGGCTGTCT 910
GGGAGTAGTGTCACTGACGTGACTGAGAGGGCCCGAGTCAGGAGAGGAGCTCACGGTCCGACAGA

CAGGGAGGAAGGAGAGAATTCCCTGAGGTAAACAACAGCTGCTCCCTCAGTCCCTGAGCCTGTGTC 980
GTCCCTCCCTCTCTAAAGGGACTCATTGTTGTCAGCAGGGAGTCAGGGAGCTGGAGACAG

AGCCATGGCCTCTCCA 997
TCGGTACCGGAGAGGGT
tcggtacccggagagggt-5'
<- Primer 'h'

```

Figure 5. Sequence of the recombinant product. Positions at which the various oligomers prime are shown. Sequences derived from class II genes (Ao.k or A.k) are denoted with a solid line between strands.

genes because it does not rely on restriction sites. Thus it allows much finer control of recombination for genetic engineering. Also, the sequence of the overlap region is determined by the primer design, making it possible to perform mutagenesis and recombination simultaneously. This technically simple and rapid approach is entirely different from methods now used for generating modified and recombinant DNA fragments, and we believe it marks the beginning of a new generation of recombinant DNA technology.

Acknowledgments

R.M.H and Z.C. were supported by predoctoral fellowships from the Mayo Graduate School, Mayo Foundation, and S.N.H. was supported by a Medical Scientist Scholarship from the Life and Health Insurance Research Fund. We would like to thank Drs. Henry Hunt and Jeffrey Pullen for sharing their gene SOEing experiences with us.

References

- Benoist, C.O., D.J. Mathis, M.R. Kanter, V.E. Williams II and H.O. McDevitt. 1983. Regions of allelic hypervariability in the murine Aa. immune response gene. *Cell* 34:169-177.
- Dulau, L., A. Cheyrou and M. Aigle. 1989. Directed mutagenesis using PCR. *Nucleic Acids Res.* 17:2873.
- Estess, P., A.B. Begovich, M. Koo, P.P. Jones and H.O. McDevitt. 1986. Sequence analysis and structure-function correlations of murine q, k, u, s, and f haplotype I-A cDNA clones. *Proc. Natl. Acad. Sci. USA* 83:3594-3598.
- Feinstone, S., C. Wychowski, S. Emerson and J. Silver. 1989. Synthesis of chimeric hepatitis A virus/poliovirus subgenomic cDNA by a PCR mutagenesis system. *J. Cell. Biol. Supplement* 13E:283 (abstract #WH130).
- Higuchi, R., B. Krummel and R.K. Saiki. 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* 15:7351-7367.
- Ho, S.N., H.D. Hunt, R.M. Horton, J.K. Pullen and L.R. Pease. 1989. Site-directed mutagenesis by overlap extension using the Polymerase Chain Reaction. *Gene* 77:51-59.
- Horton, R.M., H.D. Hunt, S.N. Ho, J.K. Pullen and L.R. Pease. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77:61-68.
- Kadowaki, H., T. Kadowaki, F.E. Wondifossford and S.I. Taylor. 1989. Use of Polymerase Chain Reaction catalyzed by *Taq* DNA polymerase for site-specific mutagenesis. *Gene* 76:161-166.
- Kraft, T., J. Tardiff, K.S. Krauter and L.A. Leinwand. 1988. Using mini-prep plasmid DNA for sequencing double stranded templates with Sequenase™. *BioTechniques* 6:544-547.
- Mullis, K. and F. Falloona. 1987. Specific synthesis of DNA in vitro via a polymerase
- catalyzed Chain Reaction. *Methods Enzymol.* 155:335-350.
- Ilis, K., F. Falloona, S. Scharf, R. Saiki, G. Horn and H. Erlich. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spr. Harb. Symp. Quant. Biol.* L1:263-273.
- Paabo, S., R.G. Higuchi and A.C. Wilson. 1989. Ancient DNA and the polymerase chain reaction: the emerging field of molecular archaeology. *J. Biol. Chem.* 264:9709-9712.
- Pullen, J.K., H.D. Hunt, R.M. Horton and L.R. Pease. 1989. The functional significance of two amino acid polymorphisms in the antigen-presenting domain of class I MHC molecules: molecular dissection of K^bm³. *J. Immunol.* 143:1674-1679.
- Sarkar, G. and S.S. Sommer. 1990. The "megaprimer" method of site-directed mutagenesis. *BioTechniques* 8:404-407.
- Suggs, S.V., T. Hirose, T. Miyake, E.H. Kawashima, M.J. Johnson, K. Itakura and R.B. Wallace. 1981. Use of synthetic oligodeoxyribonucleotides for the isolation of specific cloned DNA sequences, p. 683-693. In D.D. Brown and C.F. Fox (Eds.) *Developmental Biology Using Purified Genes*. Academic Press, New York.
- Yon, J. and M. Fried. 1989. Precise gene fusion by PCR. *Nucleic Acids Res.* 17:4895.

To purchase reprints of this article, contact:
biotechniques@fosterprinting.com

Illuminating Solutions For Your Life Science Research!

Life science and laboratory research professionals depend on high-quality Spectroline® ultraviolet products to make their work faster and easier. Our products are built to exacting standards, ensuring years of reliable, trouble-free service.



UV Crosslinkers

Provide accurate, error-free and super-fast membrane crosslinking and sample UV dosing.



MiniMAX™ UV Viewing Station

Space-saving mini-cabinet and UV lamp combination makes the perfect portable UV viewing station. Light weight and compact.



UV-Protective Eye and Face Wear

Specially designed to protect users against unwanted UV exposure.



Miniature UV Pencil Lamps

Double-bore quartz lamps used for wavelength calibration and emission. Custom lamps for OEM applications also available.



UV Transilluminators

Deliver superior irradiance uniformity and UV intensity for all viewing and photodocumentation needs.



UV Bench and Display Lamps

Lamps with various wattages and UV wavelengths provide both high intensity and wide area of light coverage.



AccuMAX™ Digital Meters

Advanced microprocessor-controlled units with software-driven functions. Ensure maximum accuracy and reliability in measurement of UV and visible light sources.